Regional Ecosystem Simulation: Combining Surface- and Satellite-Based Observations to Study Linkages between Terrestrial Energy and Mass Budgets

by

Peter Edmond Thornton

B.A., The Johns Hopkins University, 1990

M.A., The Johns Hopkins University, 1992

presented in partial fulfillment of the requirements

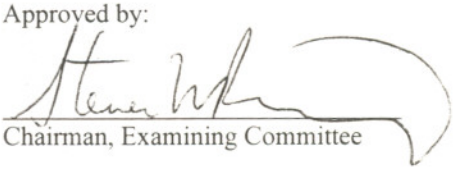
for the degree of

Doctor of Philosophy

The University of Montana

1998

Approved by:



Chairman, Examining Committee

_____

Dean, Graduate School

_____

Date

Thornton, Peter E., Ph.D., May 1998 Forestry

Regional Ecosystem Simulation: Combining Surface- and Satellite-Based Observations to Study Linkages between Terrestrial Energy and Mass Budgets (280 pp.)

Director: Steven W. Running SWR

The dissertation is presented as three separate studies. The first study presents the development and application of a set of methods for estimating near-surface meteorological parameters in complex terrain, given an arbitrarily-spaced network of meteorological measurement stations. Predicted parameters are daily maximum and minimum air temperature, daily total precipitation, daily shortwave radiation, and daily average water vapor pressure. Required observations are daily maximum and minimum air temperature, and daily total precipitation. Provisions are included for missing data, and for spatial and temporal variation in station density. Estimation proceeds in two steps: an interpolation between stations, followed by an extrapolation to account for differences in elevation between station and estimation points. Interpolation is based on a truncated Gaussian convolution filter, the radius of which varies in response to local station density. Parameters for the shape and size of the filter are determined through a cross-validation procedure. Extrapolation is based on a diagnosis of the local relationships between temperature, or precipitation, and elevation. Errors were about 2.0 and 1.0 °C, for daily and annual average temperature predictions, respectively. Annual total precipitation errors were about 20%.

The second study describes the development and testing of a numerical model for simulating terrestrial ecosystem processes. The model focuses on primary production, and includes treatments of surface water and shortwave radiation budgets, belowground carbon and nitrogen dynamics, and seasonal and interannual allocation of primary production within plants.

In the third study, meteorological parameters, derived in part by the methods of the first study, are combined with satellite-based observations of visible, near-infrared, and thermal infrared radiances, to produce a mapped estimate of the surface resistance to sensible heat transport. This resistance (r_h) is an important parameter for the surface energy budget calculations described in the second study, and has previously been poorly defined. Predictions from the surface energy budget algorithm were supported by evidence from the satellite and surface-based temperature observations. In particular, there was a strong negative relationship between vegetation cover and predicted r_h. Different vegetation types exhibited different slopes in this relationship, which could be explained by broad differences in leaf morphology and stand structure.

Acknowledgements

It has been my great fortune, since arriving in Missoula six years ago, to be in the company of people full of encouragement, good will, and happy lunacy. My everlasting gratitude goes out to Dr. Steve Running, for providing me with both the opportunity to experience new worlds of knowledge and the resources to explore them. I have always been able to rely on his guidance and generosity with time, and his confidence in my abilities was strong when my own was flagging. The ground I covered in my studies has seen other explorers: I extend a salute and my heartfelt thanks to Dr. Ramakrishna Nemani for blazing some of the more exciting trails. My warmest regards go out to Dr. Joseph White, who, although we are presently antipodal, has been my closest consort on this journey. Exemplary of character and light with mirth, his friendship has more than once given me bravery in the wilderness: may we meet again soon over a yard of the finest. For the years of harmonic cohabitation, for the hours of chess over the belly of the demon, and for losing track of how many games of pool he won, I thank Dr. Robert Kremer. Thanks to Mike White for subjecting himself to my untested code, and for sharing with me the fruits of his implacable curiosity. Thanks to Dr. Lars Pierce, Galina Churkina, Dr. Kathy Hibbard, Alisa Keyser, and Dr. John Kimball, for, among other kindnesses, poring over the details of various versions of the models described here, tracking down bugs, and suggesting many improvements. Many thanks to Joe Glassy, for fielding my endless programming questions. For the very best in technical and administrative support, thanks to Saxon Holbrook, Youngee Cho, Deb Kendall, Hal Dorsman, and Kim Hodgeson. Very special thanks are due the other members of my committee, Drs. LLoyd Queen, Bob Keane, and Hal Salwasser, for their many useful suggestions and helpful comments. To my parents, Helen and Kirby Thornton, I offer my deepest love and appreciation, for setting me out with a heart full of riches on such a rewarding path. Thanks to Chris and Matt for their lifelong encouragement, brothers in spirit as well as body; and to Polly, thank you for your songs and your love, and for showing me the glory of nature where I'd missed it before. Finally, to my wife Michele, without whom the beauty of these mountains would shine less brightly, thank you for unfailing faith, unwavering love and devotion, for hearing and helping me through my daily struggles, and for giving me joy and humor when I needed them most.

Some of the data used here was collected by researchers at the Forest Service Intermountain Fire Sciences Laboratory, and by members of the science staff at Glacier National Park. Thanks to Jim Menakis, Cameron Johnston, and Marty Beck of the Forest Service, and Carl Key at Glacier NP for their efforts on my behalf. Thanks also to Drs. Sue Ferguson and Miriam Peterson of the Forest Service Pacific Northwest research lab for helpful discussions and access to their computer resources.

I received financial support during my studies from NASA's Office of the Mission to Planet Earth, in the form of a Graduate Student Fellowship in Global Change Research, for which I am most grateful.

Table of Contents

Abstract ii
Acknowledgments iii
Table of Contents iv
List of Tables vi
List of Figures vii

Chapter 1
Introduction 1
1. Statement of research problems 1
2. Review of research problem investigated in Chapter 2 1
3. Review of research problem investigated in Chapter 3 3
4. Review of research problem investigated in Chapter 4 5

Chapter 2
Generating Surfaces of Daily Meteorological Variables over Large Regions of Complex Terrain 6
1. Background and model development 6
2. Methods 8
2.1 Interpolation 8
2.2 Temperature predictions 10
2.3 Precipitation predictions 12
2.4 Extrapolation smoothing 13
2.5 Humidity predictions 14
2.6 Radiation predictions 15
2.7 Case-study description and database filtering 16
2.8 Parameterization and validation 20
2.9 Spatial scaling analysis 21
2.10 Temporal scaling analysis 21
3. Results 22
3.1 Parameterization 22
3.2 Validation 27
3.3 Spatial and temporal scaling analysis 37
3.4 Example output 40
4. Discussion 40
5. Literature cited 52

Chapter 3
Description of a numerical simulation model for predicting the dynamics of energy, water, carbon, and nitrogen in a terrestrial ecosystem 56
1. Introduction 56
2. Overview of model structure and components 57
3. Detailed description of shortwave radiation budget 61
3.1 Shortwave radiation inputs 61
3.2 Canopy radiation absorption and fractionation between sunlit and shaded portions 62
4. Detailed description of the water budget 66
4.1 Soil water potential and volumetric water content 66
4.2 Precipitation input and canopy interception 68
4.3 Evaporation of canopy intercepted water 69
4.4 Losses from the snowpack 71
4.5 Evaporation from the litter/soil surface 72

4.6 Transpiration and controls on stomatal conductance 73
4.7 Soil water balance and hydrologic outflow estimation 79
5. Detailed description of the carbon and nitrogen budgets 80
5.1 Photosynthetic carbon fixation and canopy nitrogen distribution 81
5.2 Maintenance respiration 91
5.3 Growth respiration 91
5.4 Leaf and fine root phenology 92
5.5 Whole-plant mortality 94
5.6 Litter and soil carbon and nitrogen budgets 95
5.7 Carbon and nitrogen allocation 118
6. Example simulations 128
7. Literature cited 150

Chapter 4
Estimation of surface resistance parameters from remote sensing and surface meteorological observations 167
1. Introduction 167
2. Review of theory of mass and energy storage and transport 167
2.1 Radiation budget 167
2.2 Energy balance 169
3. Review of relevant literature 173
3.1 Early investigations, pre-1970 174
3.2 Current literature, 1970-present 176
3.2.1 Reflectance data: surface resistance and flux estimation 176
3.2.2 Thermal infrared emission data: energy balance studies 178
3.2.3 Thermal infrared emission data: empirical relationships 182
3.2.4 Combined thermal infrared and reflectance data 185
3.3 Summary of literature review with respect to current research problem 190
4. Operational statement of the research problem 192
5. Methods 198
5.1 Study area and topographic data layers 198
5.2 Estimation of daily surface meteorological parameters 200
5.2.1 Generation of Daymet input database 200
5.2.2 Parameterization of Daymet for the study region and period 201
5.2.3 Daymet simulations and summary of results 208
5.3 Estimation of subdaily meteorological parameters 211
5.3.1 Subdaily air temperature estimates 211
5.3.2 Subdaily radiation estimates 226
5.3.3 Subdaily humidity estimates 227
5.4 AVHRR database generation 227
5.4.1 Cloud and snow detection and screening 229
5.4.2 Satellite zenith angle screening 232
5.4.3 Estimation of time of day of satellite data acquisition 235
5.4.4 Estimation of sun-sensor angle 239
5.5 Estimation of land surface temperature using a split-window algorithm 239
5.6 Prediction of r_h 246
5.7 Evaluation of predicted r_h 248
6. Results 251
6.1 Multiple regression analysis 251
6.2 Predicted r_h: variation with canopy closure and NDVI 254
7. Summary of results and discussion of applications 267
8. Literature cited 270

List of Tables

Chapter 2

Table 1: Numbers of meteorological observation stations 18
Table 2: List of model parameters 27
Table 3: Cross-validation error statistics 27
Table 4: Summary of daily precipitation occurrence predictions vs. observations 33

Chapter 3

Table 1: Estimates of k_{rain} derived from other studies 69
Table 2: Annual mortality fraction from data in Sollins (1982) 95
Table 3: Regression results for ^{14}C decomposition experiments 101
Table 4: Results of numerical experiments for tests of decomposition models 103

Chapter 4

Table 1: Results of t-tests comparing constant with temporally-varying parameterization of GSP.. 204
Table 2: Final cross-validation analysis estimates of error and bias 206
Table 3: Annual and 5-year mean statistics for Daymet output 209
Table 4: Station data for 13 stations used in assessment of subdaily air temperature 213
Table 5: Regression coefficients for the prediction of $NTC14_h$ 219
Table 6: Mean absolute prediction errors (°C) for subdaily temperature 221
Table 7: Cloud detection criteria 232
Table 8: Frequencies of landcover and canopy closure classes 248
Table 9: Landcover types reclassed from the Loveland classification 251
Table 10: Multiple regression results for ΔT vs. (R_n, NDVI, VPD, P_{pre}, θ) 253
Table 11: Predicted r_h and observed NDVI, average by covertype 255
Table 12: Comparison of predicted r_h against independently determined canopy closure classes.... 260

List of Figures

Chapter 2

Figure 1: Location of study region: Northwest U.S. 17
Figure 2a: TMAX: mean absolute error for daily predictions (°C) 23
Figure 2b: PRCP: mean absolute error (% of annual total) 24
Figure 3: PRCP: mean absolute error vs. f_{max} 26
Figure 4a: TMAX: predicted vs. observed daily values 29
Figure 4b: TMIN: predicted vs. observed daily values 30
Figure 4c: TMAX: predicted vs. observed event frequency 31
Figure 4d: TMIN: predicted vs. observed event frequency 32
Figure 5a: PRCP: predicted vs. observed annual total 34
Figure 5b: PRCP: predicted vs. observed event frequency 35
Figure 6: Spatial frequency of dry stations: predicted vs. observed 36
Figure 7: Successful PO predictions vs. frequency of dry stations 38
Figure 8: Daily average PRCP vs. proportion of wet stations (log:log) 39
Figure 9: Annual vs. daily predictions of annual total PRCP 41
Figure 10a: Temperature vs. elevation regression slopes 42
Figure 10b: Precipitation vs. elevation regression slop 43
Figure 11a: Detail of study region (color plate) 44
Figure 11b: Annual average daytime temperature (color plate) 45
Figure 11c: Annual average daytime VPD (color plate) 46
Figure 11d: Annual total of wet days (color plate) 47
Figure 11e: Annual total precipitation (color plate) 48
Figure 11f: Annual average of daylight average incident shortwave radiative flux density 49

Chapter 3

Figure 1a: Stomatal conductance response to daylight average PPFD 76
Figure 1b: Stomatal conductance response to daylight average temperature 77
Figure 2a: 200-year total NPP vs. SLA ratio (no daily allocation) 86
Figure 2b: 200-year total NPP vs. SLA ratio (with daily allocation) 87
Figure 3: Decomposition model structures 98
Figure 4: Homogeneous substrate decomposition data fitted to M3 model 106
Figure 5: Litter and soil carbon dynamics for 1D-BGC 108
Figure 6a: ^{14}C glucose decomposition data 109
Figure 6b: ^{14}C cellulose decomposition data 110
Figure 6c: ^{14}C lignin decomposition data 111
Figure 7: Primary linkages between C and N in the plant-litter-soil system 116
Figure 8: Stem C vs. average stand height for unmanaged stands 126
Figure 9a: Simulated leaf C under severe N-limitation 129
Figure 9b: Simulated soil mineral N under severe N-limitation 130
Figure 10a: Simulated leaf C during and after N-limitation 132
Figure 10b: Simulated soil mineral N during and after N-limitation 133
Figure 11a: Simulated annual minimum soil water content 134
Figure 11b: Simulated annual soil water content 135
Figure 12: Simulated annual net primary production 136
Figure 13a: Simulated leaf C, constant vs. variable climate 138
Figure 13b: Simulated leaf C, constant vs. variable climate 139
Figure 14a: Simulated leaf C, normal N-limitation 141
Figure 14b: Simulated annual net primary production 142
Figure 14c: Simulated dead stem C 144

Figure 14d: Simulated annual stem growth increment 145
Figure 14e: Simulated annual stem growth increment vs. annual total precipitation 146
Figure 14f: Simulated annual net ecosystem production 148

Chapter 4

Figure 1: Variation in ($T_s - T_a$) with respect to variation in R_n 194
Figure 2: Elevation over study region 199
Figure 3: MAE for monthly averages of predicted Tmax, Tmin and Prcp, with optimal GSP 203
Figure 4: Precipitation terrain smoothing width 205
Figure 5: Monthly Prcp bias vs. monthly Prcp 207
Figure 6: Results of Daymet estimation of meteorological parameters 210
Figure 7: T_{x0} vs. growing-season average Tmax 215
Figure 8a: Predicted vs. observed 14-day deviations from T_{x0} 216
Figure 8b: Predicted vs. observed daily deviations from T_{x1} 217
Figure 9: Observed diurnal temperature curve (mean and 9 stations) 220
Figure 10: Schematic of diurnal temperature prediction logic 222
Figure 11: Daily mean prediction error for subdaily air temperature 223
Figure 12: Predicted vs. observed subdaily air temperature 224
Figure 13: Observed and predicted hourly air temperatures at Missoula, MT 225
Figure 14: Number of cloud/snow contaminated periods 233
Figure 15: Frequency of cloud/snow contaminated compositing periods 234
Figure 16: Number of periods with excessive SATZ 236
Figure 17: Frequency of compositing periods with excessive SATZ 237
Figure 18: Solar zenith angle vs. local solar time 238
Figure 19: Average time of satellite overpass 240
Figure 20: Average emissivity difference vs. average NDVI 245
Figure 21: Landcover classes, recoded from Loveland 252
Figure 22: Predicted r_h, also showing landcover subsets 256
Figure 23: Frequency of predicted r_h 257
Figure 24: Average growing-season NDVI, 1991-1994 258
Figure 25: Frequency of average growing-season NDVI 259
Figure 26: Predicted r_h vs. NDVI, landcover averages 262
Figure 27: Average predicted r_h vs. average NDVI 263
Figure 28: Graph-of-averages, r_h vs. NDVI, by landcover 266

Chapter 1

Introduction

1. Statement of research problems

The broad objective of my dissertation research was to develop a set of numerical tools to facilitate studies in terrestrial ecology over large spatial domains. I approached this objective through the investigation of the following three narrow suppositions: (a) Accurate estimates of land-surface meteorological parameters at an arbitrary landscape location can be derived from a knowledge of those same parameters at a collection of distant points and a knowledge of the relative horizontal and vertical positions of the observation points and the prediction point. (b) A small number of well-established physical, biological, and ecological principles can be integrated, in conjunction with the land-surface meteorological parameters from (a) and a specification of an initial state, to estimate the dominant components of the storage and transport of mass and energy at the land surface as they evolve over time. (c) The spatial and temporal patterns of observed radiometric surface temperature as made from a polar-orbiting satellite can be explained using a combination of estimated surface meteorological parameters and surface energy balance theory as developed in the investigations of suppositions (a) and (b). These three suppositions are addressed in turn in Chapters 2, 3, and 4.

2. Review of research problem investigated in Chapter 2

The hypothesis investigated in Chapter 2, that accurate fields of surface meteorology can be derived from distant observations and their spatial context, is based on the presupposition that a network of observation points exists in the vicinity of the point of desired prediction, and that the spatial density of observations in the network is high enough to record the dominant meteorological dynamics at the spatial scale of relevance to the intended application of the estimated fields. For example, if the intended application required an annual estimate of the number of days with above- and sub-freezing surface air temperature, then a density of observation points with several hundred kilometers between points might be

adequate. If, however, the intended application required a discrimination on each day of the year of above- and sub-freezing surface air temperature, then an adequate observation density would necessarily be higher, with a spacing between points that would resolve the passage of frontal air masses of greatly different temperatures. The required density in this more demanding case would depend on the average speed of frontal passage and the average size of frontal air masses. As an additional constraint, when the intended application requires the resolution of topographic effects in rugged terrain, the hypothesis presupposes that the vertical distribution of observation points is sufficient to sample the important aspects of topographic influence on the meteorological fields of interest. A further logical antecedent to my hypothesis, when applied to estimation in rugged terrain, is that a particular meteorological field will be most sensitive to a particular scale of topography, depending on the physical processes sampled in the measurement network and on the network density. It has already been demonstrated by other research that this is true for estimates of annual total precipitation over very large (sub-continental) regions of complex terrain, and part of the effort in Chapter 2 is directed toward the establishment of this presupposition for daily estimates of both precipitation and temperature. A final presupposition is that a particular meteorological field, as sampled within a particular station network, will be best estimated when the observed values are temporally aggregated or smoothed to some characteristic degree. This should be so when the spatial sampling density and/or the temporal sampling frequency in the network are too low to completely capture the meteorological dynamics of interest. If some aspect of these dynamics is stable or slowly varying over time, and is sampled in the network over a period of time, the temporally smoothed observations should give a better prediction than the observations from any single period.

In order to demonstrate that accurate surface meteorological fields can be derived from a network of observations with a known spatial context, I used as an experimental system the network of daily observations of minimum temperature, maximum temperature, and total precipitation operated in the United States by the National Weather Service and the Natural Resources Conservation Service. I gathered all of the available daily observations for these three variables for one year over an area of approximately 800,000 km^2 which included parts of several mountain chains as well as extensive areas of gentle terrain. I

developed an interpolation and extrapolation strategy which incorporated dependencies on station density, station vertical distribution, topographic scale, and temporal aggregation or smoothing. I then tested the hypothesis that changes in these dependencies resulted in changes in the accuracy of estimates, and that these changes were coherently related, such that optimal values for the strengths of the various dependencies could be derived by minimizing the estimation errors.

3. Review of research problem investigated in Chapter 3

Having demonstrated the validity of my first supposition in Chapter 2, I set out in Chapter 3 to demonstrate that these same meteorological parameters, together with descriptive information concerning the initial states of mass at the land surface, can be incorporated in a mechanistic numerical framework of physical, biological and ecological principles in such a way that the dominant processes of storage and transport of mass and energy at the land-surface can be estimated with some confidence. The fundamental presupposition in this investigation is that the core components of the numerical framework, the underlying physical, biological, and ecological principles, are each individually already well-defined. For example, evaporation of water is one of the dominant processes of water (and energy) transport at the land surface, and has been studied in depth by many investigators under many conditions. My basic assumption with respect to the treatment of the evaporation process is that the results of these prior investigations are valid under certain circumstances, and that by considering the logical constraints and intended application of my integrated numerical framework, I could incorporate them and thereby gain the benefits of earlier efforts. My objective was to identify the dominant processes of storage and transport and to determine whether or not the existing knowledge of mechanisms affecting individual processes, together with the meteorological and initialization data, was sufficient to generate estimates of all the dominant storage and transport processes. I assumed that the land-surface vegetation is the appropriate point of focus for an assessment of dominant processes of storage and exchange. Vegetation is a functional crossroads for the storage and exchange of incident solar radiation, water, carbon dioxide (CO_2) and nitrogen (N) at the land surface, and its presence or absence can serve to indicate the most general features of land surface processes. I further

assumed that processes in the upper layers of the soil were important to the overall dynamics at the land surface, through the interactions of vegetation with soil water and organic matter and mineral nutrient content, as well as through the influence of sensible and latent heat exchange at the soil surface on the surface energy balance.

With these basic assumptions as a foundation, I developed a logical framework that connected processes of storage and transport together and relied only on the available meteorological data and the specification of an initial state. Through literature survey, I identified individual components of this framework that had already been the subject of investigation, and used relevant studies to derive quantitative formulations for implementing my logic. From the field of biophysics, I used the well-developed theories of radiation interception and its conversion to sensible and latent heat fluxes, including assessments of the variation in those processes with respect to varying vegetation cover. From the field of soil physics I used a combination of physical theory and empirical evidence to formulate a relationship between soil texture and soil water-holding characteristics. From investigations of the biology of decomposition in litter and soil I derived a quantitative treatment of the below-ground dynamics of carbon and nitrogen in organic and mineralized forms. From studies in plant ecology I derived a treatment of the acquisition by plants of soil mineral nitrogen, the distribution of that nitrogen to various plant parts, and the sensitivity of that distribution to meteorological conditions. Also from the ecological research literature I derived a simulation logic for allocating plant growth to various organs, dependent on meteorological conditions and the plant state. This effort amounts to a synthesis of the research literature on the dominant terrestrial mass and energy storage and transport processes, in the form of a single operational algorithm. Although I was able to make some novel deductions in certain corners of the mechanistic detail, the primary innovation in this part of my study was the construction of a logical framework which permits the coherent integration of theory and empiricism from a broad range of disciplines.

4. Review of research problem investigated in Chapter 4

I set out in the final chapter to demonstrate that estimated meteorological fields, derived from methods described in Chapter 2, in combination with the theoretical framework for estimating mass and energy storage and fluxes, described in Chapter 3, can provide a reasonable explanation for the variability of radiometric land surface temperature, as observed over sub-continental regions and for multiple years from a sensor onboard a polar-orbiting satellite. Because the surface temperature is an integral component of the surface energy balance, I make the supposition that a correspondence will exist between observed surface temperatures and the variables that are predicted, by the theory presented in Chapter 3, to have the strongest influence on the energy balance. This hypothesis is founded on the following three presuppositions. (a) Factors which complicate the determination of surface temperature from the satellite sensor observations can be identified and either corrected for or safely ignored. (b) The true range of variation in surface temperature is large enough to exceed the lower limits of detection at the instrument, given instrument accuracy and non-instrument sources of error. (c) The physical and analytical theory for predictions of the surface energy balance is generally valid, and is applicable at the spatial and temporal sampling resolution of the sensor. A significant part of my effort in Chapter 4 is in testing the validity of these presuppositions. The purpose of such tests is to provide confidence in the usefulness of my methods in verifying the hypothesis that observed surface temperatures can be explained using results from Chapters 2 and 3, and to help me identify likely causes in the event that the hypothesis is deemed invalid.

Chapter 2

Generating Surfaces of Daily Meteorological Variables over Large Regions of Complex Terrain

1. Background and model development

Much of the recent literature concerning spatial interpolation of meteorological fields has focused on the generation of surfaces of long-term average or climatological precipitation. Particular attention has been given to the development of sophisticated statistical methods which, given certain assumptions, generate explicit optimality criteria and guarantees of unbiased predictions. Some examples are optimal interpolation (Gandin, 1965), kriging and its variants (e.g. Phillips et al., 1992), and smoothing splines (Hutchinson and Bischof, 1983). Simpler methods which lack such optimality criteria and guarantees of unbiasedness have been applied extensively for the determination of mean areal precipitation. The method of nearest neighbors (Thiessen, 1911) is an early example, and others include multiple nearest neighbors, inverse-distance weighting schemes, and arithmetic means.

Several studies have compared the performance of various sophisticated and simple spatial interpolation methods in the context of rainfall predictions (Creutin and Obled, 1982, Tabios and Salas, 1985, Phillips et al., 1992). Other studies have focused on one method but offered comparisons to others (Chua and Bras, 1982, Hevesi et al., 1992). These studies show that although the statistical methods are for the most part more accurate than the simple methods, they are not overwhelmingly so. For example, from Tabios and Salas (1985, Table 10), there is no significant difference between inverse distance methods and a suite of statistical methods, judged on the basis of coefficients of determination for estimates of mean annual precipitation from five sites in homogeneous terrain. In the same study the nearest neighbor method was significantly inferior to the more sophisticated methods, but only by approximately 10%.

These results suggest that an effective, efficient interpolation method could be developed using the simple methods as a starting point.

A recent example in this vein is the work of Daly et al. (1994), who developed an approach for distributing climatological precipitation, combining geographical and statistical elements, which they demonstrate to be both more flexible and more accurate than kriging and some of its variants. Another example is the method of climatologically aided interpolation (CAI), developed by Willmott and Robeson (1995) and applied to the interpolation of yearly temperature averages. CAI uses a relatively simple inverse-distance weighting scheme to adjust a spatially high-resolution climatology. The method produces low validation errors, and its accuracy is attributed in part to the incorporation of terrain effects provided by the high-resolution climatology. A third example, pertinent for its utilitarian and flexible approach, is the recursive filter objective analysis used in the operational analysis of meteorological satellite soundings (Hayden and Purser, 1995). The recursive filter method is designed to handle large data volumes in an operational setting, and employs an iterative algorithm that makes the method sensitive to spatial variability in data density. Although none of these methods are perfectly suited to my purposes, they do embody a desirable spirit of simplicity and utility.

Since I require large interpolated surfaces for a suite of daily meteorological variables, with simulations (typically) of 1-5 years, grids on the order of 500 x 500 cells, and incorporating hundreds of observation sites, computational efficiency is an important factor. The faithful application of any of the statistical methods would require at least one parameterization for each variable for each day, and previous studies suggest that even this would be insufficient for methods such as kriging, given the large and climatically heterogeneous domains of interest. An alternative is to resort to a single parameterization of a statistical method, but this negates what seems to me the principal attraction of such methods, that they generate unbiased results. Rejecting both of these alternatives, I proceed with the development of a method which lacks both formal optimality criteria and guarantees of unbiasedness, but which can be parameterized once for a given set of observations and applied effectively to the daily observations as often as needed.

In developing a new method, I borrow from the nearest-neighbor method the assertion that the area of relative influence for a given observation should be inversely related to the local observation density. In other words, a relatively isolated observation should influence predictions over a larger area than should an observation in a data-rich region. The most serious fault of the nearest neighbor method is that it generates a discontinuous surface, the familiar tessellated surface of Thiessen polygons. A continuous interpolation surface would be preferable. From the inverse-distance method I borrow the assertion that relative influence should decrease with increasing distance from an observation. The most serious flaw of the usual implementation of the inverse-distance method is, in my view, that its asymptotic condition forces the surface through all observations, generating spatially anomalous distributions. I desire a method that is, in this sense, a smoother as opposed to an interpolator, in that the resultant surface is not required to pass through the observations.

I adopt, as a basic interpolation framework, the spatial convolution of a truncated Gaussian filter with a surface containing the horizontal projections of the observation locations. In the spirit of efficiency, truncation of the filter serves to reduce the number of observations included in predictions at a given point: an untruncated filter gives finite weight to all observations at each point of prediction, but the majority of those weights are diminishingly small. Truncation causes a loss of higher-order smoothness, but still results in a continuous surface. I choose a Gaussian function because it is simple to evaluate, and has the desired features of being both an inverse distance algorithm and a smoothing filter. The descriptions here are given with respect to interpolation over an evenly spaced grid of prediction points, but the same methods could be applied to the generation of predicted values at arbitrarily placed points.

2. Methods

2.1. Interpolation

The general form of the truncated Gaussian filter, with respect to its central point, p, is:

Eq. 1 $$W(r) = \left\{ \begin{array}{c} 0;\ r > R_p \\ e^{-\left(\frac{r}{R_p}\right)^2 \alpha} - e^{-\alpha};\ r \le R_p \end{array} \right\}$$

where W(r) is the filter weight associated with a radial distance r from p, R_p is the truncation distance from p, and α is a unitless shape parameter.

The spatial convolution of this filter with a set of horizontal station locations results, for each point of prediction, in a list of weights associated with observations. Because of the spatially heterogeneous distribution of observations, a constant value for R_p results in a large disparity in the number of observations with non-zero weights between points in the least and the most densely populated regions of the prediction grid. I desire a system by which R_p can be reduced in data-rich regions, using information from a smaller radius, and increased in data-poor regions. One possibility is the specification of a fixed number of observations to be used at every prediction point, but this can be shown to violate the requirement for a continuous surface. Instead, I specify N, the average number of observations to be included at each point. R_p is then varied as a smooth function of the local station density in such a way that this average is achieved over the spatial domain. The smooth variation of R_p ensures a continuous interpolation surface, and is accomplished through the iterative estimation of local station density at each prediction point, as follows:

(1) For all grid cells, the same user-specified value, R, is used to initialize R_p.

(2) Given R_p, Eq. 1 is used to calculate weights W_i, where i = (1,...,n) are observation locations, and the local station density D_p (# of stations/area) is then determined as:

$$D_p = \frac{\sum_{i=1}^{n} \frac{W_i}{\overline{W}}}{\pi R_p^2}$$

where W is the average weight over the untruncated region of the kernel, defined as:

$$\overline{W} = \frac{\int_0^{R_p} W(r)\, dr}{\pi R_p^2} = \left(\frac{1 - e^{-\alpha}}{\alpha}\right) - e^{-\alpha}$$

(3) A new R_p is calculated as a function of the desired average number of observations, N, and the most recent calculation of D_p as:

$$R_p = \sqrt{\frac{N^*}{D_p \pi}}$$

where $N^* = 2N$ for the first I-1 iterations, and $N^* = N$ for the final iteration, where I is the number of iterations to be performed. This modification of N^* is a result of filter truncation, and helps to avoid the occassional occurrence of excessively large station counts in regions of strongly heterogeneous station spacing.

(4) The new R_p is substituted in step (2), and steps (2)-(4) are iterated I times.

Final values of R_p are incorporated in Eq. 1 to generate the interpolation weights W_i used in predictions for all days at the point in question.

The interpolation method for a given set of observations and a given prediction grid is defined by the four parameters R, I, N, and α. Given an arbitrary variable x_i, measured at each of the i = (1,...,n) observation points, values for the interpolation parameters are specified once and held constant over all days and all prediction points. Taking the case of a single prediction point on a single day, the interpolated value x_p is determined in general as:

Eq. 2 $$x_p = \frac{\sum_{i=1}^{n} W_i x_i}{\sum_{i=1}^{n} W_i}$$

This general method is refined below, making it specific to predictions of daily temperature extremes and daily total precipitation, incorporating an objective analysis of the influence of elevation differences.

2.2. Temperature Predictions

Prediction methods for TMAX and TMIN are identical, and I will refer here to a general daily temperature variable, T. I focus on the generation of a prediction T_p at a single point and for a single day, based on observations T_i, and interpolation weights W_i, for the i = (1,...,n) measurement sites. Prediction of

T_p requires a modification of Eq. 2 to include a correction for the effects of elevation differences between the observation points and the prediction point. This correction is based on an empirical analysis of the relationship of T to elevation, which is performed once for each day of prediction.

I introduce a set of transformed variables for the empirical analysis of elevation relationships, under the hypothesis that these relationships may have characteristic spatial and temporal scales which are not well represented by the recorded station elevations (Z_i) and the daily temperature observations (T_i). These new variables are z_i and t_i, a spatial transform of the recorded station elevations and a temporal transform of the daily observations, respectively. In a later section I examine the explicit connection between the transformed and untransformed variables. Note that the use of these transformed variables is limited to the assessment of influence of elevation on predictions of T_p, and that the untransformed daily observations T_i are incorporated in the eventual prediction algorithm.

A weighted least squares regression is used to assess the relationship between t and z. The daily regression is performed over all unique pairs of stations, and the regression weight associated with each point is the product of the interpolation weights associated with the stations in a pair. The independent variable is the difference in the transformed elevations associated with a pair of stations, and the dependent variable is the corresponding difference in the transformed temperatures associated with the pair. This gives a regression of the form:

$$(t_1 - t_2) = \beta_0 + \beta_1 (z_1 - z_2)$$

where subscripts 1 and 2 indicate the two stations of a unique pair, and β_0 and β_1 are the regression coefficients.

T_p is then predicted as follows:

Eq. 3
$$T_p = \frac{\sum_{i=1}^{n} W_i \left(T_i + \beta_0 + \beta_1 \left(z_p - z_i \right) \right)}{\sum_{i=1}^{n} W_i}$$

where z_p is the elevation assigned to the prediction point, and T_p would, in practice, be replaced by either $TMAX_p$ or $TMIN_p$.

2.3. Precipitation Predictions

Predictions of precipitation are complicated by the need to predict both daily occurrence and, conditional on that result, daily total precipitation. Under the assumption that there is some spatial coherence to the patterns of precipitation occurrence (wet vs. dry) when measured at the time scale of a day, I define a simple binomial predictor of spatial precipitation occurrence as a function of the weighted occurrence at surrounding stations. Taking the case of a single prediction point on a given day, and given observations of daily total precipitation P_i, and interpolation weights W_i, I estimate a precipitation occurrence probability, POP_p:

$$POP_p = \frac{\sum_{i=1}^{n} W_i PO_i}{\sum_{i=1}^{n} W_i}$$

$$PO_i = \left\{ \begin{array}{l} 0 \; ; P_i = 0 \\ 1 \; ; P_i > 0 \end{array} \right\}$$

where the PO_i are binomial variables related to observed precipitation occurrence. Daily binomial predictions of precipitation occurrence at a given point, PO_p, are based on the comparison of POP_p with a specified critical value, POP_{crit}:

$$PO_p = \left\{ \begin{array}{l} 0 \; ; POP_p < POP_{crit} \\ 1 \; ; POP_p \geq POP_{crit} \end{array} \right\}$$

POP_{crit} is held constant for the entire spatial and temporal domain of the simulation.

Conditional on precipitation occurrence ($PO_p = 1$), I proceed with the prediction of daily total precipitation, P_p. Under the same assumptions outlined for temperature predictions, I introduce the transformed variables p_i and z_i in the objective analysis of the relationship of precipitation to elevation.

Again I use a weighted least squares regression, with the same form for the weights and the independent variable as described for temperature predictions. The dependent variable in this case is defined as the normalized difference of the transformed precipitation observations, giving a regression of the form:

$$\left(\frac{p_1 - p_2}{p_1 + p_2} \right) = \beta_0 + \beta_1 (z_1 - z_2)$$

In generating the predicted daily total, P_p, the interpolation weights are multiplied by the station occurrence variable, PO_i, giving weight only to those stations which record precipitation for the day in question, as follows:

$$P_p = \frac{\sum_{i=1}^{n} W_i \, PO_i \left(\frac{1+f}{1-f} \right)}{\sum_{i=1}^{n} W_i \, PO_i}$$

Eq. 4

$$f = \beta_0 + \beta_1 (z_p - z_i)$$

The form of the precipitation prediction requires that $|f| < 1$. I introduce another parameter, f_{max} (< 1.0), and force $|f| = f_{max}$ whenever $|f| > f_{max}$.

2.4. Extrapolation Smoothing

I hypothesize that the processes controlling the observed variation of temperature and precipitation with elevation may have characteristic spatial and temporal scales different from those implied by the recorded station elevations or by observations at a daily time-step. For example, it may be that better predictions are obtained from regressions using a spatially smoothed elevation field as opposed to recorded station elevations for the calculation of elevation differences, or that the variation with elevation is better explained using multiple-day running averages of observations as opposed to the daily observations themselves. I introduce the parameters S_S and S_T to describe the spatial and temporal smoothing characteristics for regressions of each variable with elevation.

The parameter S_S describes the degree of spatial smoothing incorporated in the transformation from Z_i to z_i for a particular variable, and S_T describes the degree of temporal smoothing incorporated in

the transformation from T_i or P_i to t_i or p_i. S_S, measured in km, defines the width of a rectangular region around the presumed location of a particular station for which elevation data from a digital terrain grid is averaged to generate the transformed elevation, z_i (see below for a discussion of station location). S_T, measured in days, defines the width of a two-sided linearly tapered smoothing filter applied to the time series of T or P to generate t or p. Ends of the series are padded with zeros for the purpose of this smoothing filter. In the case of precipitation time series, the smoothing weights for days with no precipitation are set to zero, so the resulting smoothed value represents a weighted average of daily precipitation events. This correction is required to avoid the "constant drizzle" bias associated with a simple smoothing filter.

2.5. Humidity Predictions

Predictions of humidity are based on the assumption that minimum daily air temperature (T_m) is a reasonable surrogate for dew-point temperature (T_d). Tests of this relationship over a wide range of climatic settings indicate that its accuracy decreases with increasing aridity (Kimball et al., 1995), but that in general it is an adequate approximation in the absence of high-quality humidity measurements (Running et al., 1987, Glassy and Running, 1994). I generate estimates of humidity in terms of the average daytime saturation vapor pressure deficit VPD (Pa), as:

$$VPD = e_s(T_a) - e_m$$

where $e_s(T_a)$ is the saturated vapor pressure (Pa) at the average daytime site temperature T_a (°C), and e_m is the ambient vapor pressure (Pa) as inferred from the assumption that $T_m = T_d$. Vapor pressures are calculated using Murray's (1967) formulation:

$$e_s(T_a) = 610.78\, e^{\left[\frac{17.269\, T_a}{237.3 + T_a}\right]}$$

$$e_m = 610.78\, e^{\left[\frac{17.269\, T_m}{237.3 + T_m}\right]}$$

Tests of the assumption $T_d = T_m$ have not focused explicitly on the sensitivity of the relationship to variation in T_m due to local topography. While the mole fraction of water vapor in a well-mixed air column is insensitive to variation in pressure and temperature, partial pressure of water vapor is not. I therefore

expect e_m to vary with terrain height, and I assume that the methods described above to analyze the dependence of TMIN on elevation are also applicable to the purpose of estimating variation in e_m, and I set $T_m = TMIN_p$. Similarly, the variation in T_a with respect to elevation has an important effect on VPD, and following Running et al. (1987) I specify:

$$T_a = 0.606\,TMAX_p + 0.394\,TMIN_p$$

2.6. Radiation Predictions

Direct and diffuse components of incident shortwave radiation are calculated on a sub-daily timestep (typically 10 min.) using expressions that analyze the sun-earth geometry, including corrections for slope and aspect in complex terrain. Radiative fluxes at the top of the atmosphere are attenuated as a function of atmospheric transmissivity, including corrections for optical air mass and cloudiness. I use the same equations for sun-slope geometry and the empirical treatment of diffuse radiation described by Hungerford et al. (1989), but substitute the following calculation of daylength, DL (s), for a flat surface with unimpeded horizons:

$$HSS = \mathrm{acos}\left(- \frac{\sin(LAT)\sin(DECL)}{\cos(LAT)\cos(DECL)} \right)$$

$$\text{if } HSS < -\pi \text{ then } HSS = -\pi \quad (24\,\text{hr daylight})$$
$$\text{if } HSS > \pi \text{ then } HSS = \pi \quad (0\,\text{hr daylight})$$

$$DL = 2\ HSS \cdot 13751.0$$

where HSS (radians) is the hour angle of sunset, measured from solar noon, LAT (radians) is the latitude, DECL (radians) is the declination angle of the sun and the constant 13751.0 converts from radians of hour angle to seconds of daylength. This algorithm reduces errors at high (and southern) latitudes associated with the original daylength algorithm in MTCLIM (Forsythe et al., 1995). DL from this formula is truncated for sloping surfaces as in the original model documentation, but corrections to horizon angles for shading from adjacent terrain elements are not implemented here.

The algorithm of Bristow and Campbell (1984) is used to derive a daily average cloudiness correction to atmospheric transmissivity from the observed diurnal temperature range, DTR, as:

Eq. 5 $$PCST = 1.0 - e^{(-B \cdot DTR^C)}$$

where PCST is the proportion of clear-sky transmissivity on the day in question, and B and C are empirical parameters (see Glassy and Running, 1994, for a discussion of the B and C parameters; here I use the values -0.003 and 2.4, respectively). This method has been shown to successfully predict a large proportion of the variation in daily radiation fluxes (Running et al., 1987, Glassy and Running, 1994), but no analysis has related its parameters to temperature variability imposed by topographic features. I observe that the environmental lapse rate for minimum temperature is generally of smaller magnitude than that for maximum temperature, and so DTR in a neighborhood will generally decrease with increasing elevation. A strict application of Eq. 5 to the predicted surfaces of DTR ($TMAX_p$ - $TMIN_p$) would yield consistently lower transmissivity at the higher elevations in a neighborhood of complex terrain. In general, transmissivity is expected to *increase* with increasing elevation, due to a reduction in optical air mass. While some argument could be made for more frequent cloudy conditions over high terrain, I think it is unlikely that this phenomenon is responsible for the observed differences in maximum and minimum temperature lapse rates. My solution is to neglect the influence of elevation on $TMAX_p$ and $TMIN_p$ for the purpose of calculating DTR_p, performing a simple interpolation of the observed DTR_i, using the same form as in Eq. 3 and regarding β_0 and β_1 as 0.0.

Final predictions of SRAD are made by summing the direct and diffuse shortwave components of radiation over the day and dividing by the daylength, giving the daylight average instantaneous shortwave flux density in W/m^2.

2.7. Case-Study Description and Database Filtering

I implemented these methods for a one-year simulation of daily TMAX, TMIN, PRCP, VPD, and SRAD over an area of 399,360 km^2 (832 km x 480 km) in the inland northwestern United States (Figure 1, topographic detail shown in color plate 11a). The region includes southeastern Washington, northeastern Oregon, central Idaho, and southwestern Montana, and was selected to include a diversity of climatic and

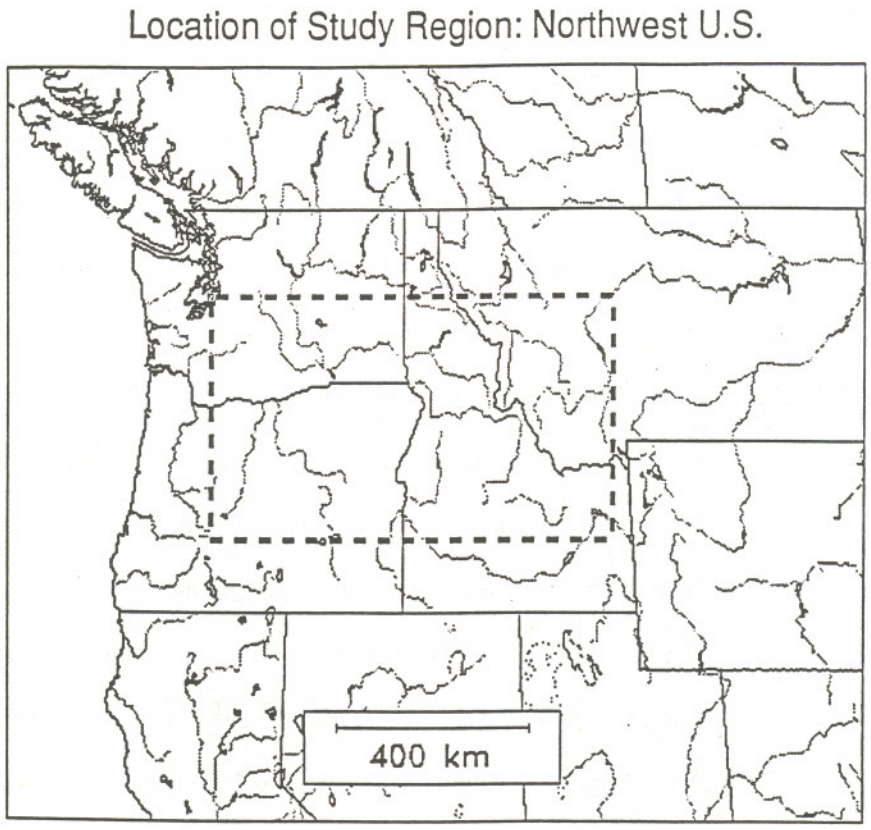


Figure 1.
Northwestern United States, showing state outlines, major river systems, and the approximate boundaries of the study region.

topographic regions. Elevations over the region range from sea level to 4000 m. Although there is considerable variation, vegetation is generally grassland, agriculture, and desert at elevations up to 700 m, with coniferous forest dominating at higher elevations, and limited alpine tundra at very high elevations. A notable exception is the dense forest cover ranging from sea level to about 1000 m on the west slope of the Cascade range. The region extends just to the west of the Cascade Range in Washington and Oregon, encompassing the transition from maritime to continental climates across the Cascade divide. The western Cascade slope is characterized by frequent heavy precipitation, with a gradual increase in annual total precipitation with elevation. The eastern Cascade slope is characterized by a dramatic precipitation gradient, with semi-arid conditions extending 200-300 km eastward of the foot of the range. Relatively hot and dry conditions prevail across the southern extent of the region, through the northern end of the Basin and Range province, to the Snake River Plain. The eastern portion of the region is dominated by a multitude of Northern Rocky Mountain ranges with elevation ranging from 800 to 4000 m, and very complex topography. Storm tracks are generally from the west, and the west side of this group of ranges receives more precipitation than the east. A minority of storms track from the southeast, and the southern ranges receive the bulk of the moisture from these storms. For the region as a whole, wintertime precipitation comes from large frontal systems, while most summer precipitation is due to local convective activity. An exception is the region west of the Cascade divide, where frontal precipitation dominates year-round.

Daily meteorological data for 1989 were obtained from the National Climatic Data Center (NWS) and from the Western Regional Climatic Data Center (USDA). Data for stations outside the validation area was incorporated in the predictions for the validation stations in order to mitigate the influence of data-sparse edges on interpolation errors. Not all stations recorded all three of the primary variables, and so the number of stations differs somewhat for predictions of temperatures and precipitation. Many SNOTEL stations measured precipitation but not temperature, while most NWS stations recorded all three primary variables, resulting in larger numbers of precipitation observations (Table 1).

Table 1: Numbers of meteorological observation stations.

Organized by station type (WS = National Weather Service Cooperative Observers Network station, ST = Natural Resources Conservation Service SNOTEL station) and by inclusion in validation data set (IN = inside validation region, USED = inside region or in bordering area).

	IN			USED		
Variable	Total	WS	ST	Total	WS	ST
TMAX/TMIN	280	213	67	436	344	92
PRCP	365	220	145	498	321	177

The original station list was filtered separately for temperature and precipitation data to exclude stations with excessive missing data. Stations were dropped from the initial database if they contained more than 25 days of missing data for the year or if they contained more than 5 consecutive days of missing data. Stations included in temperature predictions were required to pass this criteria for both TMAX and TMIN, since these are required in tandem for predictions of radiation. Days with missing data for stations passing these criteria were excluded from parameterization and validation analyses.

Station records include the longitude, latitude, and elevation for each station. Longitude and latitude are recorded by the NWS to the nearest arc minute, and elevations are recorded to the nearest meter. An accuracy of ± 1 arc-minute corresponds to a potential error in station location of 3.7 km for latitude and 2.5 km for longitude (at 47 °N). This is a considerable error, and it creates some difficulties in the registration of station locations to digital terrain maps. In this study I used a digital terrain grid with a grid-cell spacing of 500 m (from USGS), with elevations accurate to about ± 6 m. The projection of recorded station longitude and latitude onto this elevation grid results in absolute differences between grid and station elevations which average about 90 m, with a standard deviation of 151 m. Some of this variation is likely due to the sampling methods used to generate the terrain grid, but I suspect a large proportion of the variation is due to inadequate station location data. Given the relative accuracy of station and terrain grid elevations compared to horizontal station locations, I reassigned the station locations to the central point of the 500 m grid cell within a 4.5 km x 4.5 km neighborhood around the recorded station location for which the error between recorded and terrain grid elevations was minimized. The average

absolute difference between grid and station elevation after this location adjustment was 11 m, with a standard deviation of 39 m. The average change in horizontal location was 1.9 km, with a standard deviation of 0.7 km. There were no significant differences between precipitation and temperature stations or between NWS and SNOTEL stations in relocation distances.

2.8. Parameterization and Validation

Cross-validation analysis was used to test the sensitivity of these methods to variation of parameters and to estimate the prediction errors associated with the final selected parameters. The general cross-validation protocol is to withhold one observation at a time from a sample, generating a prediction error for the withheld case by comparison with the observed value, and repeating over all observations in the sample to generate an average prediction error. The sample in this case is the set of stations which record TMAX, TMIN, or PRCP on a given day. VPD and SRAD are derived from TMAX and TMIN and I am unable to validate them in this framework, having no pertinent observations. I am interested in both the absolute value and the sign of prediction errors generated in this manner, and I chose the mean absolute error (MAE) and the bias as prediction error statistics. MAE does not exaggerate the influence of outliers, as does the more common root mean squared error (RMSE) statistic, and it therefore provides a more robust parameterization framework than RMSE.

The parameterization of these methods is iterative, in that all parameters relevant to the prediction of one of the primary variables must be specified in order to generate cross-validation prediction errors, even though optimal values for some (or, at first, all) of the relevant parameters are unknown. I isolated sets of parameters which were not strongly mutually dependent, and tested the covariation of parameters within these sets independently, and afterwards combined the results and performed the covariation tests again to correct for the weaker between-set dependencies. After values for all parameters were established, a final cross-validation analysis was performed, in which predictions were compared against observations for both daily values and annual averages (or totals, in the case of precipitation). In the assessment of validation results I give particular attention to daily event frequency histograms, and to the predicted and observed relationships between precipitation occurrence and amount.

2.9. Spatial Scaling Analysis

My methods are designed to be independent of prediction grid resolution: the process of parameterization and validation is carried out with what are essentially point observations, and I assume that these predictions maintain their validity when applied to the points of a prediction grid. There is another level of abstraction involved in translating these predictions to areal totals or averages as determined by the area of grid boxes centered on the prediction points. I examined the effects of prediction grid resolution on results expressed as areal totals or averages by generating a sequence of increasingly larger prediction grids, ranging in resolution (grid point separation) from 500 m to 32 km. Digital terrain data with a resolution of 500 m provided the starting point for this analysis, and I aggregated this data to successively larger grids with resolutions of 1, 2, 4, 8, 16, and 32 km, taking care to maintain the areal correspondence of all grids. Step-wise resampling with bilinear interpolation was used to generate grids with progressively larger cell sizes. Daily simulations for one year were performed over each grid, and a comparison made of the areal results.

2.10. Temporal Scaling Analysis

Although these methods are formulated with a daily timestep in mind, they can be transformed to longer timesteps in a relatively straightforward way. I am encouraged to attempt this transformation because the majority of other published methods operate on monthly or annual timesteps, and because there is ongoing interest in the intercomparison of methods, including this one, with respect to the prediction of annual total precipitation. Here I make a simple analysis comparing the annual average (for temperature) or total (for precipitation) of daily predictions with an implementation that predicts the annual averages or totals from averages or totals of the observations. This transformation is quite simple for temperature predictions, where I replace the daily observations with the corresponding annual averages, ignoring the S_T parameter. For the case of precipitation, I neglect the occurrence prediction, and proceed with the amount prediction in the same way, assuming all $PO_i = 1$ and ignoring S_T. All other parameters are retained as the optimized values for the case of daily predictions.

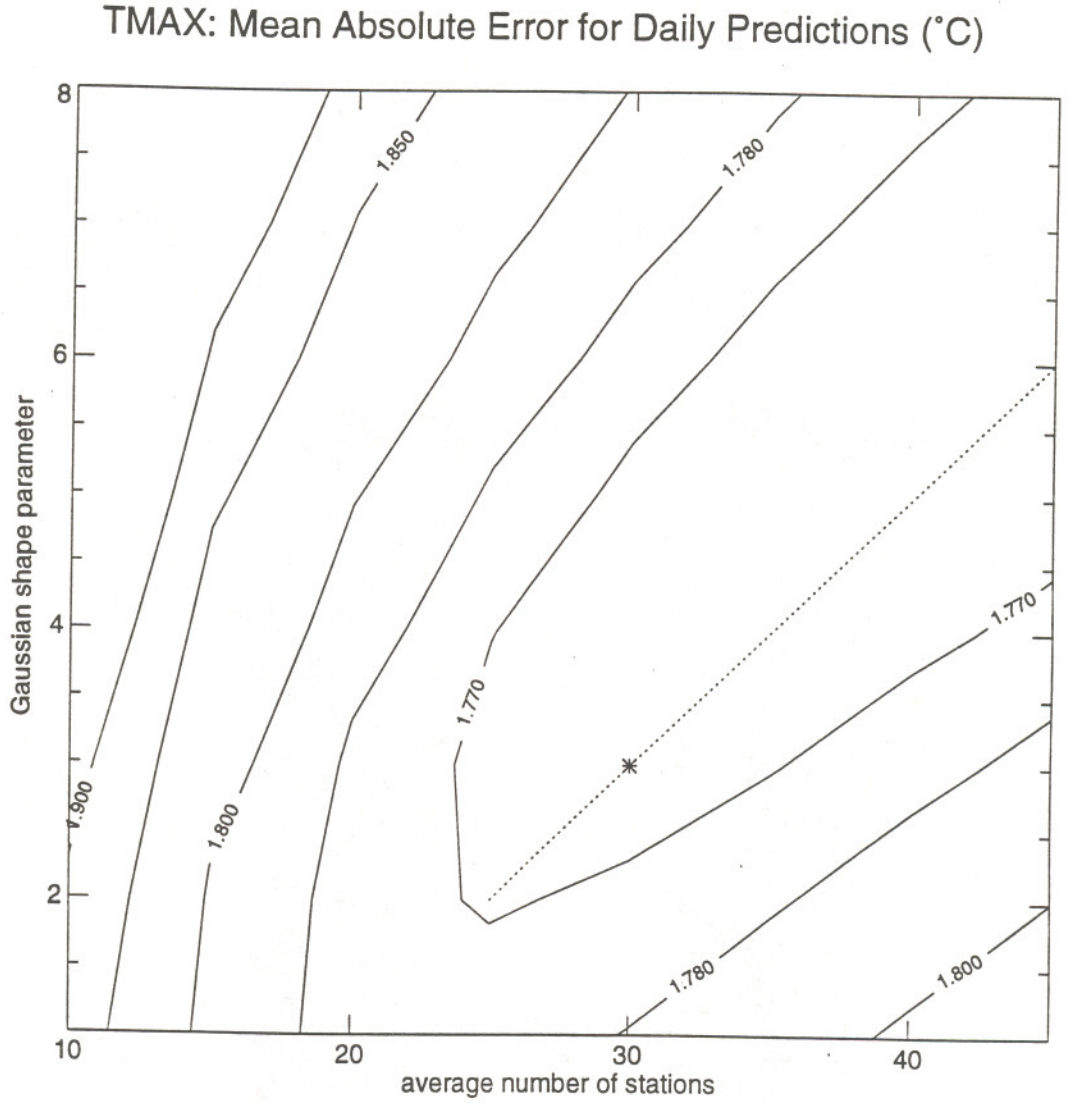
3. Results

3.1. Parameterization

For each of the three primary variables, the interpolation parameters (α: shape parameter, N: average number of stations with non-zero weights, I: number of station density iterations, and R: initial truncation radius for iterative density algorithm) are estimated independently. I find that the prediction errors are quite insensitive to the choice of R for all variables, as long as it is large enough that on the first density iteration at least one station is found inside the truncation radius for each point in the prediction grid. Given the average station density in this case, I assign R = 140 km for interpolations for all three primary variables. I also find that values for I > 3 do not generate substantially different smoothed surfaces of R_p than I = 3, and so I = 3 is used by default in all interpolations.

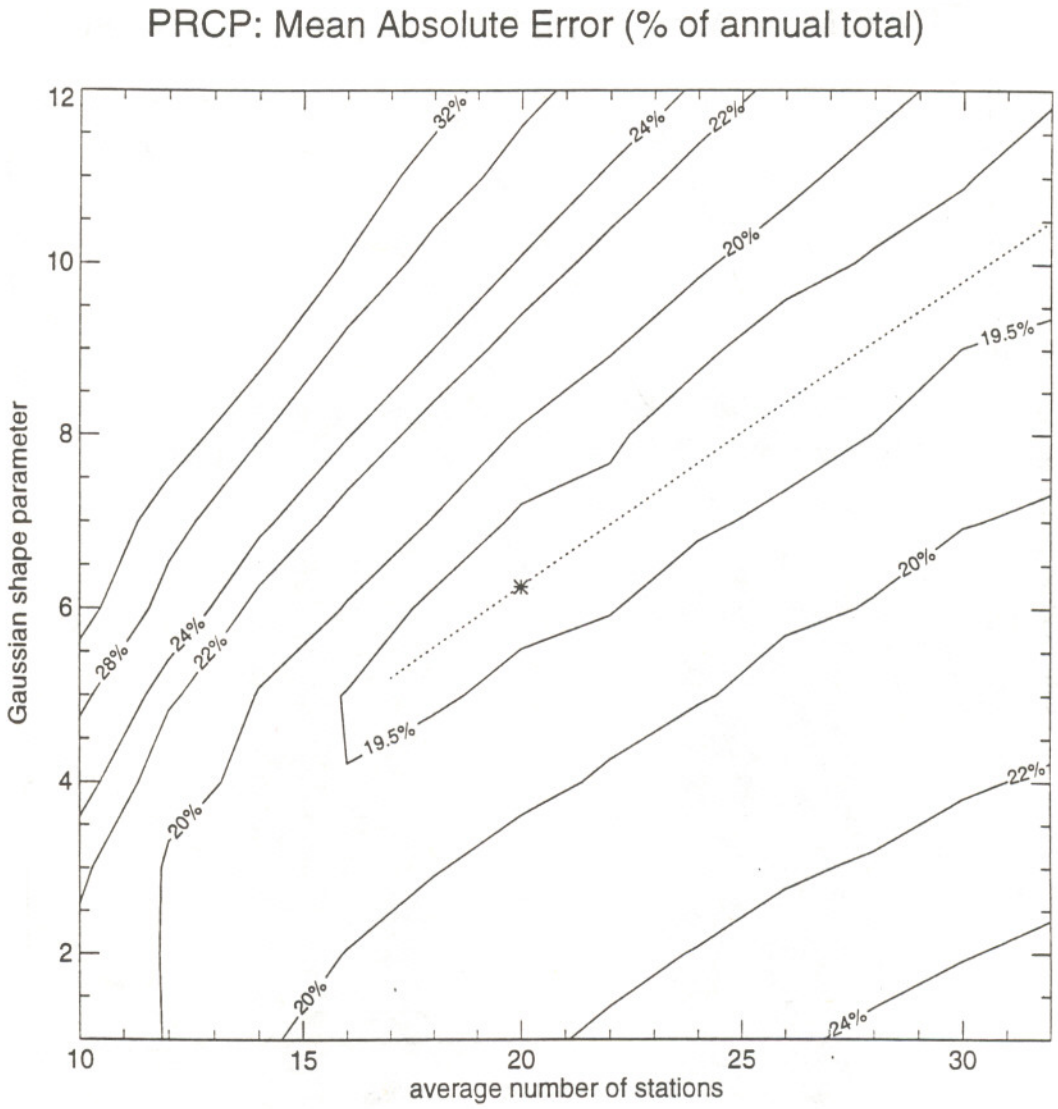
Prediction errors associated with the two remaining interpolation parameters, α and N, are found for all three primary variables to exhibit substantial covariation. Prediction errors for TMAX and TMIN were examined for daily predictions and for annual averages of those daily predictions, and in all cases a linear trough of minimized MAE extends from (N,α) = (25,2.0) to (45,6.0) and beyond. Because a low value of N leads to computational savings in the interpolation process, I chose N = 30 and α = 3.0 for interpolations of both TMAX and TMIN. The error surfaces for TMAX and TMIN predictions are similar, and an example is given for the daily prediction of TMAX (Figure 2a). Prediction error for annual total precipitation, summed from daily predictions and expressed as a percentage of the total observed precipitation, also shows a linear trough over a range of N and α, but with optimal values markedly different than for the temperature predictions (Figure 2b). I chose N = 20 and α = 6.25 for the precipitation interpolations.

S_S and S_T were tested in tandem for each of the primary variables. For both TMAX and TMIN, the lowest prediction errors were associated with the use of recorded station elevations in the elevation regressions; prediction error increased linearly with increasing spatial smoothing, S_S. Similarly, prediction errors increased linearly for both TMAX and TMIN with increasing S_T, with minimum errors obtained using the unsmoothed temperature observations. Prediction errors for precipitation, on the other hand, were



a)

Figure 2.
Contours of mean absolute error plotted against the two most sensitive interpolation parameters, α (Gaussian shape parameter), and N (average number of stations with non-zero weights), for a) daily predictions of TMAX, and b) annual totals of daily predictions of PRCP. Dashed lines represent the approximate location of the trough of minimized error, as referenced in the text. Star (*) indicates the coordinate pair selected as the final parameter values.


PRCP: Mean Absolute Error (% of annual total)
Gaussian shape parameter
average number of stations
12
10
8
6
4
2
10
15
20
25
30
32%
28%
24%
22%
20%
19.5%

b)

found to be substantially reduced by both spatial and temporal smoothing: S_S between 2 and 8 km, and S_T = 5 days, gave optimal results.

The final two model parameters, POP_{crit} and f_{max}, are specific to the daily precipitation algorithm. The value for POP_{crit} should be close to 0.5, since it sets the probability of rainfall given the weighted occurrence at a sample of nearby stations: If half or more of these stations record rainfall, one would intuitively predict an event occurrence, otherwise not. Values lower than 0.5 should result in overprediction of events, and therefore a positive bias and a large MAE for rainfall amount. Conversely, values much higher than 0.5 should result in too few predicted events, a negative bias, and again a large MAE for rainfall amount. This is in fact what is observed, with the smallest MAE and bias closest to zero at a value of POP_{crit} = 0.50 when the annual total error statistics are measured as cm of precipitation. Expressing MAE as a proportion of the observed totals gives an optimal value of 0.55. As I am not sure which error statistic to favor, I chose a value of 0.52 as a compromise.

As noted in the description of the precipitation algorithm, the parameter f_{max} is used to constrain the righthand side of Eq. 4 in the case of large elevation differences and strong precipitation-elevation gradients measured at relatively low stations. This parameter essentially embodies the method's lack of predictive ability for the very highest and wettest terrain. It must assume a value less than 1.0, and should ideally be set very close to this limit in order to allow as much information as possible from the regression with elevation to enter the predictions. Values too close to 1.0 will result in excessive predictions at high elevations, increasing the MAE estimated from predicted annual totals. There is a gradual decline in MAE as f_{max} increases from 0.1 to 0.95, with a sharp increase in MAE for values approaching 1.0. (Figure 3). Results are the same for MAE measured in cm of total annual precipitation and as proportions of observed totals, and f_{max} is set to 0.95 for these simulations. All parameters and their values are listed in Table 2.

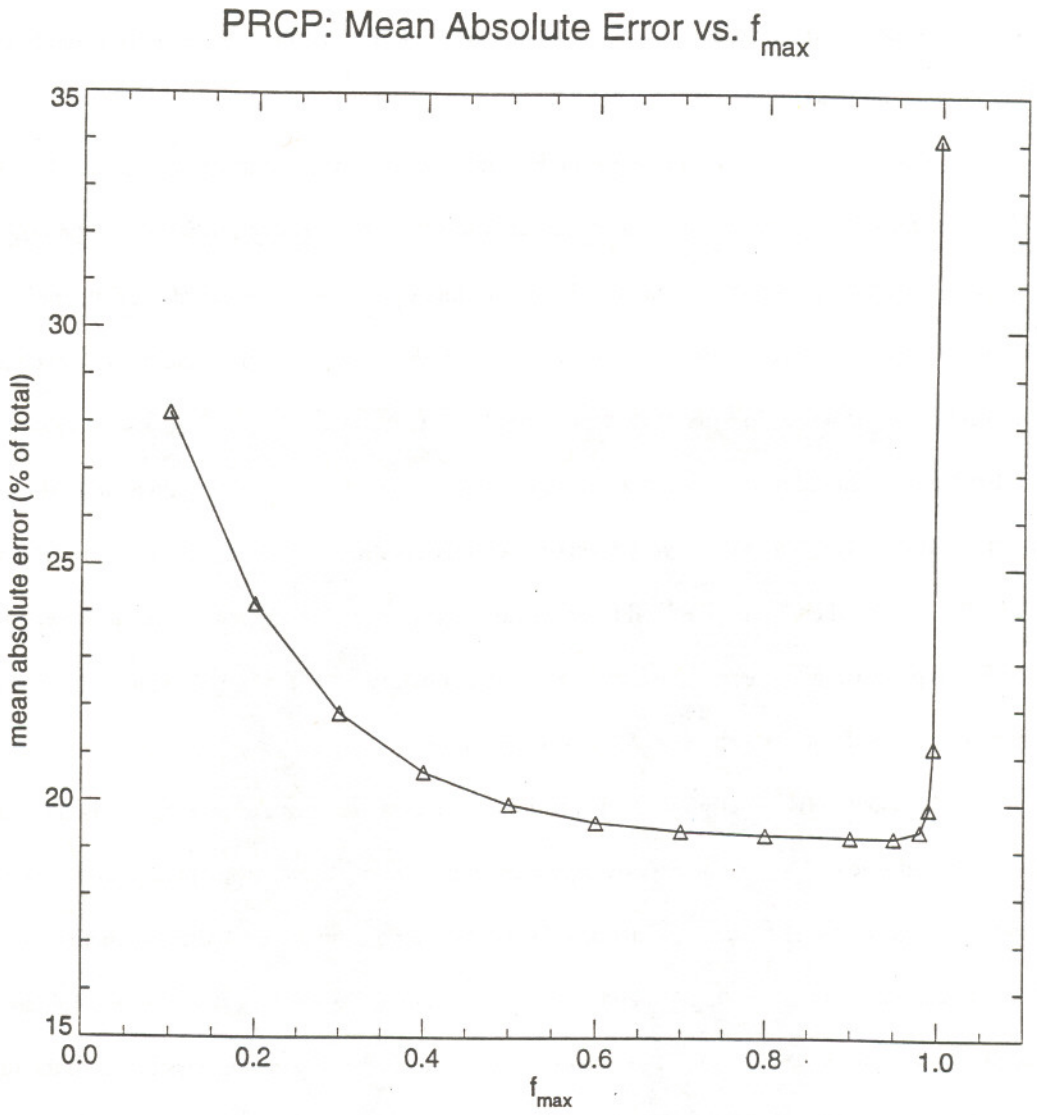


Figure 3.
Influence of f_{max} on MAE (% of annual total) for predicted annual total precipitation. MAE is minimized at $f_{max} = 0.95$. Similar results are obtained for MAE measured in cm.

Table 2: List of model parameters.
(ALL indicates that this parameter is used for TMAX, TMIN, and PRCP with the same value)

Param.	Description	Variable	Value	Units
I	Number of station density iterations	ALL	3	none
R	Truncation radius for Gaussian filter	ALL	140	km
α	Shape parameter for Gaussian filter	TMAX	3.0	none
		TMIN	3.0	none
		PRCP	6.25	none
N	Average number of stations with non-zero weight	TMAX	30	none
		TMIN	30	none
		PRCP	20	none
S_S	Spatial smoothing width for elevation regressions	TMAX	0	km
		TMIN	0	km
		PRCP	3.5	km
S_T	Temporal smoothing width for elevation regressions	TMAX	1	days
		TMIN	1	days
		PRCP	5	days
POP_{crit}	Critical precipitation occurrence probability	PRCP	0.52	none
f_{max}	Precipitation regression constraint	PRCP	0.95	none

3.2. Validation

Given the parameter values established above (Table 2), I performed a final cross-validation analysis to assess the accuracy of these methods. I examine here the performance of the daily predictions as well as their annual averages and the frequency distribution of daily observations and predictions. MAE and bias statistics for TMAX, TMIN, DTR, and PRCP are summarized in Table 3.

Table 3: Cross-validation error statistics.
Statistics for predictions of daily temperatures, annual average temperatures, and annual total precipitation, averaged over all stations within the validation region, for 1989. Precipitation statistics are given both in cm of precipitation per year and % of observed annual total.

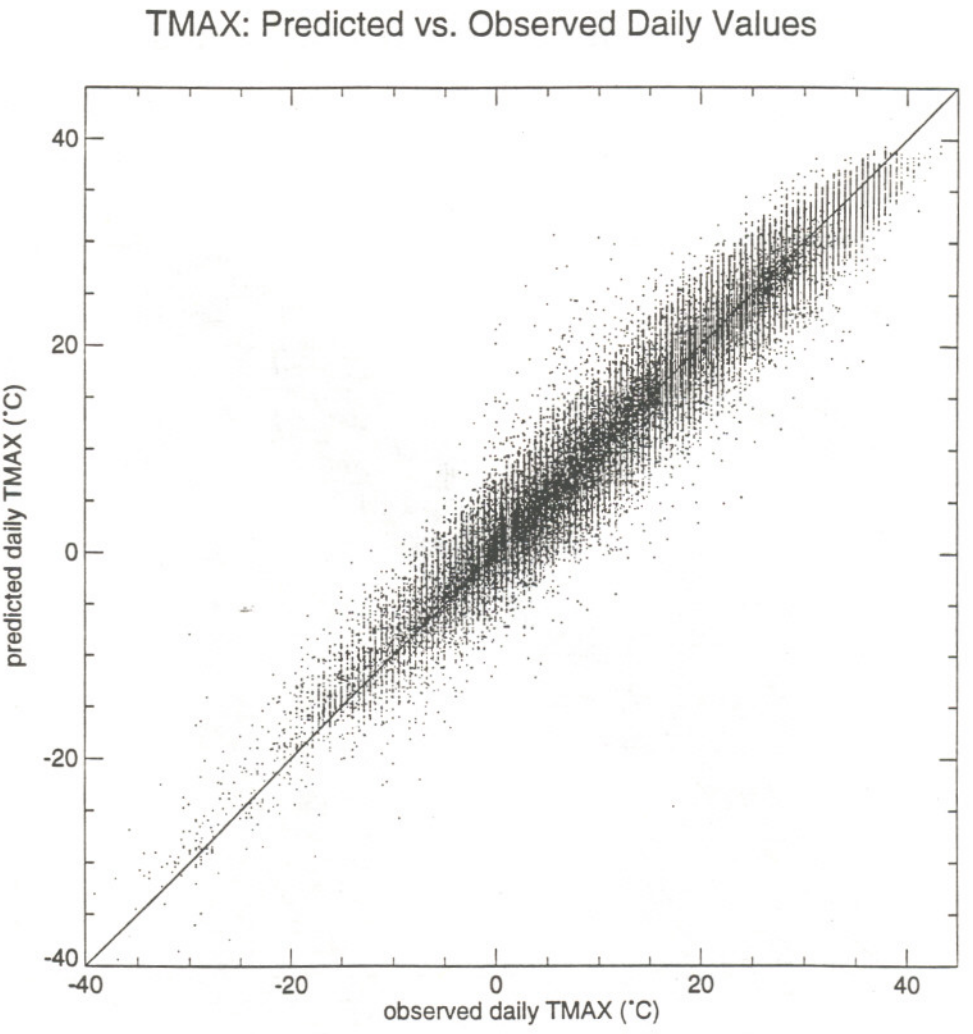
	Daily predictions from daily observations		Annual predictions from daily observations		Annual predictions from annual observations	
Variable (units)	MAE	Bias	MAE	Bias	MAE	Bias
TMAX (°C)	1.76	-0.0002	0.72	-0.07	0.72	-0.07
TMIN (°C)	1.95	+0.0001	1.24	+0.05	1.24	+0.05
DTR (°C)	2.30	-0.0003	1.41	-0.12	1.41	-0.12

PRCP (cm)	NA	NA	13.36	-2.21	12.24	+0.27
PRCP (%)	NA	NA	19.27	+3.03	18.47	+7.09

Cross-validation MAEs for daily predicted vs. observed TMAX and TMIN are 1.8 and 2.0 °C, and MAEs for annual averages of daily estimates are 0.7 and 1.2 °C, respectively. Bias for annual average TMAX and TMIN are -0.1 and +0.1 °C, respectively. These errors are very similar in magnitude to those reported for annual predictions from a global database, using a simple interpolation method (Wilmott and Robeson, 1995), and to errors for a recent point estimation method (Degaetano et al., 1995), where a larger prediction error for TMIN than for TMAX is also reported. Scatter-plots and frequency histograms of all daily cross-validation temperature predictions show good agreement over most of the observed range, with a tendency to under-estimate very warm temperatures and to over-estimate very cold temperatures (Figure 4).

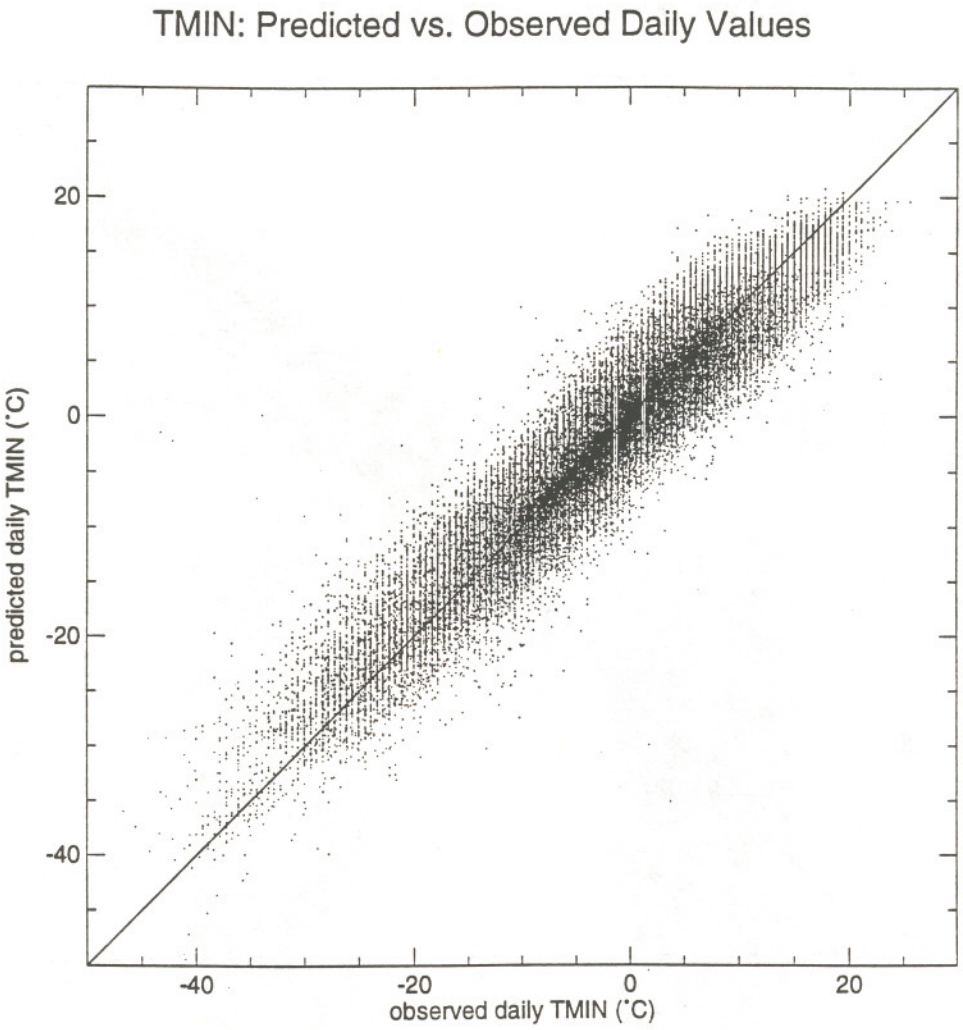
MAE for daily and annual estimates of DTR are 2.3 and 1.4 °C, respectively, with a bias in the annual averages of -0.1 °C. The frequency histogram for DTR (not shown) indicates a more serious error in prediction of extreme values than observed for TMAX and TMIN, which is the expected result for an aggregate variable. This information is included to give a general notion of the likely errors in the radiation routine as a result of errors in DTR: since corrections to DTR for elevation are neglected in the radiation algorithm there are no strictly applicable error statistics.

The validation of the daily precipitation model is somewhat more involved due to the influence of daily occurrence predictions. Strict estimates of daily error, as performed for temperature variables, are unenlightening: It is a curious fact that these errors are minimized by assuming no precipitation whatsoever, since for the climate in question this is true on a large majority of station-days. MAE estimated from annual totals of daily predictions, as used in the model parameterization, is a much more robust indicator of model performance, but it sheds no light on the relationship between occurrence and amount. For example, a reasonable annual total could be attained with a large number of very small daily events or

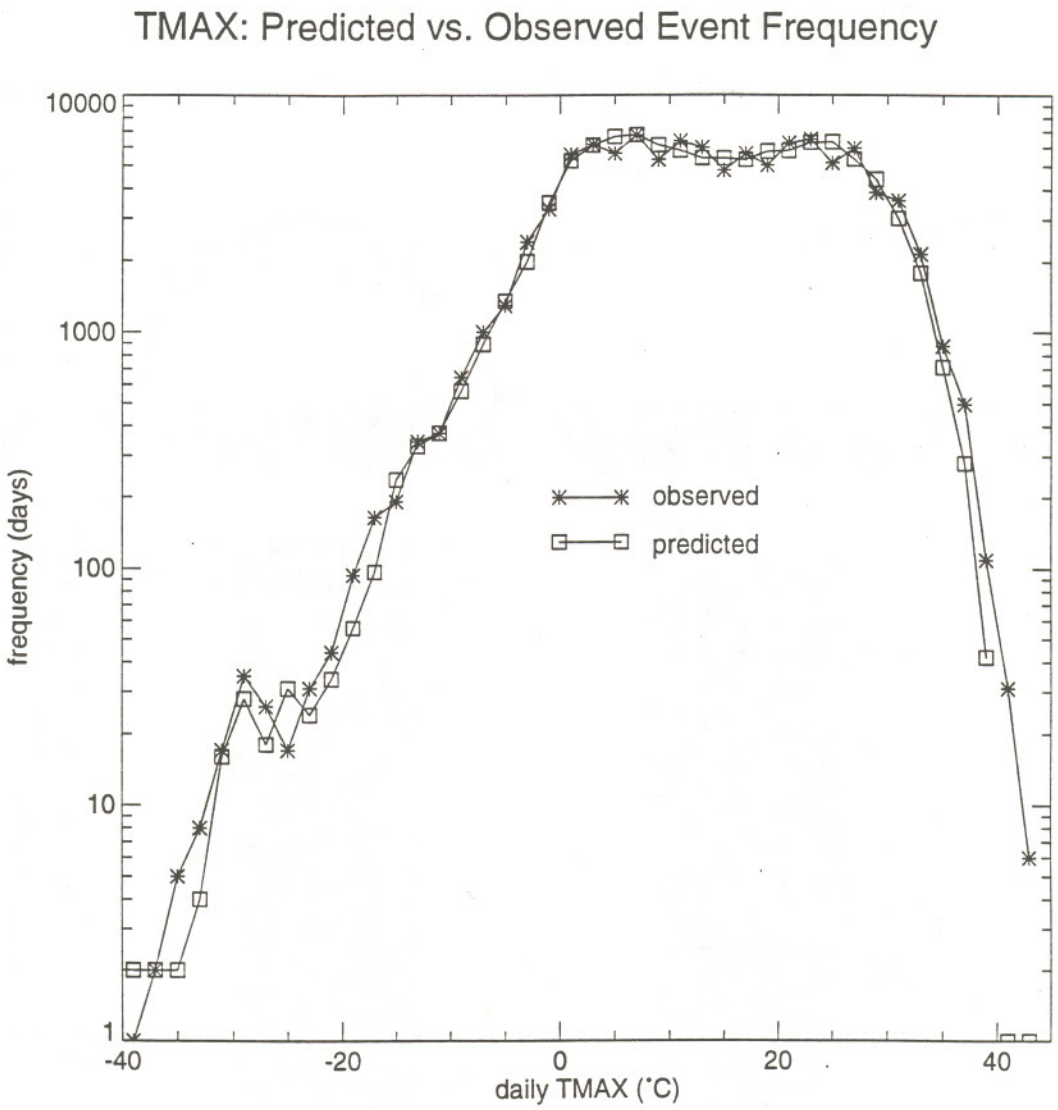


a)

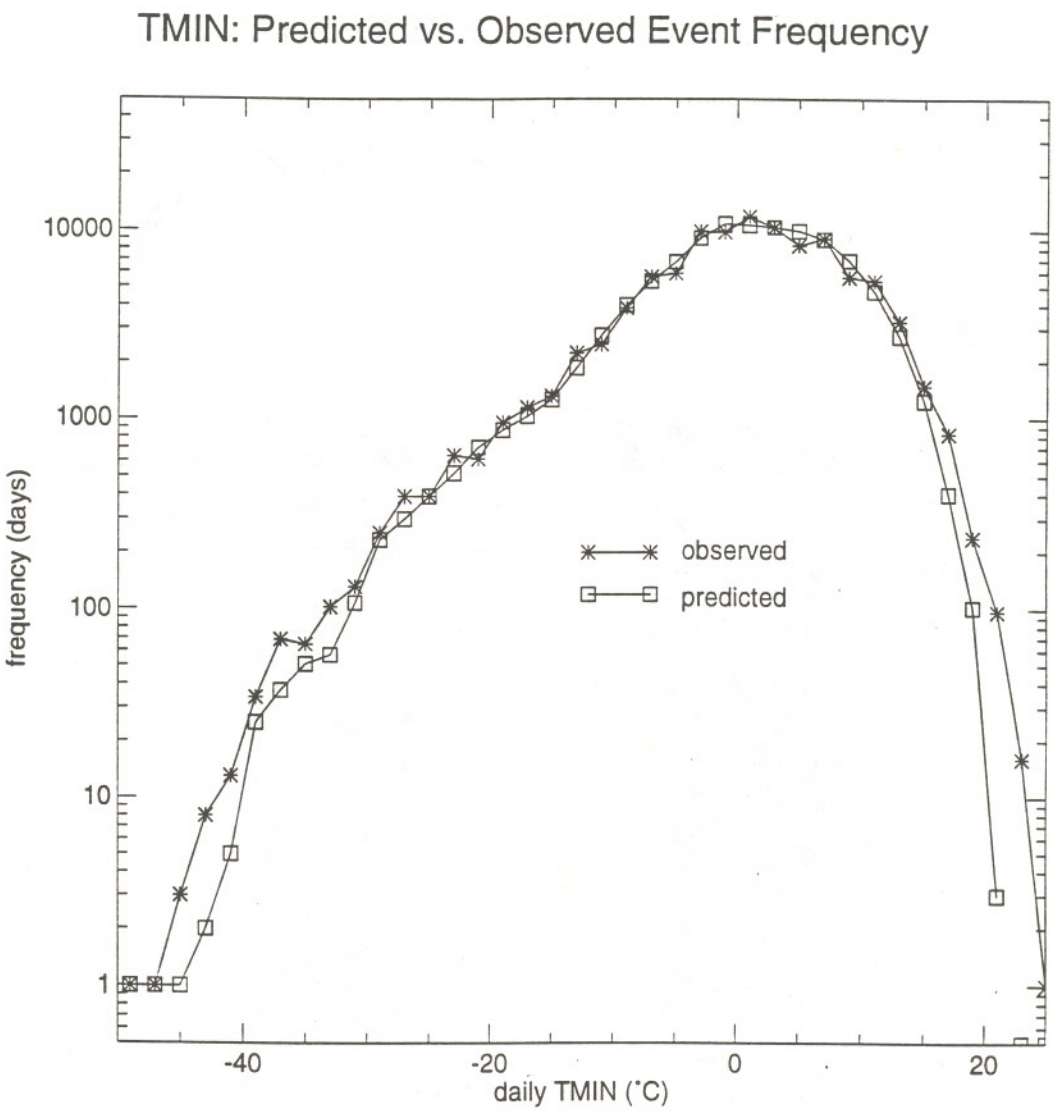
Figure 4.
Scatter-plots and frequency histograms for daily cross-validation predictions and observations of temperature extrema: a) TMAX scatterplot, b) TMIN scatterplot, c) TMAX frequency histograms, d) TMIN frequency histograms. Solid line in scatterplots shows 1:1 relationship. Vertical histogram axes are log-scaled to show detail in the ends of the distributions.


TMIN: Predicted vs. Observed Daily Values
predicted daily TMIN (˚C)
observed daily TMIN (˚C)
20
0
-20
-40
-40
-20
0
20

b)


TMAX: Predicted vs. Observed Event Frequency
10000
1000
100
10
1
frequency (days)
observed
predicted
-40
-20
0
20
40
daily TMAX (˚C)

c)


TMIN: Predicted vs. Observed Event Frequency
frequency (days)
10000
1000
100
10
1
-40
-20
0
20
daily TMIN (˚C)
observed
predicted

d)

with a small number of large events, and both cases are likely inaccurate. Assessments of both annual totals and the relationship between occurrence frequency and amount are required.

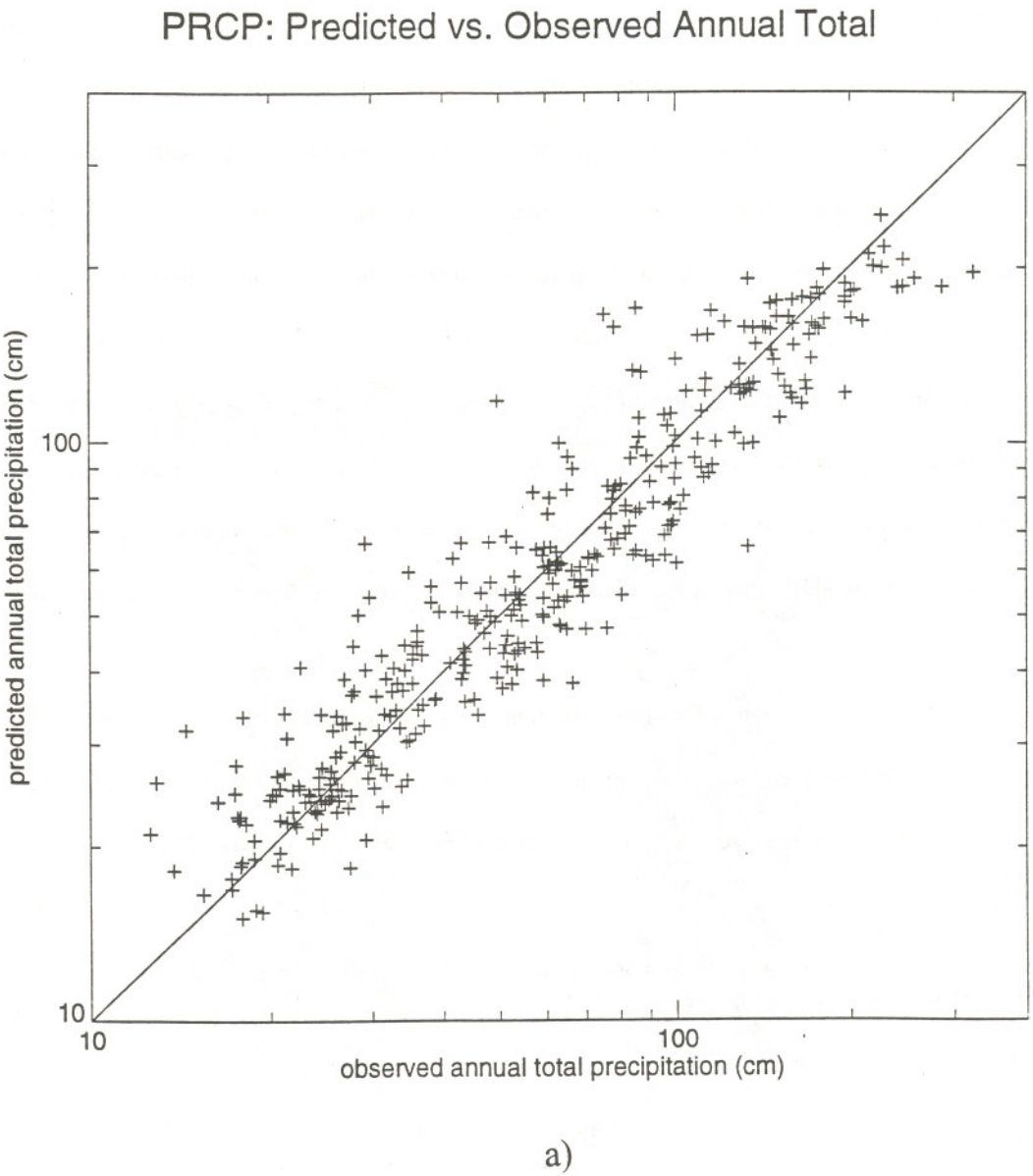
MAEs obtained from simple differences in annual totals and from percents of observed annual totals are 13.4 cm and 19.3%, respectively. Prediction errors increase with increasing observed totals, but are normally distributed on a log scale (Figure 5a). Estimated biases in annual totals are -2.1 cm and +3.0%, by the two methods, and the difference in sign between these two estimates reflects the compromise in the parameterization of POP_{crit} mentioned earlier. Frequency histograms of predicted and observed daily precipitation amounts (Figure 5b) show a small but consistent underprediction of event frequency in the middle of the range (2-8 cm/day), and an overprediction of event frequency around 1 cm/day. My methods accurately reproduce the frequency of both dry days and extreme precipitation events.

Daily precipitation occurrence predictions are summarized in Table 4. Of all simulated station-days, these methods predicted 71.5% dry and 28.5% wet, compared to observed values of 75.8% dry and 24.2% wet, with an overall success rate for occurrence prediction of 83.3% (91.4% for dry days, and 63.0% for wet days).

Table 4: Summary of daily precipitation occurrence predictions vs. observations.
For 365 stations, over a single year (1989), for a total of 133,225 station-days.

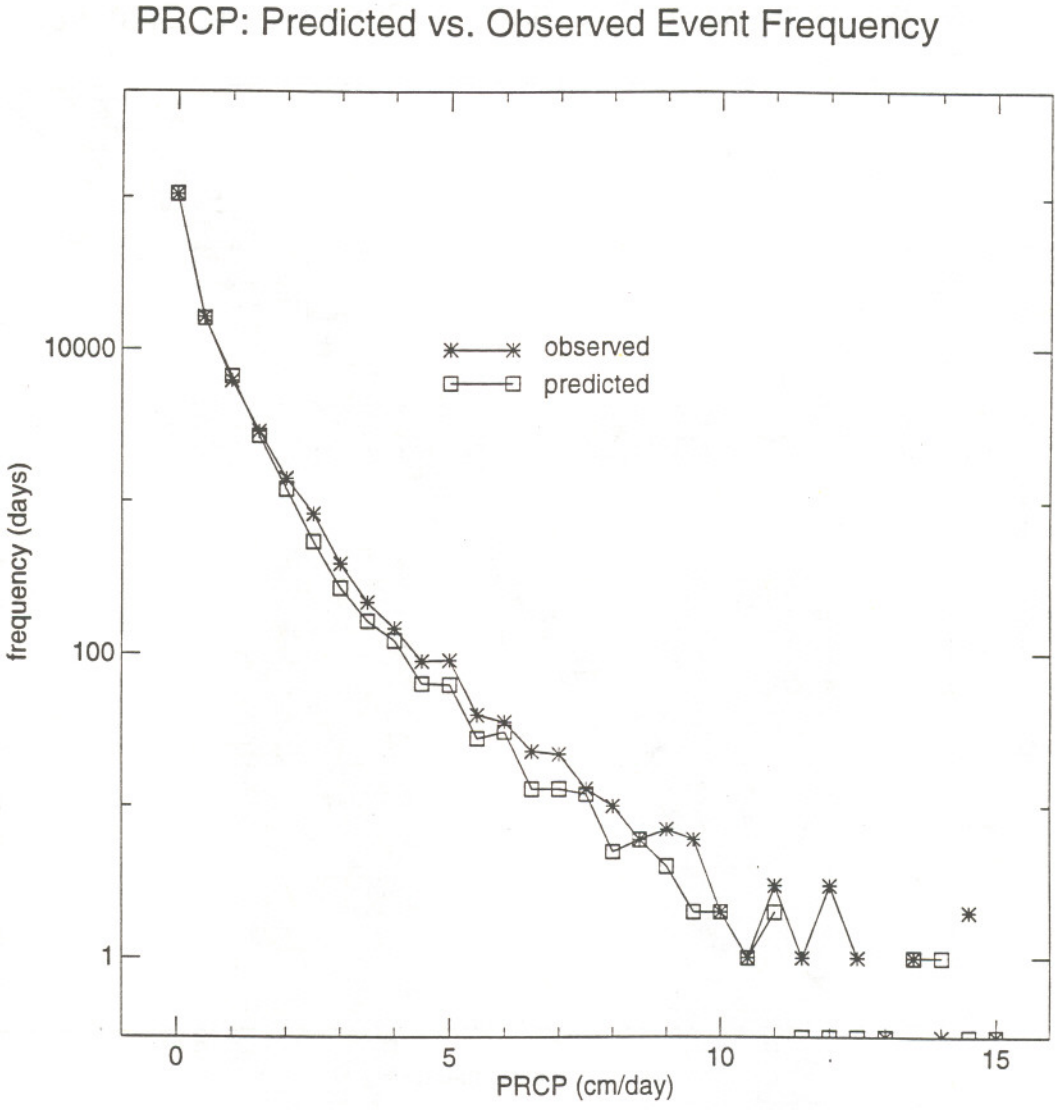
Error matrix, % of station-days		Predicted	
		Dry	Wet
Observed	Dry	65.3	10.5
	Wet	6.2	18.0

As an assessment of the ability of these methods to reproduce, for a given day, the relative proportions of wet and dry areas in a large region, the predicted vs. observed spatial frequency of dry stations is plotted in Figure 6. The frequency of dry stations is underestimated on days with widespread precipitation, and overestimated on days with scattered precipitation. The spatial frequency is predicted correctly in the middle of the range, when about half of the stations are wet and half dry. Figure 6



a)

Figure 5.
a) Scatterplot of annual total of daily cross-validation predictions and observations of precipitation, log-log scales, soild line shows 1:1 relationship. b) Frequency histogram of daily cross-validation predictions and observations of precipitation, vertical axis is log-scaled to show detail for extreme events.


PRCP: Predicted vs. Observed Event Frequency
observed
predicted
frequency (days)
10000
100
1
0
5
10
15
PRCP (cm/day)

b)

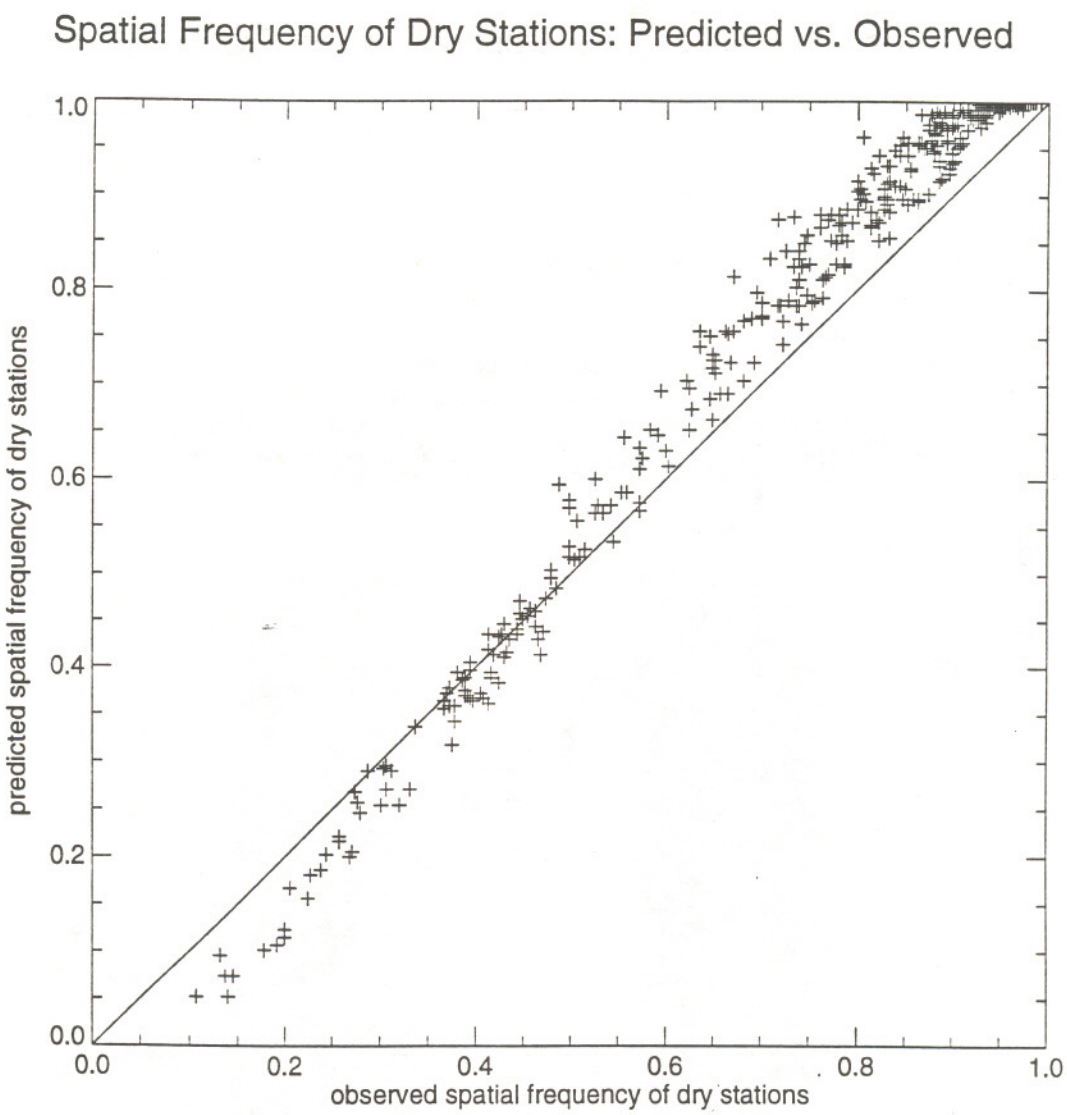


Figure 6.
Predicted vs. observed spatial frequency of dry stations (scaled as a proportion of all stations); one point per day.

illustrates the accuracy of the average predicted occurrence distribution over the study area on any day, but does not provide any information on the accuracy of the predicted spatial distribution of occurrence. Plotting the spatial frequency of correct occurrence predictions against the observed spatial frequency of dry stations (Figure 7) shows that predicted spatial distributions of occurrence are most accurate on very wet and very dry days, and suffer in the middle range where wet and dry stations are mixed evenly.

As an indication of the accuracy with which these methods reproduce the observed relationship between daily areal coverage of precipitation and daily average precipitation intensity, I plot observations and predictions of the daily total precipitation (averaged over stations) vs. the daily proportion of wet stations. The predicted distribution matches the observations very well, with the exception of the underprediction of the proportion of wet stations on very dry days (Figure 8).

3.3. Spatial and Temporal Scaling Analysis

I found that areal means (for temperature predictions) or totals (for precipitation) over the study region are preserved for a wide range of prediction grid resolutions. Areal mean annual average TMAX and TMIN of 12.9 and -6.7 °C, respectively, and areal mean annual total PRCP of 65.2 cm were obtained for independent simulations over nested prediction grids ranging in grid-point spacing from 500 m to 32 km. A shift to larger grid spacing reduced the range of variation over the grid, as a result of reduced variability in the elevations represented by the grid points.

Temporal scaling of the temperature outputs is perfect for all time steps larger than one day: the exact same result is obtained by generating daily predictions from daily observations and averaging over a longer time period as is obtained by generating predictions for the longer time period directly from averaged observations for the period. This result is a consequence of the linear nature of the prediction algorithms. Temporal scaling of the precipitation algorithms is confounded somewhat by the binary predictions of precipitation occurrence at the daily timestep. For predictions of annual totals from observed annual totals, cross-validation MAEs were 12.2 cm and 18.5 %, lower by 1.2 cm and 0.8 % than MAEs obtained from totals of daily predictions. These differences between the two time scales of prediction are significant at the 0.01 level for both methods of estimating MAE, and provide a crude estimate of the

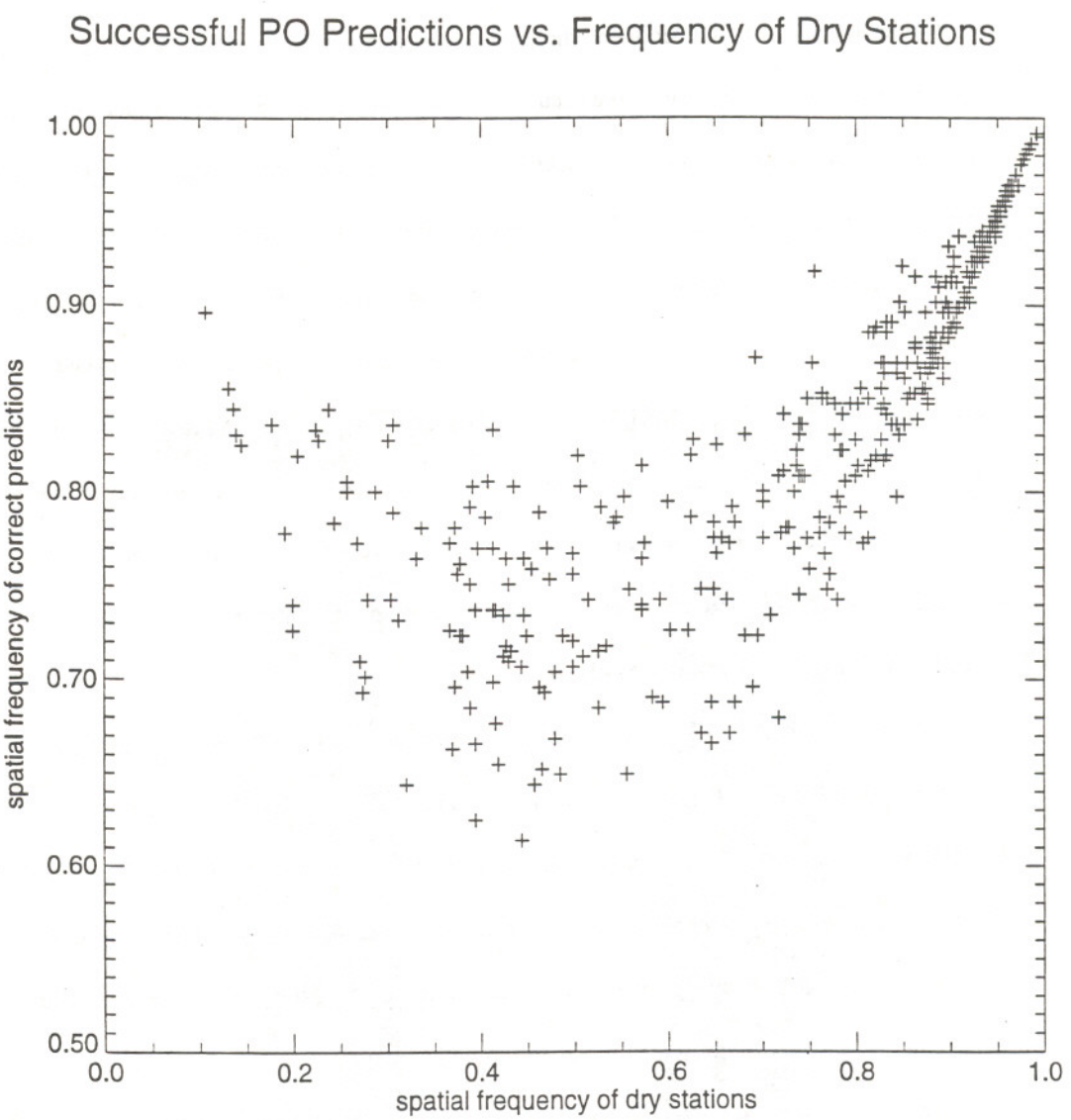


Figure 7.
Spatial frequency of correct precipitation occurrence predictions plotted against the spatial frequency of dry stations (frequencies scaled as proportions of all stations); one point per day.

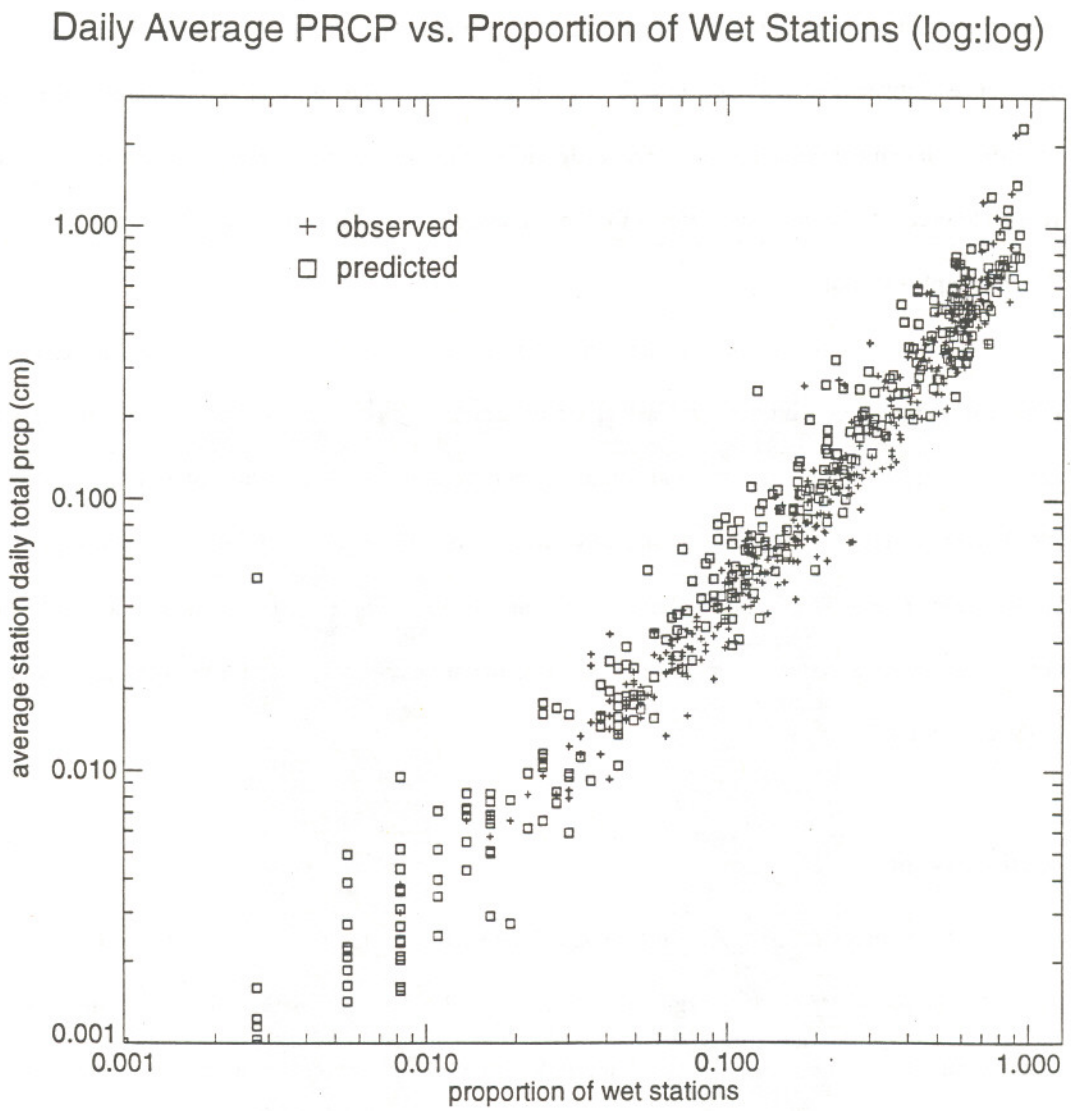


Figure 8.
Precipitation intensity vs. areal coverage for spatial averages of predictions and observations on each day (log-log scales).

contribution of occurrence prediction error to the daily prediction errors. There is no trend in the differences between the daily and annual predictions over the range of predicted annual totals (Figure 9), an indication that the prediction occurrence algorithm is not introducing significant biases which are related to precipitation intensity. Biases for the annual predictions are +0.3 cm and +7.1 %.

3.4. Example Output

It is not possible to illustrate here the daily sequences of predicted surfaces, and instead I have provided spatial output summaries. I selected a prediction grid point spacing of 2 km for these example results. As an example of the temporal variability in the diagnosed relationships of TMAX, TMIN, and PRCP to elevation, time series of the spatially averaged regression slopes (β_1) are shown in Figure 10. Annual aggregates of the predicted daily surfaces are shown in Figure 11 (color plates), with TAVG, VPD and SRAD shown as annual averages, PO shown as the number of wet days for the year, and PRCP shown as the annual total.

4. Discussion

The methods described and implemented here are essentially ad hoc, in that their design has been guided by the particular needs of regional hydroecological process simulation. However, since one of these needs is that the methods be applicable to multiple variables and purposes (e.g. the use of the same interpolation weights to drive predictions of both precipitation occurrence and intensity), I arrive at the curious case of an ad hoc methodology which is also quite general. Much of this generality stems from a conscious decision to allow the observations to dictate, to a large extent, the interpolation and extrapolation parameters. For example, while Daly et al. (1994) have stressed the importance of an explicit accounting of the influence of leeward and windward aspect in distributing precipitation in mountainous terrain, I find that these methods faithfully reproduce the extreme differences in precipitation gradient on the west and east sides of the Cascade range without recourse to anisotropic filtering criteria (Figure 11e). This is not to suggest that prevailing wind conditions are an unimportant component of precipitation distribution, but

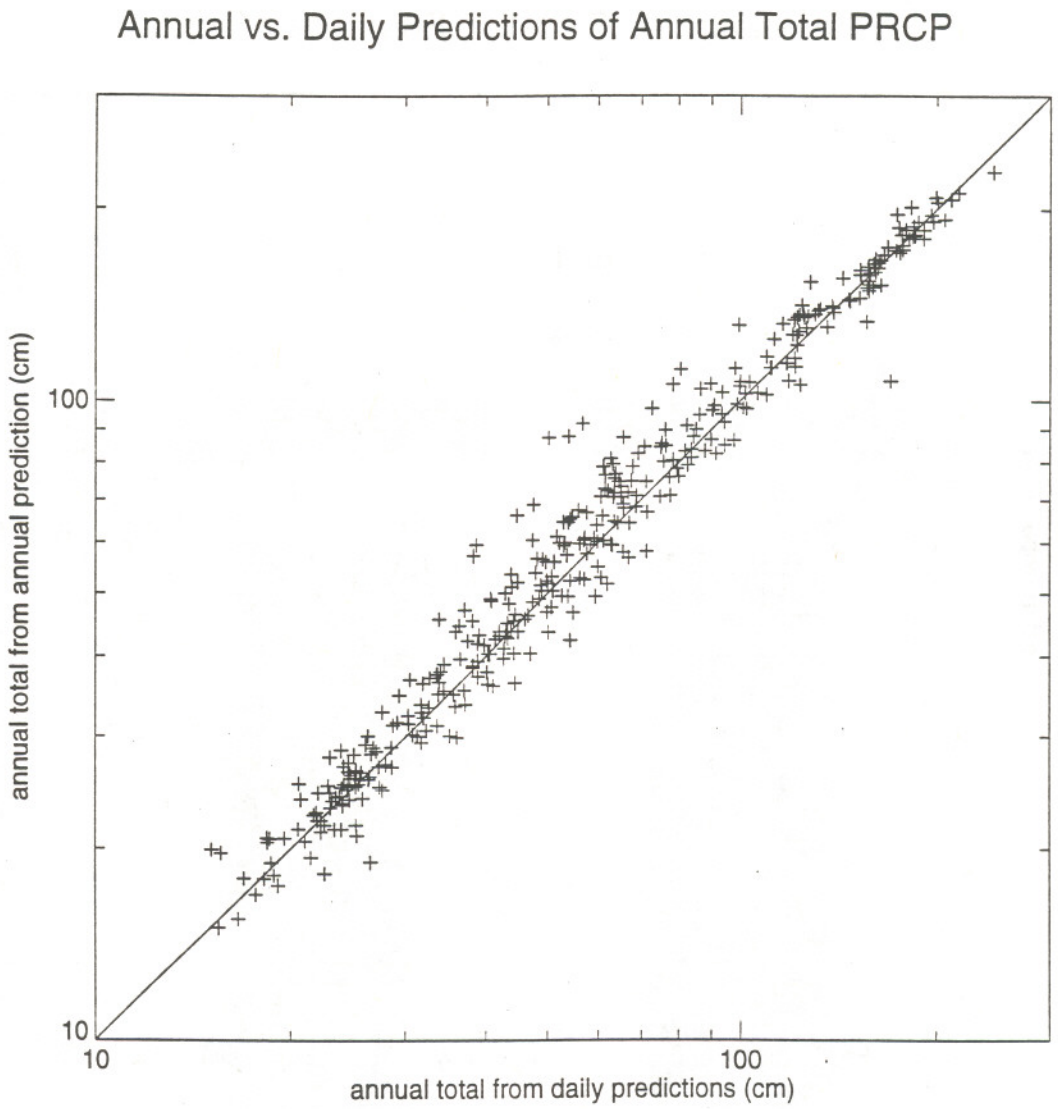
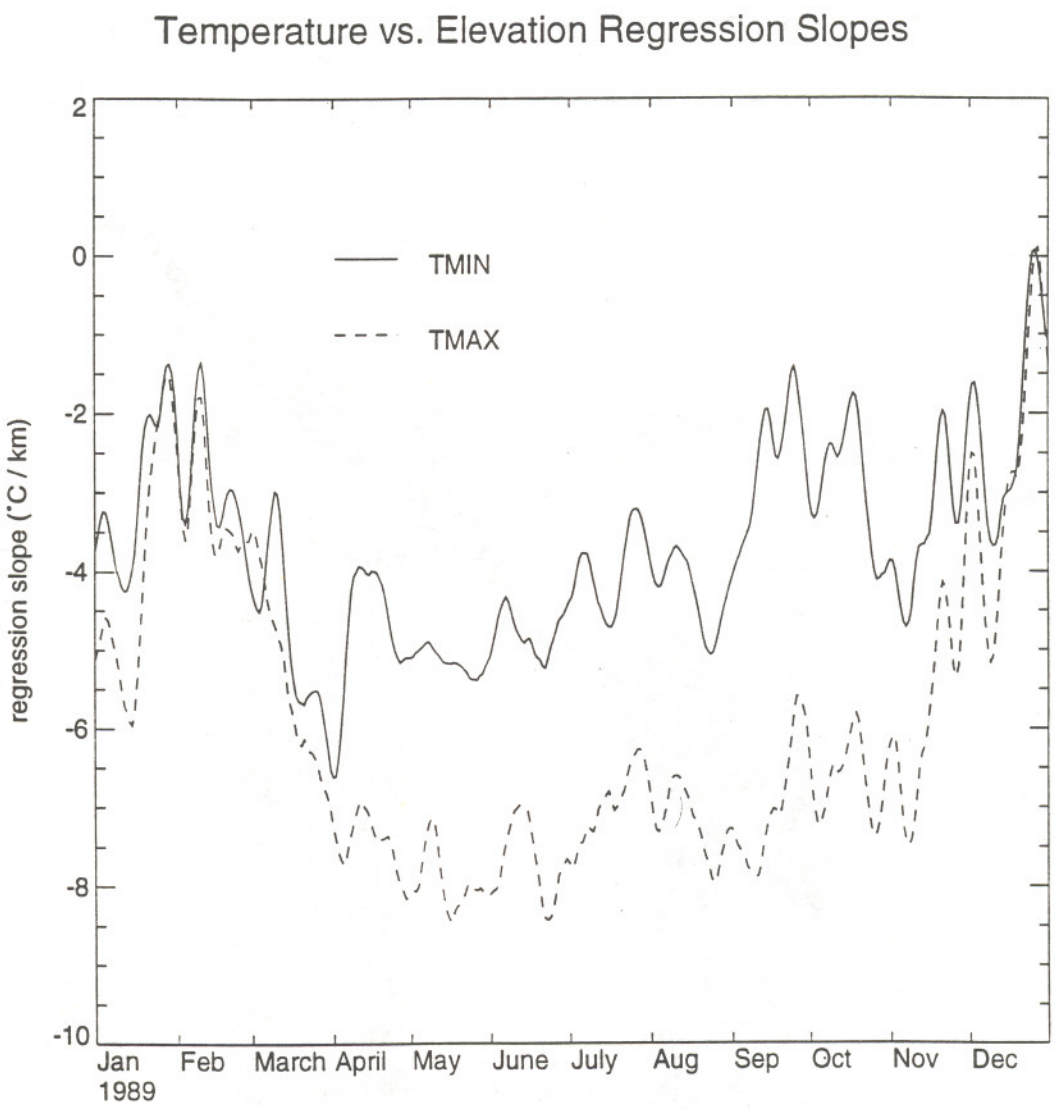
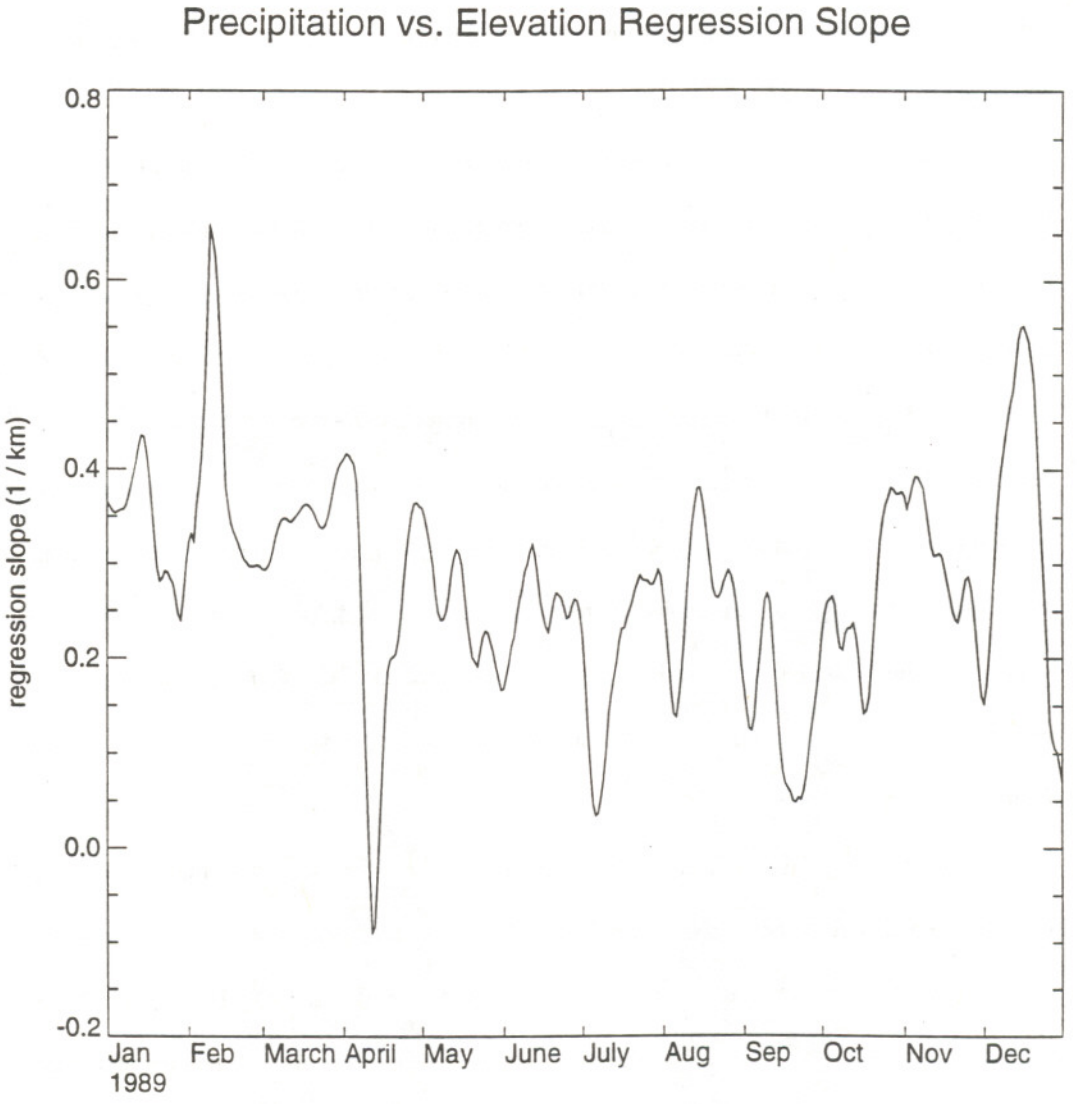


Figure 9.
Influence of temporal scaling on predictions of annual total precipitation: annual totals from daily predictions plotted against annual totals from a single annual prediction (log-log scales).



a)

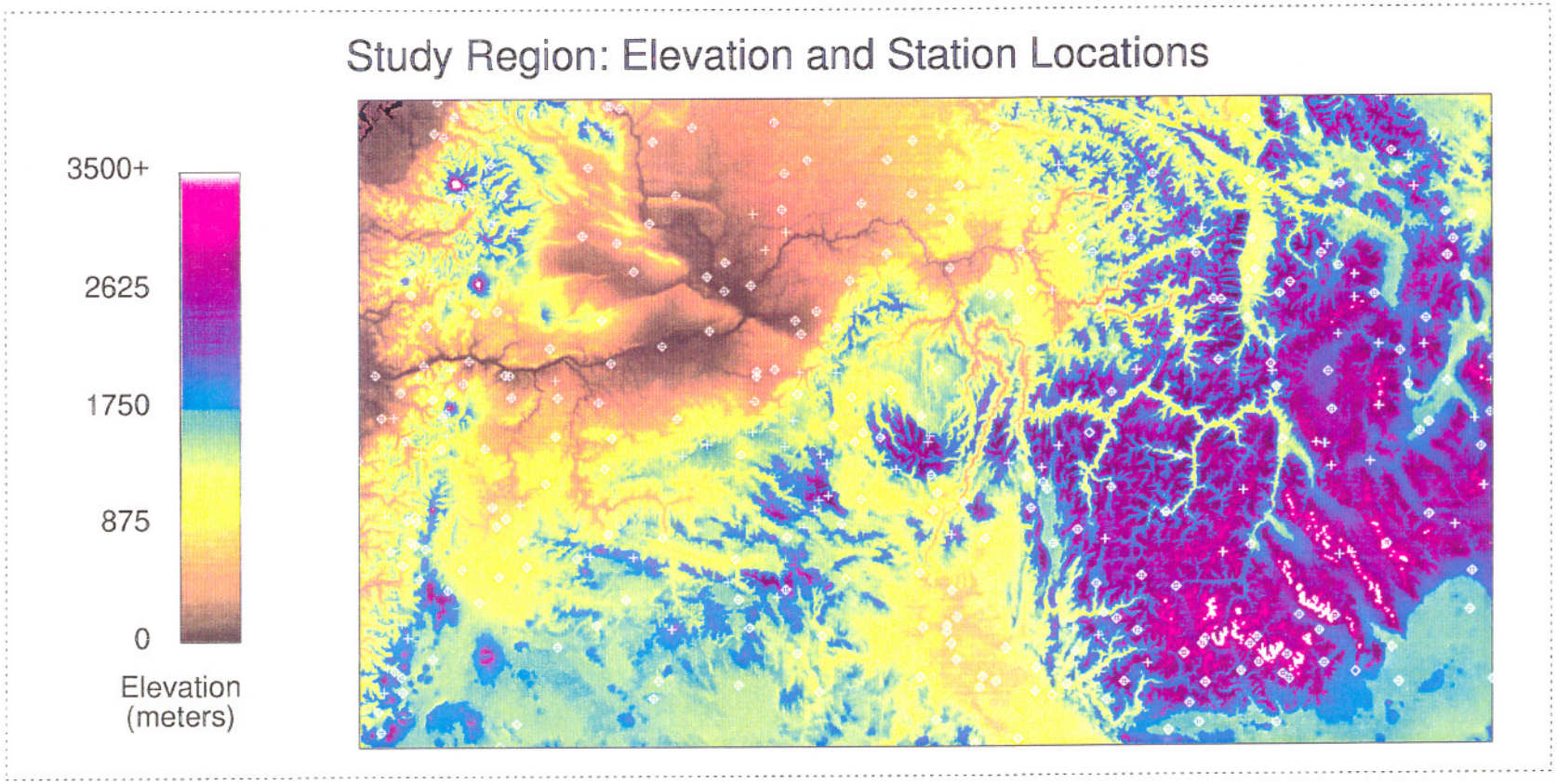
Figure 10.
Time series of spatially averaged regression slopes (β_1) for a) TMAX and TMIN, and b) PRCP. Only grid points for which the weighted average difference in elevation between observation and grid-cell elevation exceeds 200 m are included in this average. This restricts the spatial average to areas of complex terrain. For clarity, an 11-day tapered smoothing filter has been applied to these time series.


Precipitation vs. Elevation Regression Slope
regression slope (1 / km)
0.8
0.6
0.4
0.2
0.0
-0.2
Jan
1989
Feb
March
April
May
June
July
Aug
Sep
Oct
Nov
Dec

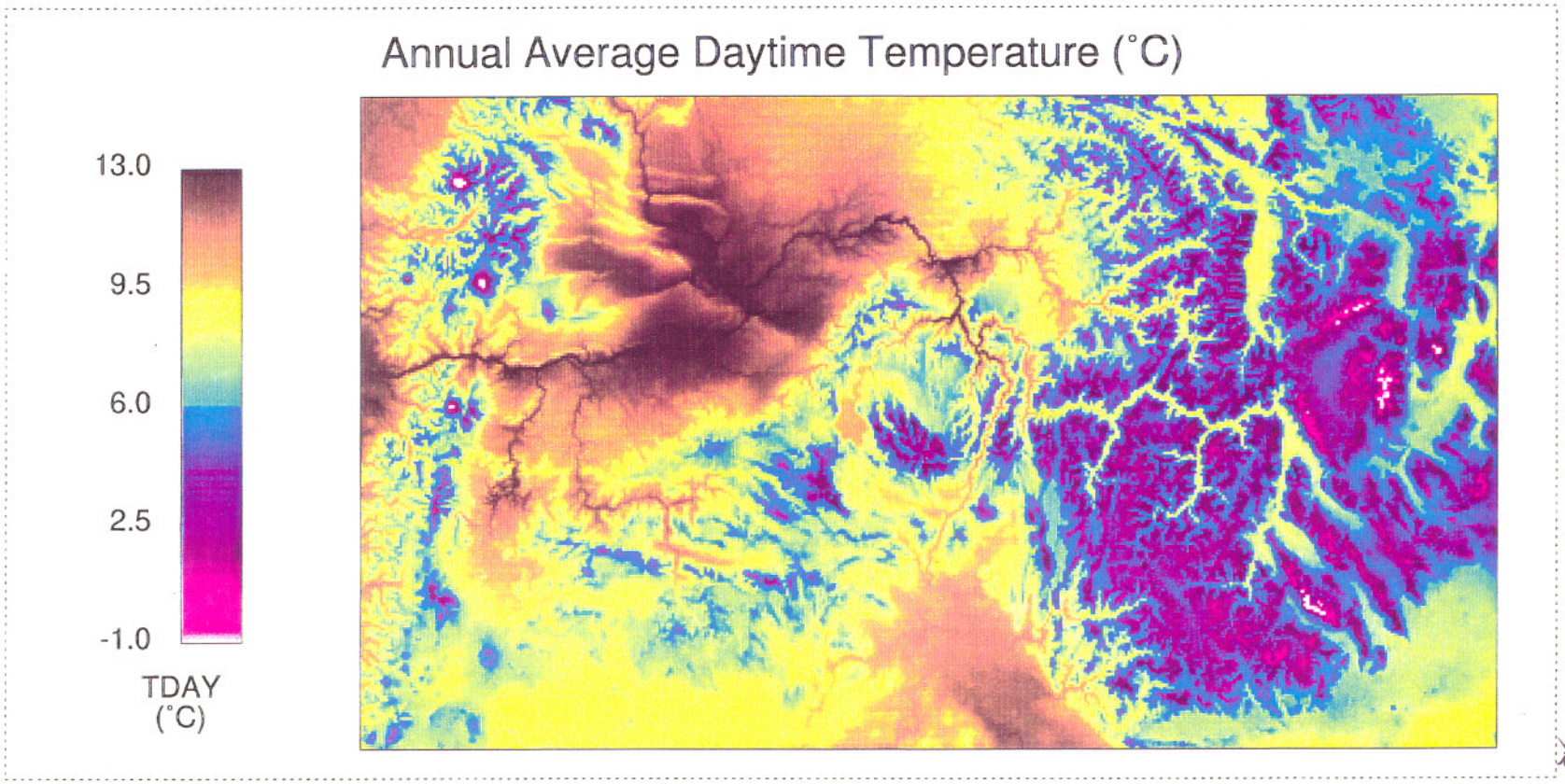
b)

Figure 11.
a) Detail of study region, showing topography and station locations. Crosses indicate PRCP stations, boxes indicate TMAX/TMIN stations, and overlapping symbols indicate the observation of all variables at the site. Major topographic features are the Cascade Range on the far western edge of the region, Blue Mtns. center, Salmon River Mountains, Bitterroot range, and other associated ranges of the Northern Rocky Mtns. to the east, and the Snake River Plain south-east to south-central. Remaining color plates show annual averages or totals of selected outputs: b) annual average daytime temperature, c) annual average daytime VPD, d) annual total of wet days, e) annual total precipitation, and f) annual average of the daylight average incident shortwave radiative flux density. Scale bars shown for figures c) and f) apply to all figures.

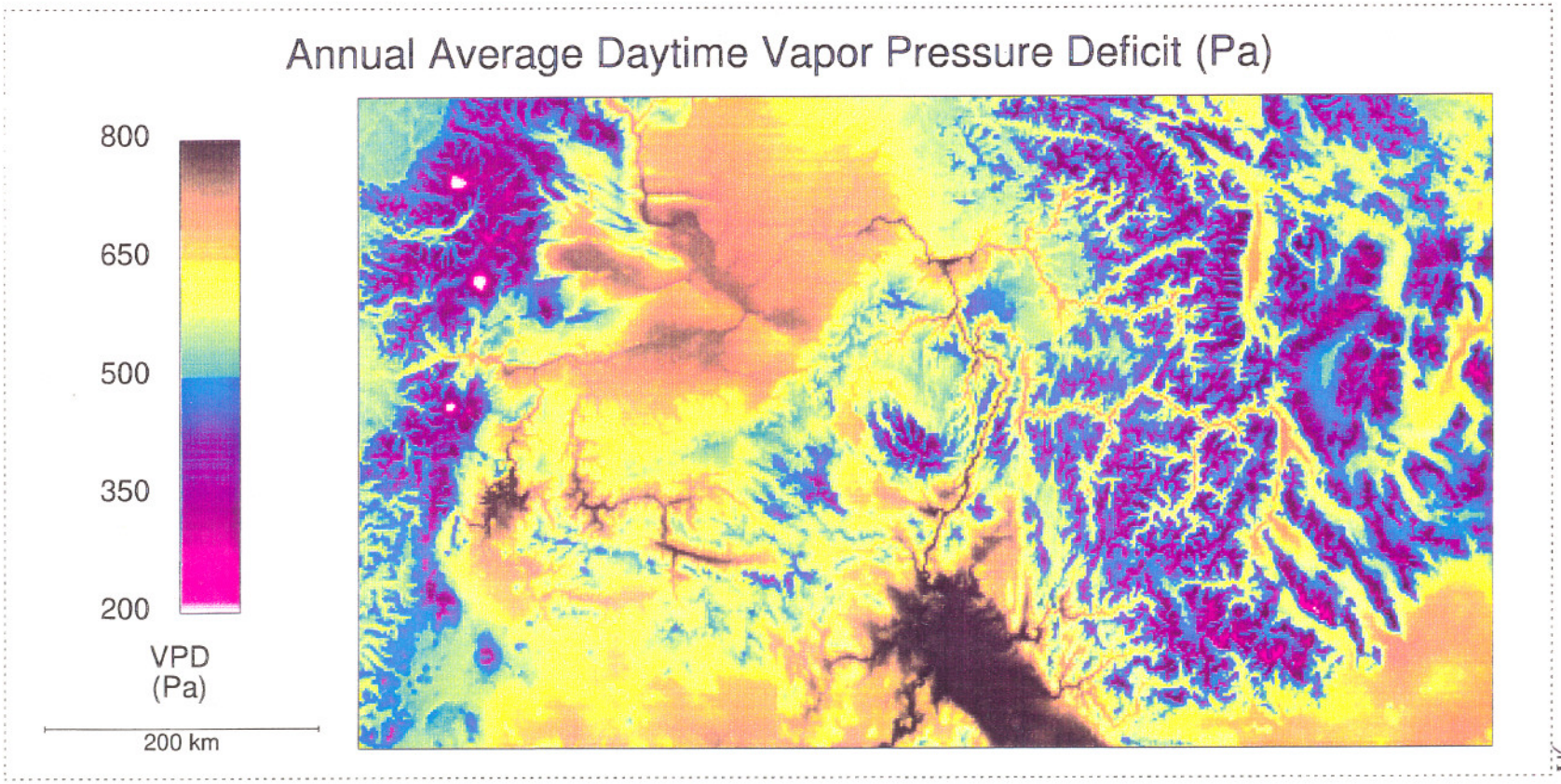
a)



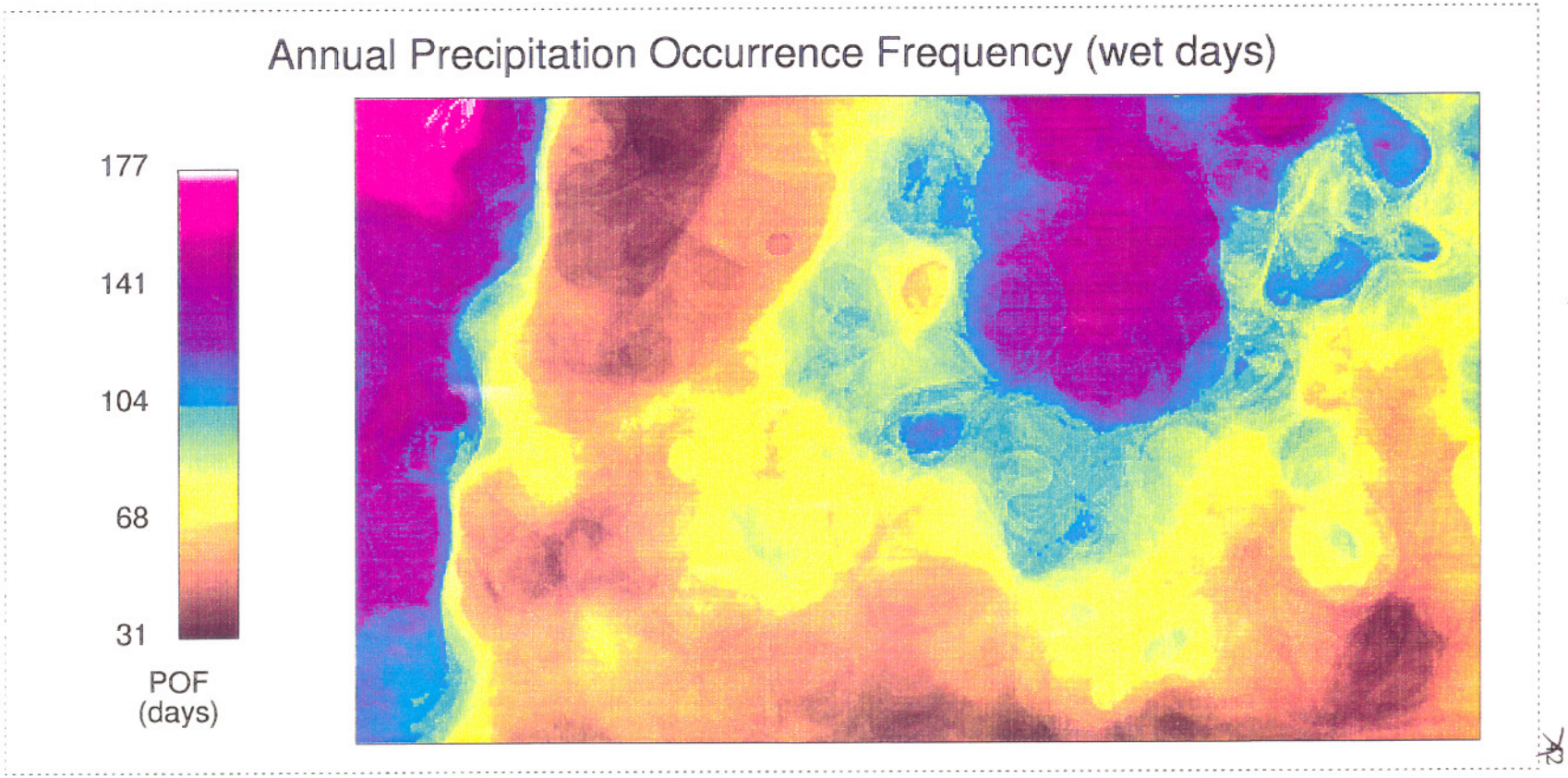
b)



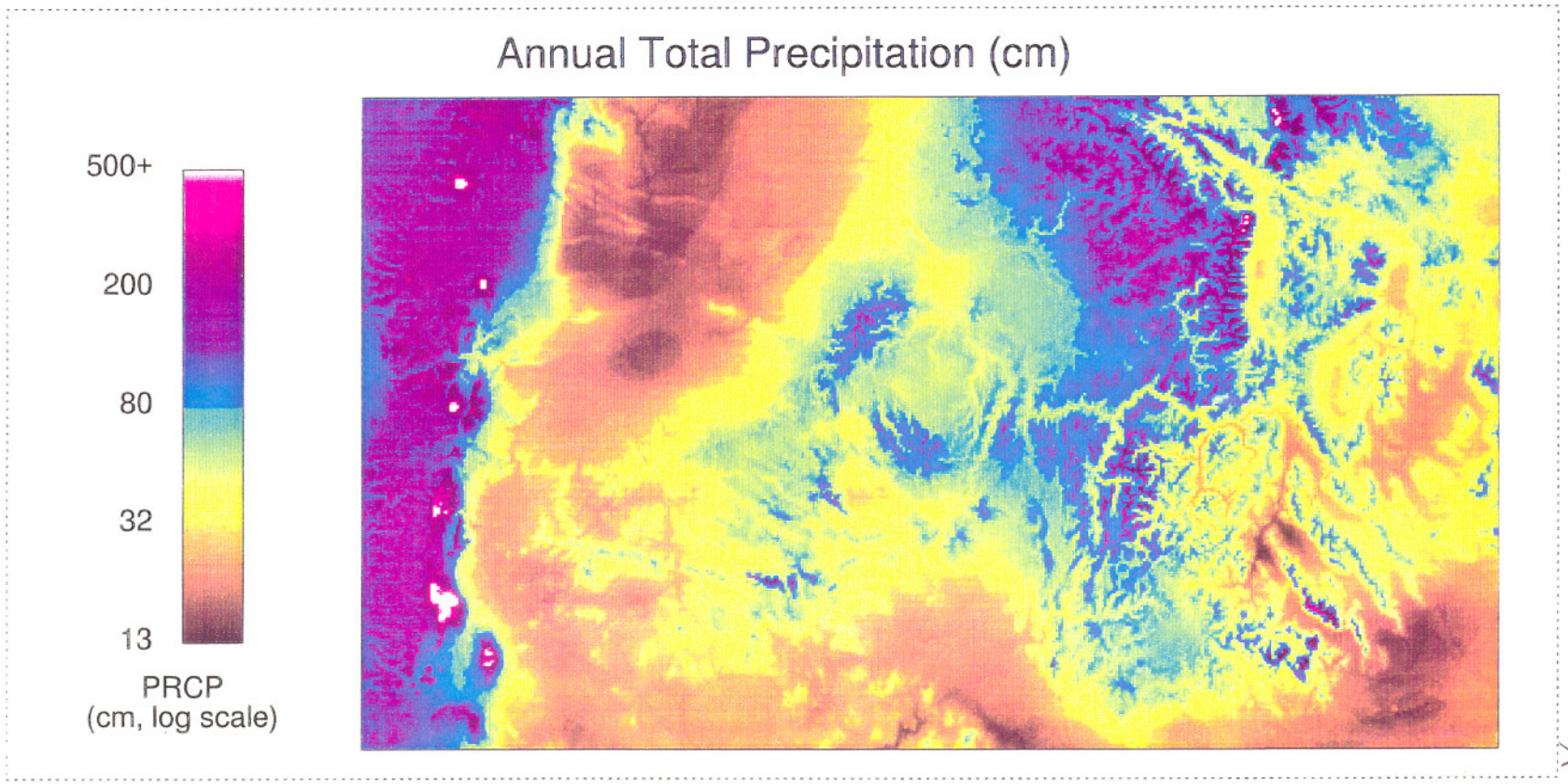
c)


Annual Average Daytime Vapor Pressure Deficit (Pa)
800
650
500
350
200
VPD
(Pa)
200 km

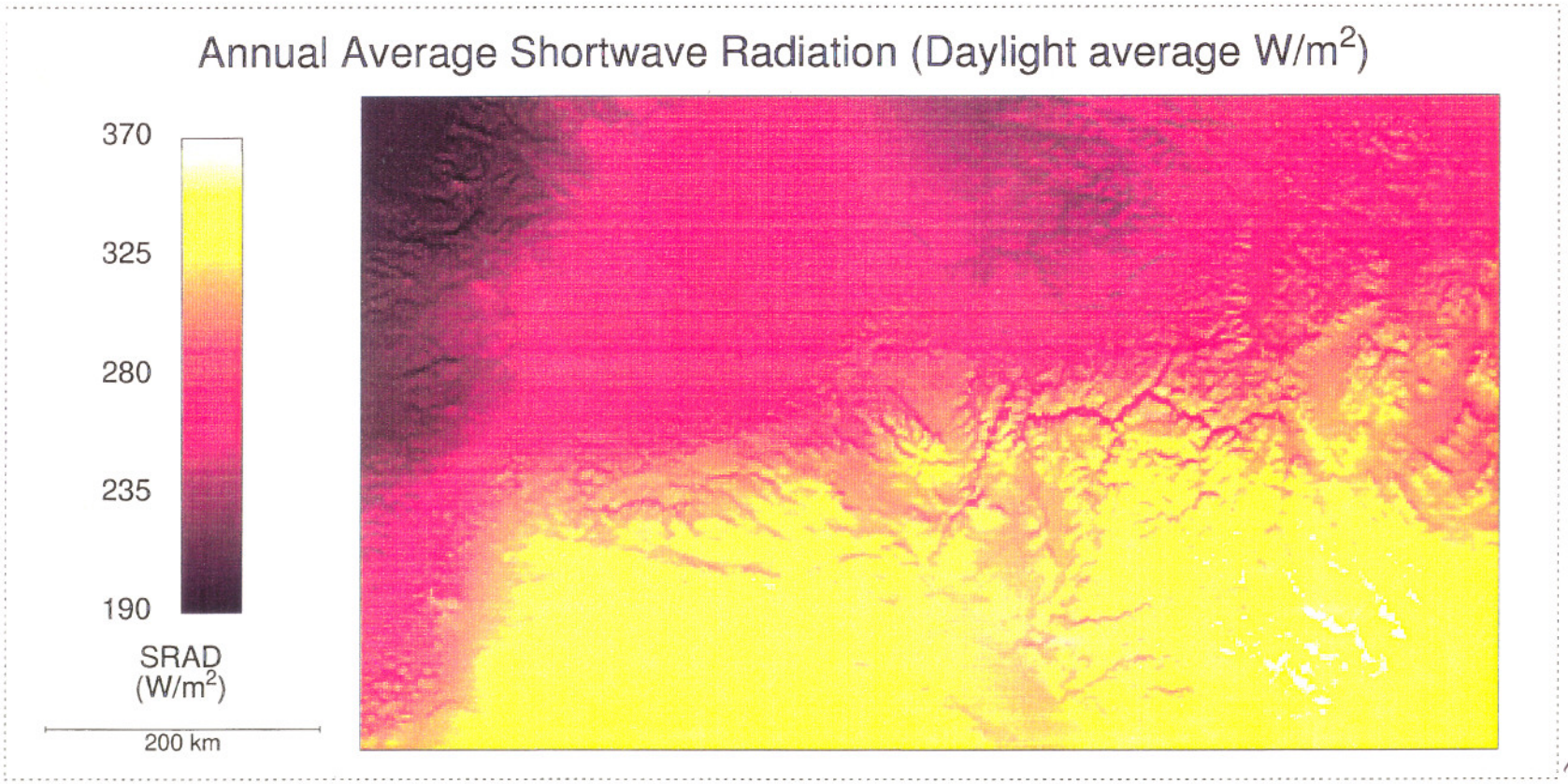
d)


Annual Precipitation Occurrence Frequency (wet days)
177
141
104
68
31
POF
(days)

e)


Annual Total Precipitation (cm)
500+
200
80
32
13
PRCP
(cm, log scale)

f)


Annual Average Shortwave Radiation (Daylight average W/m2)
370
325
280
235
190
SRAD
(W/m2)
200 km

rather to call attention to the ability of simple, isotropic methods to extract such information from a topographic neighborhood.

The same empirical arguments for simplicity can also be made with respect to the relationship between precipitation occurrence (or areal coverage) and intensity. In a discussion of stochastic precipitation predictions, Hutchinson (1995) notes the disparity in spatial and temporal scales of correlation between event-based precipitation occurrence and intensity, and suggests that these processes should be treated separately in spatial interpolations. I find, however, that an extremely simple abstraction of occurrence likelihood and intensity, derived from the same spatial and temporal interpolation parameters, gives quite reasonable results for a daily timestep (Figures 5b, 6, and 8). These results are due in part to the use of observed as opposed to modeled or stochastic sequences of daily precipitation. Even at densities which would generally be considered too low to resolve important spatial precipitation features, enough information is retained to produce realistic daily time series of both occurrence and intensity from a single, point-wise isotropic interpolation.

The point-wise isotropic nature of these methods is in contrast to its spatially varying scaling properties. Sensitivity to local observation density, by way of the iterative station density algorithm, is a trait my method shares with the smoothing splines method as implemented by Hutchinson (1995), the recursive filter objective analysis of Hayden and Purser (1995), and to some extent the topographic facet logic of Daly et al. (1994). I suggest that it is this feature which most distinguishes these methods from other, both sophisticated and simple, methods.

The results of my tests of the characteristic spatial and temporal smoothing scales associated with regressions against elevation are in agreement with results reported for precipitation-elevation regressions by Daly et al. (1994), who suggest an optimal DEM cell-size of 4-10 km, compared with my result of 2-8 km for the smoothing width. I have not found any relevant studies with which to compare my results for temporal smoothing width, nor have I found any relevant references for either spatial or temporal smoothing parameters for temperature regressions. These regressions appear to be quite sensitive to local topography, and to atmospheric conditions with short time scales (Figure 10). It is worth noting that the

average regression slope for TMAX is -6.0 °C/km, which agrees well with general observations of environmental lapse rates and with the default lapse rate employed in the original MTCLIM logic (Running et al., 1987). I plan to examine the relationship of these diagnosed slopes to synoptic atmospheric conditions. Temporal sequences of predicted temperature surfaces show the distinct passage of fronts, and I hope to derive a relationship between frontal position and characteristic regression slopes using a logic similar to the synoptic classification scheme described by Pielke et al. (1987).

The results of the spatial and temporal scaling exercises are encouraging, although the exercises themselves are not very sophisticated. I can be reasonably certain that the choice of prediction grid resolution will not have any noticeable effect on areal averages or totals of the primary variables when measured at scales considerably larger than the grid resolution. For example, if only very coarse spatial outputs are required, a widely spaced grid will give the same result as a finer and more computationally expensive grid, given that care is taken in the translation of the elevation data between grids. The temporal scaling properties of the temperature prediction algorithms eliminate any uncertainty associated with predictions at different time-steps. The close agreement between predicted annual total precipitation at daily and annual timesteps (Figure 9) gives confidence in the application of these methods in a climatological mode, but a more detailed analysis is warranted.

Spatially distributed and relatively accurate near-surface air temperatures, in combination with remotely sensed surface temperature, can provide estimates of surface resistances (Reginato et al., 1985, Seguin et al., 1994). Such estimates, based on observation, can be used to validate and refine spatially explicit estimates of surface resistance that are derived from simulations. The combination of surface and satellite observations should allow a regional assessment of evaporation and resultant soil moisture stores as presented by Saha (1995), and could lead to improvements in the initialization of surface moisture fields for coupled atmosphere-terrestrial ecosystem simulations (Pielke et al., 1993). Ten years ago, Eagleson (1986) reported that macroscale field observations were limiting the advance of hydrological science: I see the methods presented here as an attempt to overcome such limitations.

5. Literature cited

Band, L.E., Peterson, D.L., Running, S.W., Coughlan, J.C., Lammers, R., Dungan, J., and Nemani, R.R., 1991. Forest ecosystem processes at the watershed scale: Basis for distributed simulation. Ecol. Modelling, 56:171-196.

Band, L.E., Patterson, P., Nemani, R.R., and Running, S.W., 1993. Forest ecosystem processes at the watershed scale: Incorporating hillslope hydrology. Agric. For. Meteorol., 63:93-126.

Bristow, K.L., and Campbell, G.S., 1984. On the relationship between incoming solar radiation and daily maximum and minimum temperature. Agric. For. Meteorol., 31:159-166.

Chua, S.-H., and Bras, R.L., 1982. Optimal estimators of mean areal precipitation in regions of orographic influence. J. Hydrol., 57:23-48.

Creutin, J.D., and Obled, C., 1982. Objective analyses and mapping techniques for rainfall fields: An objective comparison. Water Resour. Res., 18:413-431.

Daly, C., Neilson, R.P., and Phillips, D.L., 1994. A statistical-topographic model for mapping climatological precipitation over mountainous terrain. J. App. Meteor., 33:140-158.

DeGaetano, A.T., Eggleston, K.L., and Knapp, W.W., 1995. A method to estimate missing daily maximum and minimum temperature observations. J. App. Meteor., 34:371-380.

Dolph, J., and Marks, D., 1992. Characterizing the distribution of observed precipitation and runoff over the continental United States. Climatic Change, 22:99-119.

Eagleson, P.S., 1986. The emergence of global-scale hydrology. Water Resour. Res., 22:6S-14S.

Forsythe, W.C., Rykiel, E.J., Stahl, R.S., Wu, H., and Schoolfield, R.M., 1995. A model comparison for daylength as a function of latitude and day of year. Ecol. Modelling, 80:87-95.

Gandin, L.S., 1965. Objective analysis of meteorological fields. Israel Program for Scientific Translations, Jerusalem, 242 pp.

Glassy, J.M., and Running, S.W., 1994. Validating diurnal climatology logic of the MT-CLIM model across a climatic gradient in Oregon. Ecol. Appl., 4:248-257.

Hayden, C.M., and Purser, R.J., 1995. Recursive filter analysis of meteorological fields: Applications to NESDIS operational processing. J. App. Meteor., 34:3-15.

Hevesi, J.A., Istok, J.D., and Flint, A.L., 1992. Precipitation estimation in mountainous terrain using multivariate geostatistics. Part I: Structural Analysis. J. App. Meteor., 31:661-676.

Hungerford, R.D., Nemani, R.R., Running, S.W., and Coughlan, J.C., 1989. MTCLIM: A mountain microclimate simulation Model. Research Paper INT-414. USDA Forest Service, Intermountain Research Station, Ogden.

Hutchinson, M.F., and Bischof, R.J., 1983. A new method for estimating the spatial distribution of mean seasonal and annual rainfall applied to the Hunter Valley, New South Wales. Austr. Met. Mag., 31:179-184.

Hutchinson, M.F., 1995. Stochastic space-time weather models from ground-based data. Agric. For. Meteorol., 73:237-264.

Kimball, J., Running, S.W., and Nemani, R.R, 1995. An improved method for estimating surface humidity from daily minimum temperature. Agric. For. Meteorol., accepted.

McMurtrie, R.E., Leuning, R., Thompson, W.A., and Wheeler, A.M., 1992. A model of canopy photosynthesis and water use incorporating a mechanistic formulation of leaf CO_2 exchange. For. Ecol. Manage., 52:261-278.

Murray, F.W., 1967. On the computation of saturation vapor pressure. J. Appl. Meteor., 6:203-204.

Nemani, R.R., Pierce, L., Band, L.E., and Running, S.W., 1993. Forest ecosystem processes at the watershed scale: Sensitivity to remotely sensed leaf area index observations. Int. J. Remote Sens., 14:2519-2534.

Phillips, D.L., Dolph, J., and Marks, D., 1992. A comparison of geostatistical procedures for spatial analysis of precipitation in mountainous terrain. Agric. For. Meteorol., 58:119-141.

Pielke, R.A., Garstang, M., Lindsey, C., and Gusdorf, J., 1987. Use of a synoptic classification scheme to define systems. Theor. Appl. Climatol., 38:57-68.

Pielke, R.A., Schimel, D.S., Lee, T.J., Kittel, T.G.F., and Zeng, X., 1993. Atmosphere-terrestrial ecosystem interactions: Implications for coupled modeling. Ecol. Modelling, 67:5-18.

Reginato, R.J., Jackson, R.D., and Pinter, P.J., 1985. Evapotranspiration calculated form remote multispectral and ground station meteorological data. Remote Sens. Environ., 18:75-89.

Running, S.W., 1994. Testing FOREST-BGC ecosystem process simulations across a climatic gradient in Oregon. Ecol. Appl., 4:238-247.

Running, S.W., and Coughlan, J.C., 1988. A general model of forest ecosystem processes for regional applications. I. Hydrologic balance, canopy gas exchange, and primary production processes. Ecol. Modelling, 42:125-154.

Running, S.W., and Nemani, R.R., 1991. Regional hydrologic and carbon balance responses of forests resulting from potential climate change. Climatic Change, 19:349-368.

Running, S.W., Nemani, R.R., and Hungerford, R.D., 1987. Extrapolation of synoptic meteorological data in mountainous terrain and its use for simulating forest evapotranspiration and photosynthesis. Can J. For. Res., 17:472-483.

Saha, S.K., 1995. Assessment of regional soil moisture conditions by coupling satellite sensor data with a soil-plant system heat and moisture balance model. Int. J. Remote Sens., 16:973-980.

Seguin, B., Courault, D., and Guérif, M., 1994. Surface temperature and evapotranspiration: Application of local scale methods to regional scales using satellite data. Remote Sens. Environ., 49:287-295.

Shuttleworth, W.J., 1988. Macrohydrology - the new challenge for process hydrology. J. Hydrol., 100:31-56.

Tabios, G.Q., and Salas, J.D., 1985. A comparative analysis of techniques for spatial interpolation of precipitation. Water Resour. Bull., 21:365-380.

Thiessen, A.H., 1911. Precipitation averages for large areas. Mon. Wea. Rev., 39:1082-1084.

Troch, P.A., Mancini, M., Paniconi, C., and Wood, E.F., 1993. Evaluation of a distributed catchment scale water balance model. Water Resour. Res., 29:1805-1817.

USDA Soil Conservation Service, 1988. Snow Survey and Water Supply Products Reference. West National Technical Center, Snow Survey Program, 12pp.

Wallis, T.W.R., and Griffiths, J.F., 1995. An assessment of the weather generator (WXGEN) used in the erosion/productivity impact calculator (EPIC). Agric. For. Meteorol., 73:115-133.

White, J.D., and Running, S.W., 1994. Testing scale dependent assumptions in regional ecosystem simulations. J. Veg. Sci., 5:687-702.

Willmott, C.J., and Robeson, S.M., 1995. Climatologically aided interpolation (CAI) of terrestrial air temperature. Int. J. Climatol., 15:221-229.

Chapter 3

Description of a numerical simulation model for predicting the dynamics of energy, water, carbon, and nitrogen in a terrestrial ecosystem

1. Introduction

I will refer to the model described here as 1D-BGC (1 Dimensional BioGeoChemistry). I have three purposes in this description of 1D-BGC. My first purpose is to present the theoretical background in biophysics, physiology, and system-level ecology from which an outline of the processes controlling matter and energy dynamics in a generalized terrestrial ecosystem can be drawn. My second purpose is to distill from this outline a logic by which the most critical processes and states can be predicted, given the need for input parameters which have been, or which can be, observed for a large range of ecosystem types. My third and final purpose is to present the operational detail of an implementation of this logic, with attention to the points where I see the most potential for future improvements in the faithful representation of the essential processes and states.

The definition of "essential processes and states" represented by a model depends largely on its intended application. The design criteria for the model presented here were not set entirely by the objectives of my dissertation research, but were defined more broadly. Much of the fundamental logic for 1D-BGC springs from the FOREST-BGC family of models (Running 1984, Running and Coughlan 1988, Running and Gower 1991, Hunt and Running 1992). Some important changes in model structure and function have been implemented since the last complete model description (Running and Coughlan 1988, Running and Gower 1991) was published, and a new reference point was required for the ongoing research programs using the BGC model family. Since I had created a new "standard" model version during my participation in the Interior Columbia River Basin Ecosystem Assessment, and since that version had been found to be generally useful, I took on the task of formalizing, updating, and documenting a new version o

the model. The result is 1D-BGC, and this chapter serves as a complete description of the new model version.

2. Overview of model structure and components

The most basic component of the model is its temporal framework, which consists of a dual, discrete timestep approach similar to the implementation in previous versions of the BGC logic. Most simulated ecosystem activity occurs at a daily timestep, driven by daily values for temperature, precipitation, radiation, and humidity. Examples of processes assessed daily are soil water balance, photosynthesis, new leaf and fine root growth and litterfall, and carbon and nitrogen dynamics in the litter and soil. A few processes, including the determination of phenological timing and the allocation of plant stores of carbon and nitrogen to growth of new tissue, are treated on an annual timestep. The model structure assumes that a simulation begins on January 1 and ends on December 31 any number of years later, with all years having 365 days (no leap days). Since it is intended for general application over global domains, there is treatment of both northern and southern hemisphere simulations, without having to start them on different days.

The next most fundamental characteristic of the model is its spatial framework. The model's name, One Dimensional BioGeoChemistry, implies a simple spatial structure. The single dimension in question is defined by the vertical extent of a vegetation canopy and its rooting system. None of the simulated processes are influenced by lateral or horizontal dimensions, and so influences such as subsurface water drainage from upslope sources or the effect of adjacent land surfaces on the temperature and humidity of the near-surface air are ignored. For convenience in calculations, and to keep model parameters and outputs dimensionally consistent with observations, an arbitrary unit of horizontally projected area is defined, over which all fluxes and storages are quantified. Complete horizontal homogeneity is assumed within that unit area, which by default is one square meter. This horizontal square meter and the vertical extent of the canopy and its rooting system define the maximum physical boundaries of the simulation system, which I will refer to often as simply "the system". It includes all the living and dead plant material, the air within the leaf or shoot boundary layer, the air within the boundary layer of the

litter/soil surface, all mineral and organic matter making up the soil and litter down to a depth at which penetration by the root system is negligible, all the water held in the soil down to that same depth, and any water held on top of the soil as snow.

The structural and functional attributes of 1D-BGC are founded on the principles of conservation of energy and conservation of mass. These principles are applied to four fundamental quantities, or state variables; shortwave radiant energy, water, carbon, and nitrogen. The principle of conservation means that, for any one of these state variables, the sum of all inputs to the system, less the sum of all outputs, is equal to the net storage of that component within the system. The basic structural components of 1D-BGC are compartments, or pools, capable of accepting, releasing, and storing the state quantities. The basic functional components of the model are fluxes linking these pools to each other and to the external environment, and which are driven and controlled simultaneously by a range of external and internal factors.

System behavior over time is regulated by various control mechanisms generally consisting of feedback loops by which knowledge of a previous state or flux influences a current flux. In the development of 1D-BGC, I have tried to limit the number of feedback control mechanisms. I view them as the embodiment, in algorithm form, of the evolutionary history of the ecosystem, the anti-entropic responses of the plants and microorganisms to their environments. They are difficult aspects of the ecosystem to quantify (Hunt *et al.* 1985, Field *et al.* 1992, Mäkelä *et al.* 1996, Beringer *et al.* 1996), and are difficult aspects of an ecosystem model to defend. My goal has been to represent the physical and biochemical processes in as much detail as can be quantified for a wide range of global vegetation types, and to include only enough feedback controls, judiciously placed, to provide the plants with buffers against the more extreme intra- and inter-annual variations in their physical environments. To this purpose, I have excluded a number of widely recognized genotypic and phenotypic feedback responses, such as the influence of water and nitrogen availability on leaf to fine root allocation ratios, and the influence of soil nutrition on foliar nutrient content. The rationale for such exclusions is my belief that an understanding of system dynamics can be obscured by the excessive implementation of feedback control. I would rather

present a model that is too simple, but tractable, and improve it by adding controls as deemed essential, than present a model that is too complex, and through intractability be forced to excise inessential controls.

In the following paragraphs I outline the principal flows and storage compartments for each of the state variables: shortwave radiant energy, water, carbon, and nitrogen. These outlines are faithful condensations of the details presented later, and their purpose is to introduce the reader to the basic model structure, to highlight the points of input and output, and to demonstrate that numerous linkages exist between components.

Shortwave radiant energy input is prescribed. Upon entering the system it is divided into a reflected proportion and an absorbed proportion. The reflected proportion is assumed to leave the system without further interactions (e.g. there is no explicit within-canopy scattering of shortwave radiation). Absorbed radiation is further partitioned between a fraction absorbed by the canopy and a fraction absorbed beneath the canopy at the litter/soil surface. Of the fraction absorbed in the canopy, a further division is made between a fraction absorbed as direct radiation on sunlit leaves and a fraction absorbed as diffuse radiation on shaded leaves. Each of the absorbed fractions of the incident shortwave radiation is made available as energy which is either used for the evaporation of water at the absorbing surface or stored at the surface in the form of kinetic energy which serves to increase the surface temperature. Absorbed photosynthetically active radiation (APAR) is calculated as a function of the canopy absorption of shortwave radiation and used in the photosynthetic sub-model described below.

Water input to the system occurs only as precipitation, which is prescribed. Precipitation can occur as snow or as rain. As snow, it enters a snowpack compartment which can later be melted. Rainfall may be intercepted on the canopy, where it either evaporates or falls to the surface as leaf-drip. Rainfall not intercepted on the canopy, as well as leaf-drip and melted snow water, enter the soil water compartment. Depending on a prescribed soil texture, water entering the soil can be immediately or slowly released as runoff, or it can remain in the soil until it is either evaporated from the litter/soil surface, or transpired through the canopy. Evaporation and transpiration are driven by absorbed shortwave radiation, atmospheric vapor pressure, and surface temperature and humidity, and are mediated by conductances to the transport of sensible heat and water vapor. Water leaving the system through transpiration is strongly

controlled by the conductance of stomates to water vapor diffusion, which is a biological feedback response depending on APAR, leaf temperature, atmospheric water vapor pressure, soil water content, and the night minimum temperature. The surface conductance of the litter/soil surface to evaporated water vapor is a function of the time since a rain or snowmelt event. All evaporation and runoff flows are considered as outputs from the system.

Carbon enters the system through the process of photosynthetic fixation of atmospheric carbon dioxide (CO_2). This process is driven by APAR, the canopy air concentration of CO_2, and the concentration and activity of photosynthetic enzymes in the leaves, and is mediated by leaf temperature and the stomatal conductance to CO_2. Carbon fixed by photosynthesis can either be deposited as new growth in any of the plant tissues (leaf, live stem, dead stem, live coarse root, dead coarse root, or fine root), or respired and returned to the atmosphere. Growth of new tissue requires a one-time expenditure of fixed carbon as growth respiration, and all live tissue requires an expenditure of fixed carbon as maintenance respiration during every daily timestep. Maintenance respiration is sensitive to tissue temperature and tissue nitrogen content. Carbon fixed as growth of new leaves or new fine roots can leave the live plant compartments and enter the litter either through litterfall or through whole-plant mortality. Carbon fixed as live stem or live coarse root growth can turn over to dead stem and dead coarse root, and both live and dead stem and coarse root components can enter a coarse woody debris compartment as the result of whole-plant mortality. Through physical disintegration, coarse woody debris enters the litter pools. Litter is acted upon by decomposing microorganisms, resulting in both respiration and conversion of litter carbon to soil organic matter carbon. Soil organic matter carbon is itself acted upon by microorganisms, resulting in respiration and further organic matter conversions. Litter and soil organic matter decomposition rates are related to gross chemical composition (e.g., cellulose, lignin, and humic material have distinct decomposition rates), to substrate carbon to nitrogen ratios (C:N), to soil mineral nitrogen availability, and to soil temperature and water content. Respiration (autotrophic and heterotrophic) is the only pathway for carbon leaving the system.

Nitrogen enters the system through the deposition of mineral nitrogen from the atmosphere to the soil, where it enters the soil mineral nitrogen pool. Soil mineral nitrogen can be incorporated into soil

organic matter through the action of soil microorganisms that are feeding on litter material (immobilization), or it can be taken up by the vegetation for incorporation into new tissue growth, or it can be leached from the soil under conditions of soil water saturation. Nitrogen immobilized in soil organic matter can later be released in mineral form (mineralized) as the carbon component of the organic matter is respired. Mineralized nitrogen enters the soil mineral nitrogen pool. All the same factors controlling the decomposition of litter and soil organic matter carbon also influence the immobilization and mineralization of litter and soil nitrogen. Nitrogen taken up by plants can either be incorporated in new growth, or, in the case of excess nitrogen with respect to some other growth-limiting resource (carbon, water, or light), it can be returned to the soil mineral nitrogen pool. Nitrogen fixed in new growth can be returned to the litter through litterfall of leaves and fine roots, or it can leave the plant as a component of coarse woody debris, after which it undergoes physical disintegration before entering the litter. Nitrogen leached from the soil mineral pool leaves the system, and this is the only pathway for system nitrogen loss.

3. Detailed description of shortwave radiation budget

Shortwave radiation is defined here as all radiation in the spectral region between 300 and 3000 nm. Over 98% of the solar energy entering Earth's atmosphere occurs in this wavelength region, about 45% of that occurring as visible wavelengths (~ 400-700 nm), 53% as near- and mid-infrared wavelengths (700-3000 nm), and the remaining small fraction as ultraviolet radiation (~ 300-400 nm) (Jones 1992, Nobel 1991).

3.1 Shortwave radiation inputs

The model requires an input time series of daily shortwave radiation, quantified as the radiant flux density in Wm^{-2}, measured normal to the simulation surface, averaged over the daylight portion of the day. Daylight is defined as the period when the sun would be above a flat horizon, ignoring the influence on daylength of topographic shading of direct radiation.[1] In addition to this daylight average shortwave

[1] Note that in practice, when a 1D-BGC simulation is driven by output from Daymet or its one-dimensional predecessor MT-CLIM (chapter 1), daylength is calculated as described above, but total radiation in each discrete sub-daily time interval is calculated by summing diffuse and direct radiation, with the local slope

radiant flux density (R_S, W m^{-2}), the model also requires the daylight average radiant flux density occurring as photosynthetically active radiation (PAR), in the wavelength region from 400-700 nm (R_{PAR}, W m^{-2}), for use in driving the photosynthesis sub-model. This quantity can either be supplied from an external source, or the following simple assumption relating R_{PAR} to R_s can be employed within the model when only R_s is available as an external data source (Nobel 1991):

$$R_{PAR} = 0.45\ R_s$$

An important simplifying assumption implicit in the model is that the radiation inputs (both R_S and R_{PAR}) for each day consist of a direct component (sunlight) and a diffuse component (skylight), with the ratio between them remaining constant through the course of a day. This assumption is obviously unrealistic, since it ignores the fact that differences in cloudiness alter the proportions of daily total radiation received as direct and diffuse components. However, since the assessment of sub-daily, or even daily, proportions of direct and diffuse radiation with better accuracy than provided by the gross assumptions made here would be very difficult over large areas without intensive measurement effort (Gates 1980). The inclusion of an explicit treatment of variable direct and diffuse components of incident radiation complicates the estimation of radiation absorption by the vegetation canopy (e.g., de Pury and Farquhar 1997). I have therefore chosen to ignore the complicating details and accept the ensuing errors.

3.2 Canopy radiation absorption and fractionation between sunlit and shaded portions

A simplification made in many ecosystem process models is to assume a single average irradiance for the entire vegetation canopy as a function of the canopy projected leaf area index (L_P, the projected leaf area per unit of ground area, unitless) and a unitless coefficient describing the progressive absorption of shortwave radiation with increasing L_P (the extinction coefficient, k_S) (Running and Coughlan 1988, Running and Gower 1991, Hunt and Running 1992, Potter *et al.* 1993, Sellers *et al.* 1992, Keane *et al.* 1996a). The responses of leaf-scale stomatal conductance and photosynthesis are non-linear with respect to absorbed radiation, and so the big-leaf assumption can result in serious errors in both water and carbon budgets for the whole canopy, as has been demonstrated in theory and through comparison with

---

and aspect influencing the duration through the day of direct radiation inputs. Shading from adjacent terrain elements is still ignored.

observations by Sinclair *et al.* (1976), Norman (1980), and de Pury and Farquhar (1997). Division of the canopy into multiple layers having different radiation loads helps to correct for these errors. de Pury and Farquhar (1997) show that a division into only two layers, corresponding to sunlit and shaded canopy fractions, is a significant improvement over the single-layer approach, and is nearly as accurate as an approach using more than two layers. I elected to implement a two-layer approach in 1D-BGC, with estimates of radiation interception, canopy evaporation, stomatal conductance, transpiration, and photosynthesis treated separately for the sunlit and shaded canopy fractions. My implementation is not as refined as that described by de Pury and Farquhar (1997), having a simpler treatment of direct and diffuse radiation components. The new algorithms use the same model input parameters required by the big-leaf treatment in earlier BGC versions. Details of the canopy radiation absorption and transmission algorithm are presented below.

Total shortwave flux density absorbed at the surface ($R_{S,T}$) is calculated as a function of incident radiation and reflectance from the surface (surface shortwave albedo, α_S, unitless):

$$R_{S,T} = R_S (1.0 - \alpha_S)$$

where $R_{S,T}$ includes both canopy and litter/soil absorption. Of this total, the amount absorbed in the canopy as a whole (the total of sunlit and shaded layers) is estimated by assuming a Beer's law form for canopy radiation interception (Campbell 1977):

$$R_{S,C} = R_{S,T} (1.0 - \exp(-k_S L_P))$$

where $R_{S,C}$ is the total canopy absorbed shortwave flux density (W m^{-2}). The extinction coefficient k_S is prescribed, and is defined as the total leaf area projected on a horizontal surface (when the leaves are in their canopy orientation) divided by the total projected leaf area. The shortwave radiation transmitted through the canopy and absorbed at the ground surface ($R_{S,G}$) is calculated as:

$$R_{S,G} = R_{S,T} - R_{S,C}$$

Leaves are relatively transparent to near-infrared wavelengths, compared to visible wavelengths, and so in general the canopy extinction coefficient for PAR is higher than for total shortwave radiation, while the percent surface reflectance is lower (Jones 1992). Based on data in table 2.4 from Jones (1992), I define the surface albedo in the PAR region (α_{PAR}) and the canopy extinction coefficient for PAR (k_{PAR}) as:

$$\alpha_{PAR} = \alpha_S / 3.0$$

$$k_{PAR} = 1.2\ k_S$$

and the total par absorbed by the canopy surface ($R_{PAR,C}$) as:

$$R_{PAR,C} = R_{PAR}\ (1.0 - \alpha_{PAR})\ (1.0 - \exp(-k_{PAR}\ L_P))$$

Separate calculations for canopy absorption of shortwave radiation and of PAR are employed because leaf temperature and therefore canopy evaporation and transpiration depend on total shortwave radiation, while stomatal conductance and photosynthesis depend more directly on PAR.

The theoretical treatment of canopy fractions in sunlit and shaded conditions by Campbell (1977), Jones (1992), and de Pury and Farquhar (1997), depends on sub-daily resolution of instantaneous solar elevation angles and incident fluxes. A mechanistic description of the canopy light interception mechanism is developed most naturally for this instantaneous flux case, but it is computationally expensive, and sensitive to details of canopy leaf orientation which are frequently poorly defined over large regions. In contrast, 1D-BGC makes the simplifying assumption that a single daily estimate of canopy radiation processes can be performed using daylight average radiant flux density and extrapolated meaningfully to a daily integral through the calculated daylength. This assumption remains largely untested. Some support for the assumption comes from a recent comparison of sensible and latent heat fluxes and soil water dynamics simulated by various land-surface parameterization schemes (including Biome-BGC). The schemes showed similar results for annual totals of evapotranspiration even though some of them operated on sub-daily timesteps, some on daily timepsteps, and some on monthly timesteps. Predictions of seasonal total evaporation from the daily and monthly models came nearer to observed values than did predictions from the models with sub-daily timesteps (Shao and Henderson-Sellers 1996, Parton *et al.* 1996).

The theoretical formula for the sunlit leaf area index (L_{sun}) in a canopy with horizontally oriented leaves receiving only direct sunlight is independent of solar elevation angle, and is defined as (Jones 1992):

Eq. 3.1 $L_{sun} = 1.0 - \exp(-L_P)$

while the same quantity for a canopy with leaves oriented at an angle is dependent on an extinction coefficient (k) that varies with leaf angle and solar elevation:

Eq. 3.2 $L_{sun} = (1.0 - \exp(-k\ L_P))/k$

These equations imply that, for a horizontal leaf angle, the maximum sunlit leaf area index is 1.0, while the maximum for an angled leaf canopy is 1/k, which is typically less than 1.0 for the sun lower than 40° above the horizon and greater than 1.0 for the sun higher than 40° (Jones 1992). Eq. 3.1 then resembles an average of Eq. 3.2 over a typical low-to-mid latitude range of diurnal solar angles. Since the 1D-BGC algorithm is intended to approximate the average daylight condition, I use Eq. 3.1 to predict the leaf area receiving direct sunlight. Because the incident radiation is assumed to be calculated as a value normal to the ground surface, and because the empirical, prescribed value of k_S is intended to account for the cumulative influence of leaf angle, leaf clumping, and solar angle variations on radiation interception (Campbell 1986, Sampson and Smith 1993, Yang *et al.* 1993, Fassnacht *et al.* 1994, Hirose and Werger 1995), I use the following formula to calculate the shortwave radiation intercepted per unit projected leaf area in the sunlit fraction of the canopy ($R'_{S,sun}$) (Jones 1992):

$$R'_{S,sun} = k_S\ R_S$$

while the total radiation absorbed in the sunlit canopy fraction ($R_{S,sun}$) is:

$$R_{S,sun} = k_S\ R_S\ L_{sun}$$

The remainder of the total canopy absorbed shortwave radiation is assumed to be the total absorbed by the shaded canopy fraction ($R_{S,shade}$):

$$R_{S,shade} = R_{S,C} - R_{S,sun}$$

The shortwave radiation absorbed per unit of projected leaf area in the shaded canopy fraction ($R'_{S,shade}$) is then:

$$R'_{S,shade} = R_{S,shade} / L_{shade}$$

where L_{shade} is the shaded canopy leaf area index, defined by:

$$L_{shade} = L_P - L_{sun}$$

Because L_{sun} does not depend on any factors that differ in treatment between shortwave radiation and PAR, the same values for L_{sun} and L_{shade} are used in the equations for the fractionation of canopy absorbed PAR:

$$R_{PAR,sun} = k_{PAR}\ R_{PAR}$$

$$R'_{PAR,sun} = R_{PAR,sun} / L_{sun}$$

$$R_{PAR,shade} = R_{PAR,C} - R_{PAR,sun}$$

$$R'_{PAR,shade} = R_{PAR,shade} / L_{shade}$$

where $R_{PAR,sun}$ is the total PAR absorbed by the sunlit canopy fraction, $R'_{PAR,sun}$ is the PAR absorbed per unit sunlit projected leaf area, $R_{PAR,shade}$ is the total PAR absorbed by the shaded canopy fraction, and $R'_{PAR,shade}$ is the PAR absorbed per unit shaded projected leaf area. The values for radiation absorbed in the canopy and by the litter/soil surface are used to estimate evaporation, transpiration, and photosynthesis.

4. Detailed description of the water budget

Water state variables are denoted by a subscripted W, and are all defined with units of kg (H_2O) m^{-2} (ground surface). Water can be stored in the snowpack, in the soil water, and as canopy intercepted water. Of these, the snowpack and soil water compartments can accumulate and lose water mass over the entire duration of a simulation, while the canopy intercepted water compartment is forced to zero mass by the end of each daily timestep, with no day-to-day storage of intercepted water on the canopy. Water fluxes, denoted by a subscripted Q, are calculated either as kg m^{-2} d^{-1} or as kg m^{-2} s^{-1}, but all fluxes are converted for final reporting and integration with other sub-model components to units of kg m^{-2} d^{-1}.

Water input is entirely from precipitation. Output is either as evaporation (from litter/soil surface, from canopy intercepted water, or from transpired soil water), or hydrologic outflow. Evaporation is assumed to enter the atmosphere without changing the prescribed surface humidity. Hydrologic outflow is assumed to drain to soil layers below the level at which the vegetation rooting system becomes negligible, or to enter a stream channel either through overland or subsurface flow.

4.1 Soil water potential and volumetric water content

The balance of water inputs to and outputs from the soil water compartment determines the soil water content. All inputs and outputs are accounted for at the end of each daily timestep, and the volumetric soil water content (θ, volume of water per volume of soil in rooting zone) relevant to calculations during a given day is the value resulting from the balance between the previous day's inputs and outputs. The soil volume in the rooting zone is calculated for a reference area of 1 m^2 by multiplying

the prescribed depth of the rooting zone (d_{root}) by 1 m^2. In prescribing d_{root}, it is assumed that any fraction of the soil volume taken up by rocks is corrected for by reducing d_{root}. In the estimation of stomatal conductance to water vapor diffusion, described below, an estimate is required of the soil water potential (ψ, MPa). An empirical estimate of ψ can be obtained from the relationship:

Eq. 4.1 $$\psi = \psi_{sat}\left(\frac{\theta}{\theta_{sat}}\right)^{b}$$

where ψ_{sat} is the soil water potential at saturation, θ_{sat} is the volumetric water content at saturation, and b is an empirical shape parameter (Clapp and Hornberger 1978, Cosby *et al.* 1984, Saxton *et al.* 1986). An empirical study by Cosby *et al.* (1984) related the parameters ψ_{sat}, θ_{sat}, and b to soil texture for 1448 soils of differing textures and source regions. They defined texture as the percentage, by volume, of sand (P_{sand}), silt (P_{silt}), and clay (P_{clay}) in the soil, after correcting for rock fraction. Particle classes were defined by size according to the USDA system (Saxton *et al.* 1986), where sand = 2.0 to 0.05 mm, silt = 0.05 to 0.002 mm, and clay < 0.002 mm. Multivariate regressions from the Cosby *et al.* (1984) are used in 1D-BGC to estimate ψ_{sat}, θ_{sat}, and b, using prescribed values for P_{sand}, P_{silt}, and P_{clay}:

$$\psi_{sat} = -9.8e\text{-}5 \exp[(1.54 - 0.0095\, P_{sand} + 0.0063\, P_{silt}) \log(10.0)]$$

$$\theta_{sat} = (50.5 - 0.142\, P_{sand} - 0.037\, P_{clay}) / 100.0$$

$$b = -3.10 - 0.157\, P_{clay} + 0.003\, P_{sand}$$

The treatment of hydrologic outflow requires an assessment of the volumetric water content at field capacity (θ_{fc}), where field capacity is defined as a water potential of -0.015 Mpa (W. Parton, personal communication). θ_{fc} is calculated by inversion of eq. 4.1, with ψ_{fc} = -0.015 MPa, to give:

$$\theta_{fc} = \theta_{sat}\left(\frac{-0015}{\psi_{sat}}\right)^{\frac{1}{b}}$$

Also required in the calculation of soil water balance are the field capacity and saturation values for θ converted to soil water contents (W_{fc} and W_{sat}, kg H_2O m^{-2} ground area), and these are calculated as:

$$W_{fc} = 1000.0\, d_{soil}\, \theta_{fc}$$

$$W_{sat} = 1000.0\, d_{soil}\, \theta_{sat}$$

where 1000.0 is the density of water, assumed to be constant (kg m^{-3}).

4.2 Precipitation input and canopy interception

Water input as precipitation is prescribed as a daily total. When the daily average air temperature (T_{avg}, calculated as the average of prescribed daily maximum and minimum temperatures) is above freezing, all precipitation is assumed to enter as rain (W_{rain}), and when T_{avg} is at or below freezing precipitation is assumed to enter as snow (W_{snow}) . Snow enters the snowpack compartment ($W_{snowpack}$), ignoring the possibility of interception of snow on the canopy.

On days with rain, the daily total of rain intercepted on the vegetation canopy (W_{int}) and throughfall of rain entering the soil water compartment ($Q_{rain,soil}$) are estimated as:

Eq. 4.2 W_{int} = the lesser of $\{k_{rain}\, W_{rain}\, L_A\}$ or $\{W_{rain}\}$

$$Q_{rain,soil} = W_{rain} - W_{int}$$

where k_{rain} is the canopy rainfall interception coefficient, with units of kg water intercepted per kg water in gross rainfall per unit all-sided leaf area index per day, and L_A is the all-sided leaf area index. L_A is used here instead of L_P because the majority of literature studies of use in parameterizing k_{rain} report L_A, and conversions to projected LAI are not always available from the source literature. A more mechanistic approach to interception estimation is possible when the sub-daily timing, intensity and duration of rainfall are known, in which case interception losses can be estimated from the simultaneous accumulation and evaporation of canopy intercepted water, assuming knowledge of a canopy storage parameter (Gash 1979, Dolman 1987, Lankreijer *et al.* 1993, Gash *et al.* 1995). Since the design of 1D-BGC is predicated on the assumption that only daily total precipitation is available, a more approximate, empirical approach is required.

A comparison of several explicit models for canopy interception and evaporation, using observations from both needleleaf and broadleaf canopies in temperate climates, found that total canopy water storage capacity (related to leaf type and leaf area) and evaporation environment during precipitation events (related to general climatology) dominate the interception process (Lankreijer *et al.* 1993). This result suggests that an empirical model such as Eq. 4.2 could be parameterized on the basis of studies covering the major vegetation types and climatic regions of the world (e.g., temperate needleleaf, tropical

broadleaf, etc.). The empirical parameter k_{rain} can be determined from observations of gross rainfall (P), interception loss (I), and canopy leaf area. Several studies have reported these data for a variety of vegetation types and climates, and estimates of k_{rain} derived as:

$$k_{rain} = I / (P\ L_A)$$

fall in a relatively narrow range from 0.035 to 0.063 (Table 1).

Table 1. Estimates of k_{rain} derived from other studies

I/P	L_A	k_{rain}	climate	vegetation	reference
0.19	5.3	0.036	temperate	pine	Kelliher *et al.* 1992
0.12	2.3	0.052	temperate	pine	Gash *et al.* 1995
0.19	5.4	0.035	temperate	broadleaf	Klaassen *et al.* 1996
0.20	4.9	0.040	temperate	broadleaf	Lankreijer *et al.* 1993
0.40	6.4	0.063	tropical	broadleaf	Scatena 1990

The mean k_{rain} for temperate forests is 0.041. The single value in Table 1 for tropical forests may be higher than is typical for most tropical forests, due to relatively high VPD and low rainfall intensity at the study site (Scatena, 1990). Many studies reporting P and I do not report the canopy LAI (in either all-sided or projected units). With some additional information for these sites the number of useful studies could be doubled.

Eq. 4.2 may be an unwarranted oversimplification of the interception process, since it appears that the relationship between I and P for a constant LAI is linear, but that it has a positive intercept, such that for precipitation events smaller than some critical value, all precipitation is intercepted (Giacomin and Trucchi 1992). It is not clear from the available studies whether and how this intercept varies with changing LAI. The result of ignoring this intercept in Eq. 4.2 is that interception losses for low precipitation events will be underestimated and losses for larger events will be overestimated. Since the parameterization of k_{rain} depends on long-term average data, the resulting long-term average predictions should still be accurate. Because daily variation in rainfall intensity, duration, and atmospheric evaporative demand during precipitation results in significant variation of interception losses around the average value obtained for k_{rain} (Gash 1979, Gash *et al.* 1995), corrections to Eq. 4.2 for positive intercept may be insignificant in comparison with other confounding factors.

4.3 Evaporation of canopy intercepted water

Of the canopy daily intercepted rainfall some may be evaporated, with any remainder sent to the soil water. The Penman-Monteith equation is used to calculate the evaporation of canopy intercepted water as a function of the air temperature, air pressure, vapor pressure deficit (VPD, Pa), radiant flux density, and resistances to the transport of sensible heat and water vapor. Currently, all canopy intercepted water evaporation is assumed to take place during the daylight portion of the day. On a given day, there can be both evaporation of canopy intercepted water and transpiration of soil water through the canopy. Because these two processes, evaporation from and transpiration through the canopy, both depend on the radiation absorbed in the canopy during the day, and because transpiration will be limited when the evaporation of intercepted water is occurring (due to a low VPD between the intercellular air space and a moist leaf boundary layer), I assume that the daylight period is split between a period when only canopy evaporation of intercepted water is occurring, and another period when only transpiration is occurring. In order to calculate this split, an instantaneous rate of canopy evaporation (E_{int}, kg m^{-2} s^{-1}) is first calculated from the Penman-Monteith equation (hereafter referred to as PM(T,R,r_h,r_v), where T will be replaced with the relevant air temperature (°C), R will be replaced with the relevant radiant flux density (W m^{-2}), and r_h and r_v will be replaced with the relevant resistances to the transport of heat and water vapor (s m^{-1}), respectively) as:

$$E_{int} = PM(T_{day}, R_{S,C}, r_{Ch}, r_{Cv})$$

where T_{day} is the daylight average air temperature, and r_{Ch} and r_{Cv} are, respectively, the canopy total resistances to sensible heat and water vapor transport. For the case of evaporation from the leaf surface, r_{Ch} and r_{Cv} are equivalent, and are equal to the inverse of the leaf boundary layer conductance (g_{bl}, m s^{-1}, see discussion in section 4.6, below) scaled by the projected leaf area index for evaporation from canopy elements in parallel:

$$r_{Ch} = r_{Cv} = 1.0 / (g_{bl} L_P)$$

The time required to evaporate all the intercepted canopy water (t_E, s) is then calculated as:

$$t_E = W_{int} / E_{int}$$

If t_E is shorter than the daylight part of the day (t_{light}, s), then the daily amount of canopy water evaporated ($Q_{int,E}$) is:

$$Q_{int,E} = W_{int}$$

and the remaining daylight time is allocated for transpiration (t_T, s):

$$t_T = t_{light} - t_E$$

Otherwise, the daylight period is not long enough to evaporate all the canopy water, and the actual amount evaporated is defined as:

$$Q_{int,E} = E_{int}\, t_{light}$$

and t_T is set to 0, since the entire daylight period is used evaporating canopy water. In this case, the difference between W_{int} and $Q_{int,E}$ is sent to the soil water compartment:

$$Q_{int,soil} = W_{int} - Q_{int,E}$$

4.4 Losses from the snowpack

Water can leave the snowpack either through melting, by which it enters the soil water compartment, or through sublimation, by which it is lost to the atmosphere. Melting is only possible when T_{avg} is above freezing, and consists of two components, one driven by the air temperature and another driven by the radiation absorbed by the snowpack. The daily amount of snowmelt due to temperature ($Q_{snow,T}$) is:

$$Q_{snow,T} = 0.65\, T_{avg}$$

where 0.65 is the temperature snowmelt coefficient, with units of kg H_2O m^{-2} d^{-1} $°C^{-1}$ (Running and Coughlan 1988). The daily amount of radiation-driven snowmelt ($Q_{snow,R}$) for above-freezing conditions is:

$$Q_{snow,R} = 0.33\, R_{S,G}\, t_{light} / H_{fus}$$

where H_{fus} is the latent heat of fusion of water (J kg^{-1}) and 0.33 is the assumed absorptivity of snow in the shortwave region (Marks *et al.* 1992). No sublimation (or evaporation) occurs on days with T_{avg} above freezing. With T_{avg} at or below freezing, only sublimation occurs, and is estimated as:

$$E_{snow} = 0.6\, R_{S,G}\, t_{light} / H_{sub}$$

where H_{sub} is the latent heat of sublimation of water (J kg^{-1}). Total snowmelt water entering the soil water compartment ($Q_{snow,soil}$) is then:

$$Q_{snow,soil} = Q_{snow,T} + Q_{snow,R}$$

with the following condition testing for daily snowmelt greater than total snowpack:

if ($Q_{snow,soil} > W_{snow}$) : $Q_{snow,soil} = W_{snow}$, and $W_{snow} = 0.0$

4.5 Evaporation from the litter/soil surface

When there is no snowpack, water can be evaporated from the litter/soil surface. The resistance to water vapor transport across this surface is estimated as a function of the number of days since the last rainfall or snowmelt event, in an effort to simulate the retarding effect of soil surface drying on the diffusion of water vapor from lower soil layers to the surface. Based on a number of studies in which the apparent soil surface resistance was measured as the soil surface was progressively dried, I set the soil surface resistance to water vapor transport (r_{Sv}, s m^{-1}) over the first 8 days after the most recent rainfall or snowmelt event as:

r_{Sv} = [500, 1000, 2000, 4000, 8000, 16000, 100000, 1000000]

and r_{Sv} = 1000000 s m^{-1} for more than 8 days since the most recent rainfall or snowmelt event (Camillo and Gurney 1986, Mahfouf and Noilhan 1991, Van de Griend and Owe 1994). The Penman-Monteith equation is used to estimate the instantaneous evaporation from the soil surface under daylight average meteorological conditions (E_{soil}, kg m^{-2} s^{-1}) as:

$E_{soil} = PM(T_{soil}, R_{S,G}, r_{Sh}, r_{Sv})$

where T_{soil} is the estimated soil temperature (°C)in the upper 10 cm of the soil, determined using an 11-day running average of the daily average air temperature (T_{avg}), according to the methods in Zheng *et al.* (1993). The total flux leaving the soil water compartment due to evaporation ($Q_{soil,E}$) is:

$Q_{soil,E} = E_{soil}\ t_{light}$

The variable r_{Sh} is the soil surface resistance to sensible heat transport, which, in theory, varies with windspeed near the soil surface, soil surface roughness, and vertical air temperature profile. It has been found in practice to have a relatively small range with an average value close to 50 s m^{-1} (van de Griend and Owe, 1994). I have taken that value as a constant. Soil water loss to evaporation is removed from the soil water compartment. It may, in future model development, be advantageous to introduce a slightly more complicated treatment of soil water, which would account for the top one or two cm of the litter/soil depth as a separate compartment from which soil water evaporation would be drawn. This would allow for a more rigorous treatment of both soil surface resistance and soil surface evaporation (Mahfouf *et*

al. 1996, Shao and Henderson-Sellers 1996, Shao *et al.*, 1994, Mahfouf and Noilhan 1991). Some testing of this approach is warranted, and data from intensive field measurement campaigns are available for comparison of model and observed fluxes under several different climates (Mahfouf *et al.* 1996). Since one additional soil layer amounts to more than a doubling of the complexity of the soil water calculations, improvements would need to be significant to warrant a change from the current model logic.

4.6 Transpiration and controls on stomatal conductance

Fluxes of transpired water vapor are calculated at the level of the leaf, using the Penman-Monteith equation. Separate calculations are made for the average leaf-level conditions in the sunlit and the shaded canopy fractions. These fluxes are scaled to canopy values by multiplying by the relevant leaf area, and the total canopy transpiration flux is obtained as the sum of the canopy-level fluxes for the sunlit and shaded canopy fractions.

The Penman-Monteith equation relies on a description of resistances or conductances to sensible heat and water vapor transport away from a surface, where conductance (g, m s^{-1}) is simply the inverse of resistance ($g = 1/r$). In most of this section I use conductances, since it simplifies the algebra. The source of sensible heat for a leaf is assumed to be the leaf surface, and so the pathway of transport of sensible heat away from the leaf consists simply of passage through the leaf boundary layer, the thin layer of laminar (non-turbulent) flow adjacent to the leaf. The conductance associated with passage of sensible heat through the leaf boundary layer (g_{bl}) depends on the size and shape of the leaf and the windspeed, and is typically large for narrow or needle-shaped leaves (~0.1 m s^{-1}) and smaller for broad leaves (~0.01 m s^{-1}) (Jones 1992, Nobel 1991). The source of water vapor being transported away from a leaf during transpiration is assumed to be the exposed cell-walls inside the stomatal cavity, and the pathway for this transport is more complicated than for sensible heat. Water vapor can leave the stomatal cavity either through the stomates or through the leaf cuticle. Conductance along the stomatal pathway varies depending on stomatal aperture. Conductance along the cuticular pathway is relatively constant and generally very low. Leaf-scale stomatal conductance (g_s) approaches cuticular conductance (g_c, 0.00001 to 0.0004 m s^{-1}, Körner 1995, Jones 1992, Nobel 1991) when stomatal aperture is at its smallest, and increases with stomatal opening (maximum values 0.001 to 0.025 m s^{-1}, with an average for many vegetation types of ~0.005 m s^{-1}

^{1}). After exiting the stomatal cavity, water vapor must also cross the leaf boundary layer, and the conductance for sensible heat and water vapor transport along this pathway are generally equivalent (Jones 1992). The total conductance for leaf-level transport of water vapor during transpiration (g_{Tv}) is therefore the parallel conductances of stomate and cuticle, together in series with the boundary layer conductance, giving:

Eq. 4.3 $$g_{Tv} = \frac{g_{bl}(g_s + g_c)}{g_{bl} + g_s + g_c}$$

as the general case.

Stomatal conductance varies in response to the radiant flux incident on the leaf, the leaf temperature, the difference in water vapor pressure between the atmosphere and the interior of the leaf, the soil water potential in the rooting zone, and the atmospheric concentration of CO_2, among other factors (Kelliher *et al.* 1995, Körner 1995, Jones 1992). This variation in stomatal conductance has a strong influence on the rate of transpiration (Jarvis and McNaughton, 1986). The treatment of environmental influence on stomatal conductance in 1D-BGC is very similar to that described by Running and Coughlan (1988) for FOREST-BGC. The basic logic is to assign a level of stomatal control (from 0.0 for complete closure to 1.0 for full opening) for each of a series of environmental influences, and to arrive at the overall stomatal control as the product of all the individual controls. The following environmental parameters are assumed to have influence the stomatal conductance: radiant flux density, air temperature, soil water potential, daily minimum temperatures below freezing, and vapor pressure deficit. Of these, only radiant flux density is assumed to vary between the sunlit and shaded canopy fractions. This control is calculated separately for each fraction, and the other controls are calculated once and applied to both fractions.

The influence of radiation on stomatal conductance is generally found to follow a hyperbolic relationship, and although there is variation in the hyperbolic curvature found for different species, for different leaves (sun and shade) of the same species, and for the same species growing in different environments, many studies find that rapid increases in stomatal conductance occur between 50 and 500 $\mu mol\ m^{-2}\ s^{-1}$ of photosynthetically active photon flux density (PPFD) for trees (e.g., Holmgren *et al.* 1996, Jones 1992, Petersen *et al.* 1991, Gutschick 1995, Barradas *et al.* 1994, McCaughey and Iacobelli 1993). I

use a simple hyperbola to describe this influence on the stomatal conductance multiplier for intercepted radiation (m_{PPFD}, unitless):

$$m_{PPFD} = PPFD / (PPFD_{50} + PPFD)$$

where PPFD is defined as:

$$PPFD = 4.55\ R'_{PAR}$$

which converts values for PAR radiation intercepted per unit of projected leaf area (either for sunlit or shaded canopy fractions) from W m^{-2} to μmol m^{-2} s^{-1}, with 4.55 being the average energy (μmol J^{-1}) of PAR photons. $PPFD_{50}$ is the hyperbolic shape parameter, which defines the level of PPFD for which m_{PPFD} = 0.5. In a recent review of stomatal conductance studies, Körner (1995) suggests a range of values between 50-100 for $PPFD_{50}$ in late successional woody species, with values perhaps twice as high for grasses. Since Körner's analysis is based on the PPFD incident on a flat surface, and not corrected for leaf angle, the difference between woody and graminoid species in $PPFD_{50}$ relevant to PPFD calculated as interception per unit projected leaf area, as here, should be reduced, as should be the range for woody species, with a shift to the lower end of the $PPFD_{50}$ range. In addition, field studies report the influence of instantaneous PPFD on g_s, but the daylight average PPFD is more relevant for 1D-BGC, which also should tend to reduce the model value of $PPFD_{50}$. I therefore use a constant value of 75 μmol m^{-2} s^{-1} for $PPFD_{50}$ (Figure 1a).

The influence of air temperature on stomatal aperture is generally an increase from freezing to the typical summer maximum temperature experienced by the plant, with a decrease in stomatal aperture at temperatures substantially higher than the average mid-summer maximum temperature (Jones, 1992). As a representation of the influence of temperature on stomatal conductance, I adopt the function specified by Rastetter *et al.* (1991), and used in Hunt *et al.* (1996), which specifies the stomatal conductance multiplier for daylight average air temperature (m_{tday}, unitless) as a function of a maximum (critical) temperature, an optimal temperature, a shape parameter, and a condition that forces the multiplier to 0.0 at 0°C:

$$m_{tday} = \exp(0.1 \ln((T_{crit} - T_{day})/(T_{crit} - T_{opt}))\ (T_{crit} - T_{opt}))\ \exp(0.1\ (T_{day} - T_{opt}))$$

$$\text{and, if } (T_{day} < 5.0°C) : m_{tday} = m_{tday}\ T_{day} / 5.0$$

where T_{crit} is the critical temperature, above which m_{tday} = 0.0, T_{opt} is the optimal temperature, at which m_{tday} = 1.0, T_{day} is the daylight average air temperature, and 0.1 is a unitless shape parameter (Figure 1b).

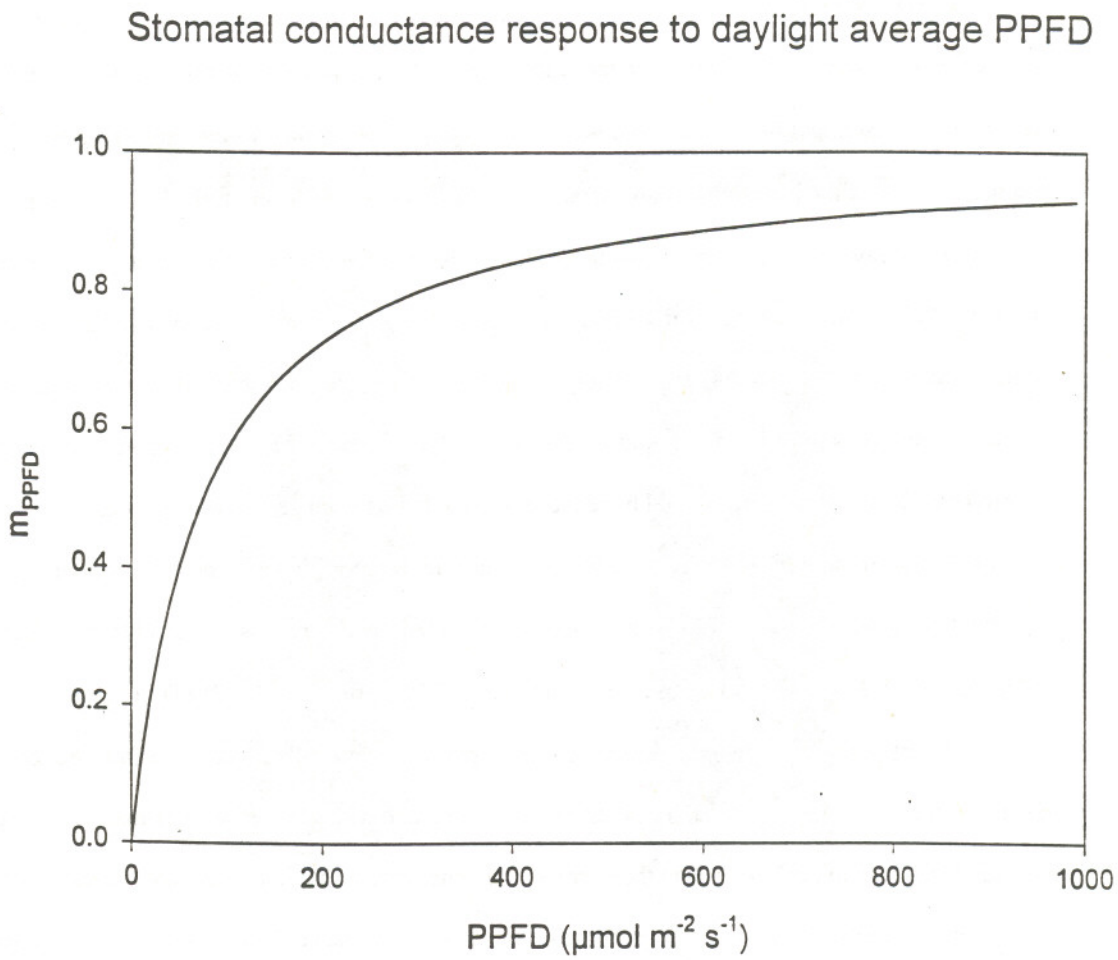

Stomatal conductance response to daylight average PPFD
1.0
0.8
0.6
0.4
0.2
0.0
m_{PPFD}
0
200
400
600
800
1000
PPFD (µmol m^{-2} s^{-1})

Figure 1a

Stomatal conductance response to daylight average temperature

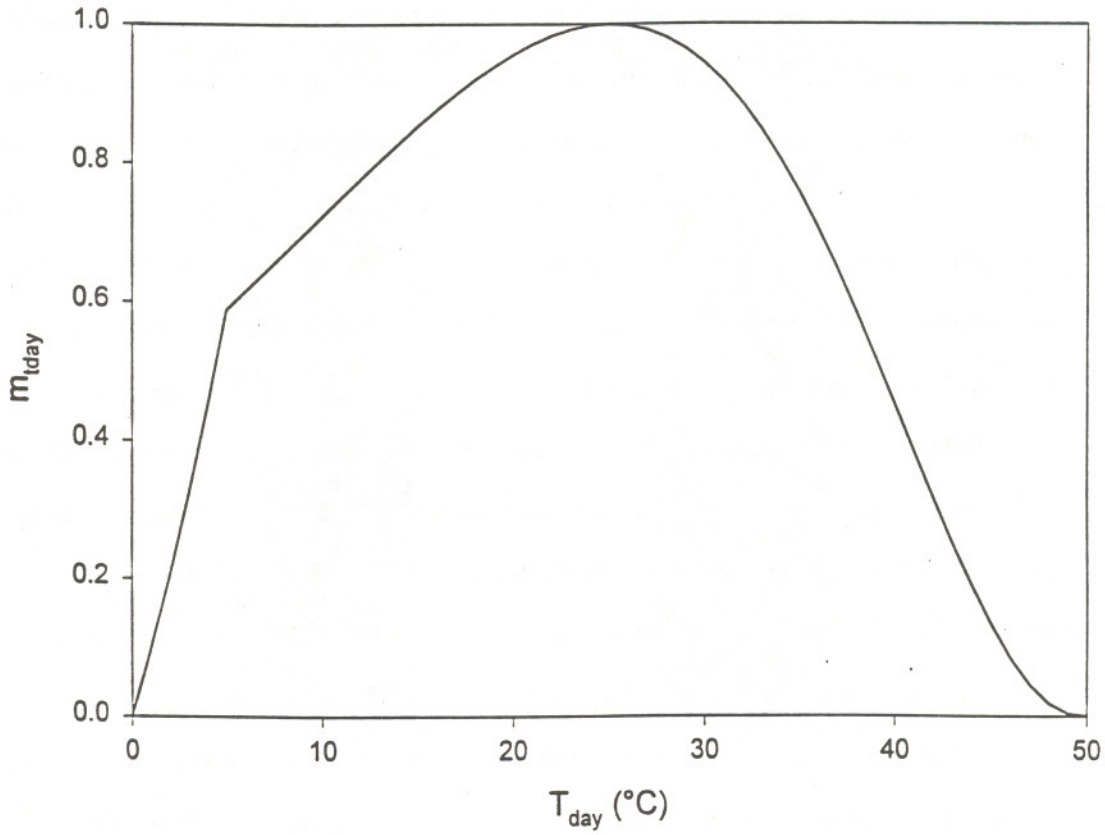


Figure 1b

Daily minimum temperatures below freezing are assumed to limit stomatal conductance, with an arbitrary lower limit of -8°C, below which stomatal conductance is forced to 0.0 (Graham and Running 1984, Smith *et al.* 1984). The multiplier for this effect (m_{tmin}, unitless) is a linear function from 0.0 at -8°C to 1.0 at 0°C.

Stomatal aperture for many species is known to be sensitive to changes in the atmospheric exaporative demand (Schulze and Hall 1982, Monteith 1995). This demand is commonly quantified as the vapor pressure deficit (VPD, Pa), the difference between the water vapor partial pressure in the air surrounding the leaf and in the stomatal cavity air. Air in the stomatal cavity is usually assumed to be saturated with respect to water vapor at the leaf temperature. I assume for the general model development that leaf temperature is the same as air temperature (in Chapter 4 an extension to the 1D-BGC logic is given which treats the leaf temperature more explicitly). The multiplier describing the influence of VPD on stomatal conductance (m_{VPD}, unitless) is defined as a simple linear relationship, varying from 1.0 at a prescribed minimum VPD (below which no stomatal control is exercised) to 0.0 at a prescribed maximum VPD, above which stomates are assumed to be completely closed. Körner (1995) suggests that the minimum VPD at which stomatal closure begins is around 1000 Pa for a variety of species, and that a relatively linear decrease in conductance occurs until VPD reaches 4000-5000 Pa, at which point stomatal conductance is around 1/5 to 1/10 of its value under high humidity.

The final control on stomatal conductance is that due to soil water potential (m_{ψ}, unitless). Stomates are known to respond to the average soil water status, by beginning to close as the soil water potential drops below a certain point and reaching full closure at a critical soil water potential (the wilting point) (Schulze *et al.* 1980, Graham and Running 1984, Tenhunen et al 1987). I define this relationship as a simple linear function, with a prescribed soil water potential at which stomatal closure begins, and another prescribed soil water potential at which full stomatal closure is experienced.

The total response of the stomatal conductance to these environmental factors (m_{total}) is defined as their product, with separate treatment for the sunlit and shaded canopy fractions:

$$m_{total,sun} = m_{PPFD,sun}\, m_{tday}\, m_{tmin}\, m_{VPD}\, m_{\psi}$$

$$m_{total,\,shade} = m_{PPFD,shade}\, m_{tday}\, m_{tmin}\, m_{VPD}\, m_{\psi}$$

The actual daily leaf-level stomatal conductance to transpired water vapor (g_s) is assessed as the maximum leaf-scale stomatal conductance to water vapor ($g_{s,max}$) scaled by the final conductance multiplier:

$$g_{s,sun} = g_{s,max}\, m_{total,sun}$$

$$g_{s,shade} = g_{s,max}\, m_{total,shade}$$

Total leaf-level conductance to transpired water vapor is calculated for sunlit and shaded fractions using these stomatal conductances and Eq. 4.3. The instantaneous transpiration per unit projected leaf area for the sunlit and shaded canopy fractions at daylight average meteorological conditions is then estimated from the Penman-Monteith formula as:

$$T'_{sun} = PM(T_{day}, R_{S,sun}, r_{Th}, r_{Tv,sun})$$

$$T'_{shade} = PM(T_{day}, R_{S,shade}, r_{Th}, r_{Tv,shade})$$

where r_{Th} is the resistance to sensible heat transport relevant to the transpiration flux calculation, which is simply $1/g_{bl}$, and $r_{Th,sun}$ and $r_{Tv,shade}$ are the total leaf-level resistances to transpired water vapor, defined as $1/g_{Tv,sun}$ and $1/g_{Tv,shade}$, respectively. The daily total transpiration flux (from both canopy fractions) is finally calculated as:

$$Q_{soil,T} = t_T\,(T'_{sun}\, L_{sun} + T'_{shade}\, L_{shade})$$

This treatment of leaf-level conductances requires prescribed values for maximum stomatal, cuticular and leaf boundary layer conductance. Since conductance through air varies with the air temperature and pressure, the prescribed values are assumed to be given for standard conditions of 20 °C and 101300 Pa. Based on the prescribed daily air temperature (converted to Kelvins) and an air pressure estimated from a prescribed elevation (p_a, Pa, assuming a standard atmosphere, Irbane and Godson 1981), the prescribed standard conductances ($g'_{s,max}$, g'_c, and g'_{bl}) are converted to actual conductances for the day according to Jones (1992):

$$g_{s,max} = g'_{s,max}\,((T_{day}+273.15)/293.15\ K)^{1.75}\,(101.3\ kPa/p_a)$$

and similarly for the other two prescribed conductances.

4.7 Soil water balance and hydrologic outflow estimation

At the end of each daily timestep, the soil water compartment is balanced and hydrologic outflo is estimated. Inputs to the soil water compartment over the course of the day (Q_{in}) are calculated as:

$$Q_{in} = Q_{rain,soil} + Q_{int,soil} + Q_{snow,soil}$$

Outputs from the soil water compartment (Q_{out}) are calculated as:

$$Q_{out} = Q_{soil,.E} + Q_{soil,T}$$

Inputs and outputs are balanced with the soil water from the end of the previous daily timestep ($W_{soil,prev}$) to estimate the new soil water content as:

$$W_{soil} = W_{soil,prev} + Q_{in} - Q_{out}$$

If the new $W_{soil} > W_{sat}$, then the soil water content in excess of W_{sat} is immediately released as saturation outflow:

$$Q_{soil,satflow} = W_{soil} - W_{sat}$$

$$W_{soil} = W_{soil} - Q_{soil,satflow}$$

If no saturation flow has occurred, and if $W_{soil} > W_{fc}$, then the soil water in excess of field capacity drains to outflow as a fraction of the excess on each timestep when the excess occurs:

$$Q_{soil,drainflow} = 0.5\ (W_{soil} - W_{fc})$$

$$W_{soil} = W_{soil} - Q_{soil,drainflow}$$

where 0.5 indicates that 50% of the soil water in excess of field capacity on each day drains to outflow. The net result of the outflow algorithm is that water in excess of saturation leaves immediately as the equivalent of overland flow, while water in excess of field capacity drains slowly, approaching field capacity as an exponential decay in the absence of other net losses.

5. Detailed description of the carbon and nitrogen budgets

I present the description of these budgets under a single heading because the model dynamics of carbon (C) and nitrogen (N) are tightly coupled, and numerous mutual dependencies make an isolated description and discussion of either budget difficult. State variables (denoted by subscripted C or N) are expressed in units of kg (C or N) per m^2 ground area, while fluxes (denoted by subscripted CF or NF) are expressed in units of kg (C or N) per m^2 ground area per day, although other intermediate units are also

used. The carbon state variables include plant compartments for leaf, live stem, dead stem, live coarse root, dead coarse root, fine root, and growth storage pools for leaf and fine root growth that have been allocated but not yet displayed. Litter and soil carbon state variables include litter soluble carbon, litter cellulose carbon, litter lignin carbon, soil microbial biomass and byproduct carbon, and recalcitrant soil organic matter carbon. A coarse woody debris compartment is used as an intermediate litter storage pool. Nitrogen state variables include compartments corresponding to all of the plant, litter, soil, and coarse woody debris carbon compartments, as well as a soil mineral nitrogen compartment representing ammonium and nitrate in the soil. The model includes two additional state variables serving as storage compartments between allocation events: photosynthetically fixed C (C_{pool}), and mineral N taken up from the soil (N_{pool}).

5.1 Photosynthetic carbon fixation and canopy nitrogen distribution

The sole input of carbon is through photosynthetic fixation of atmospheric CO_2. The photosynthetic process is described through a sequence of algorithms, some of which are motivated mainly by biochemical insight and others of which are driven more by empiricism. The representation of photosynthesis in 1D-BGC differs significantly from that in previously published descriptions for the BGC family of models. The FOREST-BGC representation of photosynthesis (Running and Coughlan 1988) relies on the parameterization of a mesophyll conductance to CO_2, with fixation estimated as a diffusion process driven by an assumed intercellular CO_2 concentration (C_i, Pa). FOREST-BGC does not implement an explicit treatment of the photosynthetic biochemical pathways. In the generalization of FOREST-BGC to other biomes (BIOME-BGC: Hunt and Running 1992) a more detailed model was implemented, with an explicit treatment of photosynthetic biochemistry (Farquhar *et al.* 1980, Leuning 1990). The BIOME-BGC implementation includes a pseudo-iterative calculation of C_i, as well as an explicit calculation of the CO_2 compensation point (Γ, Pa). The implementation of photosynthetic biochemistry in 1D-BGC is closely related to the BIOME-BGC logic in that it is based on the biochemical model of Farquhar *et al.* (1980), but the resulting set of equations is somewhat different due to differences in the logical constraints applied. I solve a quadratic system of equations by elimination of C_i, instead of specifying a value for C_i as an initial condition. Other differences from BIOME-BGC include a more detailed dependence of the kinetic and enzyme activity parameters on temperature (Woodrow and Berry 1988), and a simplifying assumption that

empirically relates the maximum rate of electron transport to the maximum carboxylation velocity. (Wullschleger 1993). 1D-BGC does not currently include a treatment of the C_4 photosynthetic pathway, although such a treatment will be required for global applications.

Required inputs to the photosynthesis subroutine are the leaf-level conductance to CO_2, the daytime leaf maintenance respiration rate, the leaf nitrogen per unit projected leaf area, and daily meteorological variables including air pressure, daylight average air temperature, and PPFD. Conductance to CO_2, PPFD, and leaf nitrogen per unit projected leaf area (N_{area}, kg m^{-2}) differ between the sunlit and shaded canopy fractions, while all other inputs are assumed constant with canopy fraction. Separate estimates of instantaneous photosynthetic rate are made for the sunlit and shaded canopy fractions, and total daily canopy photosynthesis is estimated as these instantaneous rates weighted by projected leaf area in their respective fractions and multiplied by daylength.

The leaf-level conductance is estimated directly from g_{Tv}, under the assumption that the transport pathway for water vapor and CO_2 between the leaf interior and the canopy air are equivalent. The ratio of conductances for water vapor and CO_2 is then simply the ratio of their molecular weights (Nobel 1991), which is 1.0/1.6. Because the photosynthesis submodel uses mole units instead of mass units to estimate CO_2 flux, a unit conversion for conductance is required (from (m s^{-1}) to (μmol m^{-2} s^{-1} Pa^{-1}), Jones 1992). Together, these conversions result in the following leaf-level mole-based conductance for CO_2 (g_{mTc}):

$$g_{mTc} = 1e6\ g_{Tv} / (1.6\ R\ (T_{day} + 273.15))$$

where R is the universal gas constant (8.3143 m^3 Pa mol^{-1} K^{-1}).

The estimation of daytime leaf maintenance respiration rate (CF_{ldmr}) is described in a later section. For the purpose of the photosynthesis submodel, a unit conversion is required, from the usual units (kg m^{-2} ground area d^{-1}) to mole units on a projected leaf area basis (MR_{ld}, μmol m^{-2} s^{-1}), using the molecular weight of CO_2 (12.011 g mol^{-1}) and the daylength:

$$MR_{ld} = CF_{ldmr} / (12.011e\text{-}9\ L_P\ t_{light})$$

The concentration of leaf nitrogen per unit of leaf C (N_{mass}, the inverse of leaf C:N ratio) is prescribed and assumed constant for any given canopy (although variable by species or vegetation type, Field and Mooney 1986), but the concentration of leaf nitrogen per unit of projected leaf area, N_{area}, varies

within a canopy between the sunlit and shaded portions as a function of changing specific leaf area (SLA, m^2 projected leaf area m^{-2} ground area). On any given daily timestep, the total leaf area (L_P) is determined on the basis of the total leaf C (C_l) and a prescribed average canopy SLA (SLA_{avg}), as:

$$L_P = C_l \, SLA_{avg}$$

After calculating L_{sun} and L_{shade} (section 3.2), the specific leaf area for the sunlit and shaded canopy fractions (SLA_{sun}, SLA_{shade}) are defined on the basis of the prescribed ratio between them ($SLA_{ratio} = SLA_{shade} / SLA_{sun}$) as:

$$SLA_{sun} = \frac{L_{sun} + \frac{L_{shade}}{SLA_{ratio}}}{C_l}$$

$$SLA_{shade} = SLA_{sun} \, SLA_{ratio}$$

Values for leaf N per unit of projected leaf area in the sunlit and shaded canopy fractions (N'_{sun} and N'_{shade}, kg m^{-2}) are then defined as:

$$N'_{sun} = \frac{N_{mass}}{SLA_{sun}}$$

$$N'_{shade} = \frac{N_{mass}}{SLA_{shade}}$$

My assumption of constant N_{mass} through time and with canopy depth is supported by several recent studies. Pereira (1995) found that the primary response of increased N availability in Eucalyptus forests is an increase in leaf area, rather than an increase in N_{mass}. Dewar (1996) presents a theoretical model of canopy nitrogen distribution, and reports that the model prediction of changes in leaf area as opposed to changes in N_{mass} under fertilization is supported by many observations from field-grown plants, but contrary to other observations from potted-plant studies. Similar results are reported by Reich and Walters (1994) for a variety of broadleaf rainforest communities in Venezuela, and by Ellsworth and Reich (1993) and Reich *et al.* (1995) for temperate deciduous broadleaf and evergreen needleleaf canopies, where the dominant response to a gradient in canopy radiation was an increase in SLA and N_{area} with depth (constant N_{mass}), and that changes in N_{mass} were of secondary importance and occurred mostly in response to increasing leaf age. Hollinger (1996) found a similar pattern for a broadleaf Nothofagus canopy, and

reported that while the response of decreasing SLA with increasing PPFD was very significant and explained much of the canopy variation in N_{area}, the actual canopy variation in SLA was not as strong as was predicted from a model of optimal N_{area} variation. Haxeltine and Prentice (1996b) suggest on the basis of simulations and observations of N_{mass} over a temperature gradient in deciduous broad-leaved forest that the primary influence of N limitation over large spatial scales would be on leaf area rather than on N_{mass}.

A significant change in N_{mass} is reported for sclerophylous shrubs in California (Field 1983, Field and Mooney 1983), where the primary mode of correlation between N_{mass} and mass-based photosynthetic rate was due to changes in specific leaf area at relatively constant N_{area}, and was poorly correlated with the variation in radiation at characteristic leaf age class positions in the canopy. It seems likely that this response is due to N dilution through the accumulation of low-N compounds (e.g. starch) in the aging leaves, which may be acting as temporary photosynthate storage organs (Randlett *et al.* 1996). In this case, photosynthesis on a mass basis (A_{mass}) is reduced by the dilution effect, but photosynthesis on an area basis (A_{area}) remains relatively constant (Field 1983, Reich and Walters 1994).

Sims and Pearcy (1989) found a constant N_{mass} but an increase in SLA, N_{area}, and A_{area} in response to increasing PPFD during growth for a tropical understory herb and a related crop species. Hirose and Werger (1994) showed that for an herbaceous canopy with linear leaves, variation in N_{mass} dominated the response of N_{area} to varying PPFD through the canopy for individual species, but that between-species variation was mostly the result of differing SLA. Because different species occupied characteristic canopy positions, this result can be interpreted as a community-level response dominated by changing SLA. Hikosaka *et al.* (1994) performed experiments on a vine to assess the relative importance of leaf age and PPFD in determining N_{area}, and they found the light response to be greater under nonlimiting N conditions, but that both responses are strong under N limitation. They do not report SLA, so the importance of variation in N_{mass} vs. SLA in the N_{area} response could not be determined.

The constant N_{mass} assumption with varying SLA is supported by the data reviewed by Field and Mooney (1986) for evergreen shrubs and trees and for South African shrubs, but not for old field annuals, and only partly for Death Valley annuals and deciduous chaparral shrubs. The variation in N_{mass} in the chaparral shrubs may be due to N-dilution not related to changes in the photosynthetic system, as already

noted, and the same may be true for the annuals. The relationship between increasing N_{mass} and A_{mass} explored by Field and Mooney (1986) is probably most important in describing the variation in productivity between sites with different species compositions, or in describing the trajectory of productivity at a single site as the canopy species composition changes through time. Both of these kinds of community-level differences in N_{mass} and A_{mass} could be related to variation in the availability of soil mineral N. Since the explicit goal of this model is to represent changes over time in productivity and resource dynamics for a specified vegetation community, and since I am not concerned here with possible changes in species composition or vegetation growth form under the influence of changing environmental conditions or resource availability, I believe the constant N_{mass} assumption is reasonable. Predictions of productivity under temporal gradients in resource availability would be improved by a mechanism which allowed the simulation of shifting species composition, as provided, for example, by the biogeography models of Nielson (1993), Nielson and Marks (1994), Prentice *et al.* (1992), and Haxeltine and Prentice (1996a).

Observation typically shows a range of SLA_{ratio} between 2 and 3, depending on L_P, canopy extinction coefficient, and average irradiance (Sims and Pearcy 1989, Ellsworth and Reich 1993, Reich and Walters 1994, Reich *et al.* 1995, Hollinger 1996). I performed a series of simulations to test the sensitivity of canopy net photosynthesis (gross photosynthesis minus leaf maintenance respiration) to variation in SLA_{sun} with respect to SLA_{shade}, and found that the maximum assimilation occurred in the range of SLA_{ratio} from 1.5 to 6, with higher values for higher k_{PAR}. Simulation results are shown for the 200-year total net primary productivity of a pine forest developing in a Missoula-like climate, using an allocation scheme that excludes daily leaf growth (Figure 2a) and a scheme that includes daily leaf growth (Figure 2b). Allocation schemes are discussed in a later section, but for now notice that although the total productivity is higher with daily allocation, the pattern of sensitivity to SLA_{ratio} is the same for both schemes. In studies comparing predictions of canopy nitrogen distribution based on optimality arguments with observations from real canopies, the real canopies usually show a more homogeneous nitrogen distribution than the predicted optimal distribution (Field 1983, Hollinger 1996, de Pury and Farquhar 1997). This may result from the sensitivity of the real distributions to diurnally and seasonally variable canopy light environments (Medlyn 1996, de Pury and Farquhar 1997). Note in Figure 2 that the slope of the productivity vs. SLA_{ratio}

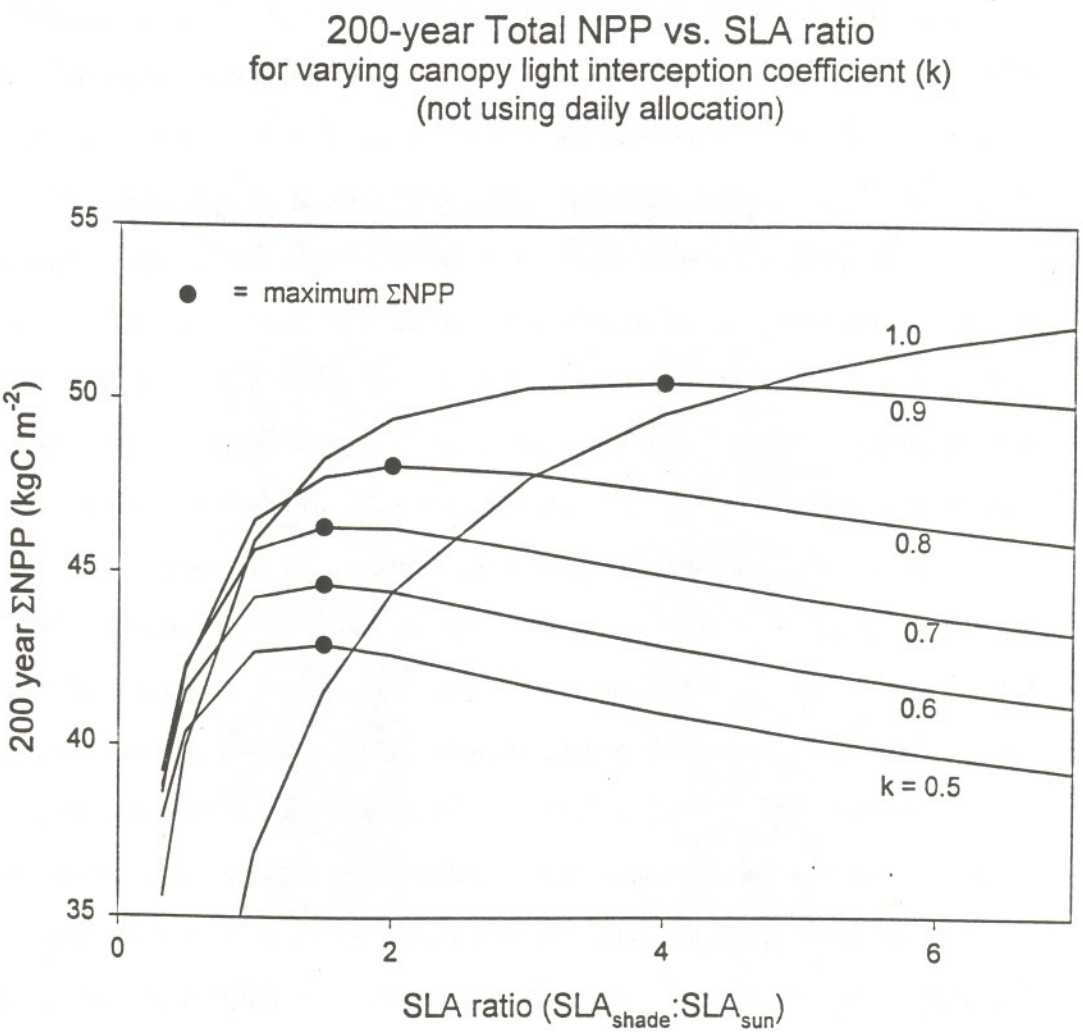

200-year Total NPP vs. SLA ratio
for varying canopy light interception coefficient (k)
(not using daily allocation)
● = maximum ΣNPP
1.0
0.9
0.8
0.7
0.6
k = 0.5
200 year ΣNPP (kgC m-2)
55
50
45
40
35
0
2
4
6
SLA ratio (SLAshade:SLAsun)

Figure 2a

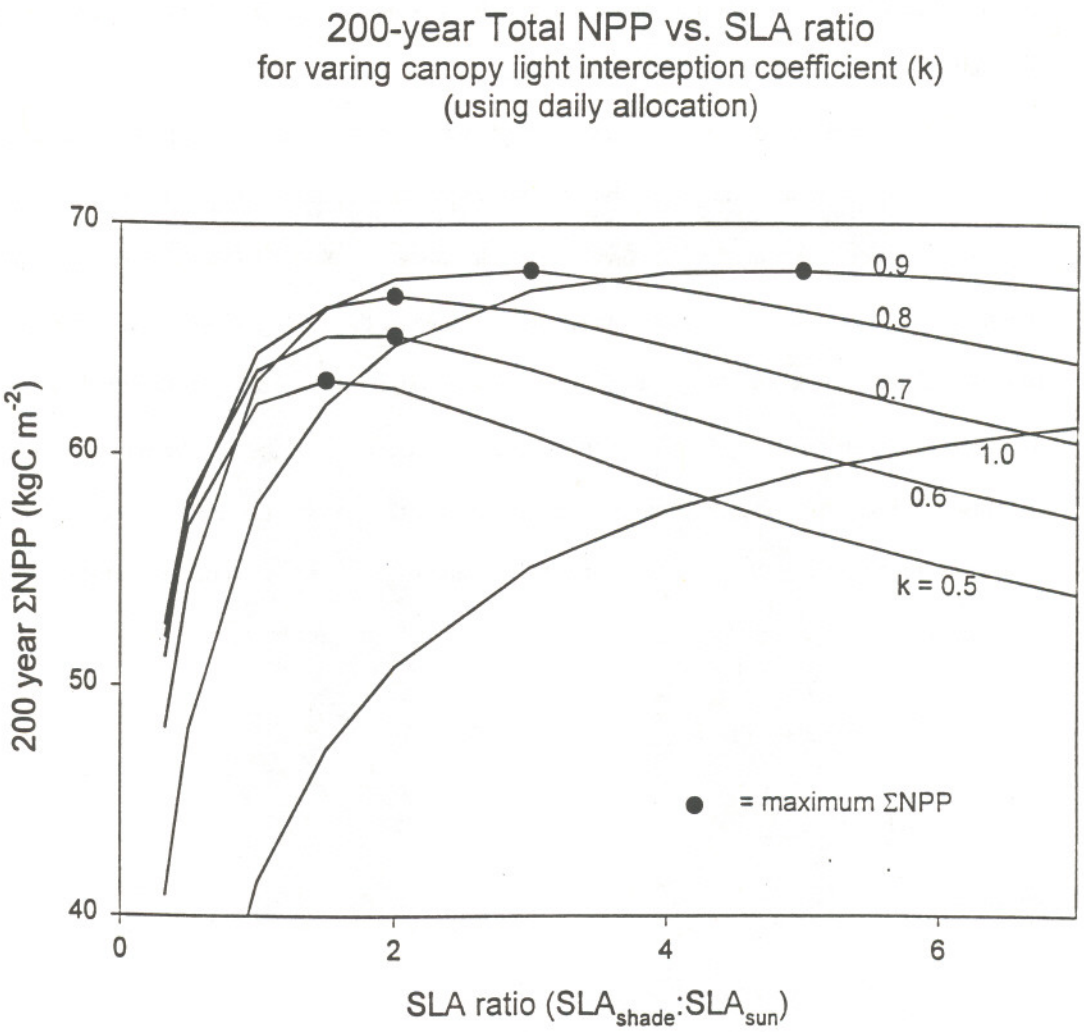

200-year Total NPP vs. SLA ratio
for varing canopy light interception coefficient (k)
(using daily allocation)
200 year ΣNPP (kgC m-2)
70
60
50
40
0
2
4
6
SLA ratio (SLAshade:SLAsun)
0.9
0.8
0.7
1.0
0.6
k = 0.5
= maximum ΣNPP

Figure 2b

curve is typically lower for values of the ratio above 2.0 than for values below 2.0. This suggests that the benefits to the plant of increasing SLA_{ratio} above 2.0 are small, and if there are significant costs associated with the plasticity in leaf morphology required to achieve high values of SLA_{ratio} it may be that the true optimum would be closer to 2.0 for canopies having light extinction coefficients close to 1.0, for example in broadleaved or heliotropic canopies (Jones 1992).

Given the required inputs, just described, photosynthesis rates for the sunlit and shaded canopy fractions are calculated according to the biochemical model for C_3 photosynthesis described by Farquhar *et al.* (1980). In the present implementation of that model, the actual rate of assimilation (A, μmol m^{-2} s^{-1}) is calculated as the lesser of two potential rates, one controlled by the activity of the carboxylating enzyme (ribulose-1,5-bisphosphate carboxylase-oxygenase, "rubisco") and the other controlled by the regeneration of the carboxylation substrate. Rubisco activity (ACT, μmol CO_2 kg^{-1} rubisco s^{-1}) and the kinetic constants for rubisco carboxylation and oxygenation reactions (K_c and K_o, respectively, Pa) are corrected for the daylight average temperature on the basis of the Q_{10} relationships given by Woodrow and Berry (1988), which are, for $T_{day} > 15$ °C:

$$ACT = ACT_{25} Q_{10ACT}^{\left(\frac{T_{day}-25}{10}\right)}$$

$$K_c = K_{c25} Q_{10K_c}^{\left(\frac{T_{day}-25}{10}\right)}$$

and, for $T_{day} <= 15$°C:

$$ACT = \frac{ACT_{25} 1.8 Q_{10ACT}^{\left(\frac{T_{day}-15}{10}\right)}}{Q_{10ACT}}$$

$$K_c = \frac{K_{c25} 1.8 Q_{10K_c}^{\left(\frac{T_{day}-15}{10}\right)}}{Q_{10K_c}}$$

and, for any T_{day}:

$$K_o = K_{o25} Q_{10K_o}^{\left(\frac{Tday-25}{10}\right)}$$

where: act_{25}, K_{c25}, and K_{o25} are the values at 25°C for ACT, K_c, and K_o; and Q_{10ACT}, Q_{10Kc}, and Q_{10Ko} are coefficients describing the sensitivity to changing temperature of ACT, K_c, and K_o, respectively. The CO_2 compensation point in the absence of leaf maintenance respiration (Γ^*, Pa) is calculated from the formula given in Farquhar et al (1980):

$$\Gamma^* = 0.105\ K_c\ O_2/K_o$$

where O_2 is the atmospheric partial pressure of oxygen (Pa).

The maximum rates of carboxylation (V_{cmax}, μmol m^{-2} s^{-1}) for sunlit and shaded canopy fractions are calculated as:

$$V_{cmax.sun} = N'_{sun}\ f_{Nrub}\ 7.16\ ACT$$

$$V_{cmax.shade} = N'_{shade}\ f_{Nrub}\ 7.16\ ACT$$

where f_{Nrub} is the fraction of leaf nitrogen in rubisco, prescribed and held constant for both sunlit and shaded fractions, and 7.16 is the weight proportion of rubisco to its nitrogen content, derived from data for the enzyme in spinach (Kuehn and McFadden 1969a, Kuehn and McFadden 1969b, Fasman 1976). The parameter f_{Nrub} relates total leaf nitrogen to the amount in the dominant photosynthetic enzyme. This value varies over different vegetation types, from typical values of 0.1 for evergreen needleleaf trees, to 0.3 for agricultural crop species (Evans 1989).

The maximum rate of electron transport (J_{max}, μmol m^{-2} s^{-1}) is related to V_{cmax} on the basis of the empirical relationship established by Wullschleger (1993) for a wide range of species:

$$J_{max} = 2.1\ V_{cmax}$$

The potential rate of electron transport (J, μmol m^{-2} s^{-1}) depends on J_{max} and PPFD, and is found as the smaller root of the quadratic solution to the following equation given by de Pury and Farquhar (1997):

$$0.7\ J^2 - (J_{max} + (PPFD/2))J + J_{max}\ (PPFD/2) = 0$$

Values for J are found for both the sunlit and shaded canopy fractions, using the appropriate values for V_{cmax}, J_{max}, and PPFD.

Finally, the following system of three equations, describing the constraints on A due to diffusion of CO_2 into the leaf (eq. 5.1), the rate of carboxylation (eq. 5.2), and the electron-transport driven rate of substrate regeneration (eq. 5.3), can be solved:

Eq. 5.1 $A_{(v\ or\ j)} = g_{mTc}\ (C_a - C_i)$

Eq. 5.2 $$A_v = \frac{V_{c\,max}\left(C_i - \Gamma^*\right)}{C_i + K_c\left(1 + O_2/K_o\right)} - MR_{ld}$$

Eq. 5.3 $$A_j = \frac{J\left(C_i - \Gamma^*\right)}{4.5C_i + 10.5\Gamma^*} - MR_{ld}$$

by solving Eq. 5.1 for C_i and substituting to eliminate C_i from Eqs. 5.2 and 5.3. Quadratic solutions for A_v and A_j are then found as:

$$A_v = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$$

where

$$a = -1/g_{mTc}$$

$$b = C_a + \frac{\left(V_{c\,max} - MR_{ld}\right)}{g_{mTc}} + K_c\left(1 + \frac{O_2}{K_o}\right)$$

$$c = V_{c\,max}\left(\Gamma^* - C_a\right) + MR_{ld}\left(C_a + K_c\left(1 + \frac{O_2}{K_o}\right)\right)$$

and

$$A_j = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$$

where

$$a = -4.5/g_{mTc}$$

$$b = 4.5C_a + 10.5\Gamma^* + \frac{J - 4.5MR_{ld}}{g_{mTc}}$$

$$c = J\left(\Gamma^* - C_a\right) + MR_{ld}\left(4.5C_a + 10.5\Gamma^*\right)$$

This system of equations is solved once during each daily timestep for each of the canopy fractions, substituting the generic values of g_{mTc}, V_{cmax}, MR_{ld}, and J shown above with their values for the sunlit and shaded canopy fractions, in turn. In each fraction, the daylight average instantaneous assimilation rate (A) is found as the lesser of A_v and A_j. Daily gross accumulation of carbon due to photosynthesis ($CF_{grossPSN}$) is

then found by multiplication of the average instantaneous rates by the daylength, weighting for the canopy leaf area in the two fractions, converting from mole to carbon units, and adding back the daytime leaf maintenance respiration that was subtracted for the instantaneous rate calculations, to give:

$$CF_A = (A_{sun} L_{sun} + A_{shade} L_{shade}) t_{light} 12.011e\text{-}9 + CF_{ldmr}$$

Calculations for gross photosynthesis are carried out only on timesteps for which leaf area is non-zero.

5.2 Maintenance respiration

The rates of maintenance respiration observed in different plant parts generally reflect the amount of living tissue in those parts (Kramer and Kozlowski 1979). Maintenance respiration in 1D-BGC is calculated as a function of tissue nitrogen concentration, using the same base respiration rate for all live plant compartments and making corrections for temperature. The empirical model of Ryan (1991) is used, in which daily total maintenance respiration at 20°C (CF_{mr20}) is determined for each compartment on the basis of the total compartment N:

$$CF_{mr20} = 0.218\ N$$

where the constant 0.218 is derived from Ryan's (1991) regressions, and represents the mass of carbon respired per mass of tissue nitrogen per day at 20°C. The actual daily maintenance respiration is calculated assuming a Q_{10} relationship with $Q_{10} = 2.0$ for all components, as:

$$CF_{mr} = CF_{mr20}\, Q_{10}^{(T-20)/10}$$

where T is the temperature relevant to the particular tissue respiration rate. For live stem tissue, $T=T_{avg}$. For fine roots and live coarse root tissue, $T=T_{soil}$. For leaves, a separate maintenance respiration rate is calculated for the day and for the night (see section 5.1), with the daylight rate calculation using $T=T_{day}$, and the night rate calculation using $T=T_{night}$. The leaf rates are also scaled for the daylength, so that they sum to the total daily respiration. Note that the dead stem wood and dead coarse root components are assumed to consist entirely of conducting xylem and/or heartwood, with any parenchyma which might in reality extend through these dead components assigned by definition to the live stem wood and live coarse root compartments. Dead woody compartments thus have no maintenance respiration associated with them.

5.3 Growth respiration

This was wrong. I misinterpreted Larcher's Table 6.1. Correct values are about 0.3 for both woody and nonwoody growth.

Growth respiration is calculated on the basis of the construction costs for different tissue types, with one construction cost applied to woody tissue growth, and another applied to non-woody tissue growth. From Table 6.1 in Larcher (1995), I estimate the growth costs in terms of mass of respired carbon per mass of new tissue carbon for woody tissue (G_w, for live and dead stem wood and coarse root) to be 2.0, and for non-woody tissue (G_n, for leaf and fine root) to be 1.2. These values are taken as constants, with the total growth costs determined by multiplying the new growth by the relevant constant. A more sophisticated treatment of the growth respiration costs could be implemented, by changing the growth constants as a function of tissue type, for example by implementing different values of $G_{nonwood}$ for different leaf growth forms. Since the variation in construction costs as indicated by the literature reviewed by Larcher (1995) is small, the current simple treatment may introduce errors which are small in comparison to other sources of unaccounted variation.

5.4 Leaf and fine root phenology

In previous versions of BGC the timing of new leaf and fine root appearance and litterfall was prescribed, with the total annual new leaf and fine root growth appearing on a single day and the total annual litterfall also occurring on a single day. The timing of leaf growth and litterfall events and the duration of the growing season greatly influence productivity (Pereira 1995), and there is abundant evidence that this timing is sensitive to climatic and meteorological conditions (e.g. Sparks and Carey 1995, also see review in White *et al.* 1997). In addition to provisions for prescribed dates of new growth and litterfall, I have incorporated the phenological model for temperate deciduous trees and temperate grasslands described in White *et al.* (1997). I have also extended the predictions from that model to evergreen vegetation. Based on long-term and current growth season meteorological data, the model predicts the date for 50% of new leaf expansion in the spring and for 50% of annual canopy loss in the fall. Instead of having all growth and litterfall occur on a single day, I allow for prescribed periods of new leaf growth and litterfall. These parameters are not generated by the phenological model. The timing of new leaf growth and fine root growth are assumed to be the same. This is likely a gross over-simplification of the general pattern of fine root growth and litterfall (Vogt *et al.* 1986, Nadelhoffer and Raich 1992, Pregitzer *et al.* 1993, Holmes and Zak 1994). The measurement of fine root growth and turnover is very

difficult and subject to a wide range of errors (Nadelhoffer and Raich 1992, Beets and Whitehead 1996), and so a more sophisticated treatment of fine root dynamics is difficult to justify. I choose to follow the logic of Raich and Nadelhoffer (1989) and Nadelhoffer and Raich (1992) in connecting the global-scale patterns of belowground production and turnover to leaf production and litterfall.

Justification for the form of the phenology model and details of its development and testing are found in White *et al.* (1997). Here I will describe the details relevant to the integration of the phenology model into the larger structure of 1D-BGC. I assume that the same logic developed for deciduous tree phenology applies to new leaf growth and leaf litterfall for evergreen canopies as well. An evergreen canopy is defined as having a leaf turnover rate less than 1.0 yr^{-1}.

The phenology model consists of two independently formulated components, one for predictions of dates of onset and offset in forest systems, and the other for such predictions in grasslands. For trees, the date of onset is predicted as a function of the summation (referenced from phenological yearday 1, which is defined as January 1 in the northern hemisphere and July 1 in the southern hemisphere) of the daily soil temperature values that exceed 0 °C. When this summation exceeds a critical value (Ph_{treeon}), onset is predicted to occur. The critical soil temperature summation is a function of the long-term mean daily average temperature ($T_{avg.clim}$) at a given site:

$$Ph_{treeon} = \exp(4.795 + 0.129T_{avg.clim})$$

The date of offset for trees is predicted as a function of daylength and soil temperature. Two scenarios can result in offset: (1) if midsummer has past, and the daylength is less than 10.9 hours, and the soil temperature has fallen below the long-term average soil temperature for the range of phenological yeardays between 244 and 305, or (2) if midsummer has past and the soil temperature is less than 2 °C.

Predictions of onset and offset in grasses is more complicated, and less accurate (White *et al.* 1997). Predicted onset depends on the annual summation of daily soil temperature in excess of 0 °C and on the summation of daily total precipitation, referenced from phenological yearday 1. These summations are compared to critical values for the soil temperature sum ($Ph_{grass,stcrit}$) and the precipitation sum ($Ph_{grass,pcrit}$) which are given as:

$$Ph_{grass,stcrit} = (0.5\ (1380.0 - 418.0)\ (t_1\text{-}1)/(t_1+1)) + 900.0$$

$$t_l = \exp(32.9\ (T_{avg.clim} - 9.0))$$

$$Ph_{grass,pcrit} = 0.15\ PRCP_{avgann}$$

where $PRCP_{avgann}$ is the long-term average annual total precipitation (cm). The conditions for onset in grasslands are that the annual soil temperature summation exceed $Ph_{grass,stcrit}$ and that the annual precipitation summation exceed $Ph_{grass,pcrit}$. Offset in grasslands is predicted to occur under two sets of conditions: hot and dry, or cold. Offset due to hot and dry conditions is not permitted within one month of onset, and offset due to cold conditions is not permitted before phenological yearday 182. The hot and dry condition for offset requires that the previous 31-day total precipitation be less than 1.14 cm, that the next 7-day total precipitation be less than 0.97 cm, and that the current daily maximum temperature be greater than 0.92 times the long-term maximum value of the daily maximum temperature. The cold condition for offset requires that the 3-day running average of daily minimum temperature be less than the long-term average daily minimum temperature.

5.5 Whole-plant mortality

In addition to the seasonal loss of leaves and fine roots, 1D-BGC includes a mechanism for the transfer of carbon and nitrogen from all plant compartments into the litter. This transfer is prescribed as an annual fraction of the plant biomass, which is then scaled to a daily loss rate and implemented once each day. For example, if the prescribed annual mortality fraction is 0.01 (1% per year), on each day an amount equal to 0.0000274 times the compartment mass would be removed from each of the plant compartments, and sent to the litter compartments. This is functionally equivalent to losses due to wind-throw or insect infestation or disease, in which individuals are removed from a stand, resulting in a loss of leaf, stem, coarse root, and fine root biomass in proportion to their ratios in the stand as a whole. In a woody system, this is the only pathway by which stem wood and coarse root wood (live or dead) can leave the plant pools and enter the litter pools. In earlier BGC versions, a turnover fraction for stem and coarse root was specified, which served a similar purpose.

In addition to the removal of mass from the plant compartments already mentioned, mass in the abstract storage pools for carbon and nitrogen is also subject to mortality losses. Although these compartments are not associated with any particular plant tissue, they are supposed to represent actual

storage mechanisms which could be localized in different compartments during different parts of the growth cycle (Kramer and Kozlowski 1979).

Observations of mortality are difficult and rare because they require permanent plots that are maintained over long periods of time (Sollins 1982). Data from Sollins (1982) indicates a relatively narrow range of mortality rates for stands of different ages and different species dominance (Table 2).

Table 2. Annual mortality fraction from data in Sollins (1982).

Forest type	age	AGB[1]	AGI[2]	AGMF[3]
Douglas fir	95	558	2.2	0.0039
Hemlock-spruce	105	647	3.1	0.0048
Douglas fir	450	654	3.0	0.0046
Douglas fir	450	654	4.5	0.0069
Mean	275	628	3.2	0.0051

1. AGB = above ground biomass (Mg ha^{-1})
2. AGI = above ground mortality inputs (Mg ha^{-1} yr^{-1})
3. AGMF = above ground mortality fraction (yr^{-1})

Other data for deciduous forest types and for both evergreen and deciduous types in different climates is desirable, but I have not yet obtained any. As will be discussed below, inputs of stem and root wood to the litter compartments have an important impact on the cycling of nitrogen through the litter, soil, and plant compartments, and so an accurate representation of long-term mortality rates is important in cases where the long-term dynamics of the nitrogen cycle are of interest. For clarification, this mortality does not represent disturbances, such as fire or grazing, that result in system losses of carbon and/or nitrogen. These disturbance processes are themselves important, but it is beyond my scope in this project to address them in any detail. I will, in a later discussion, address the importance of such disturbances on the model predictions of long-term equilibrium in the plant and soil carbon and nitrogen budgets.

5.6 Litter and soil carbon and nitrogen budgets

A representation of cycles of carbon and nitrogen in litter and soil is required because the availability of nitrogen for new plant growth is directly linked to the inputs of dead plant material to the litter and the processing of these materials by microorganisms to turn them into soil organic matter. In addition, the release of carbon dioxide to the atmosphere during this decomposition process is an important component of the carbon budget in the global biosphere, and many questions relevant to the understanding of global terrestrial ecosystem responses to variable climate and increasing atmospheric concentrations of

carbon dioxide can not be answered without a consideration of these processes (Rastetter *et al.* 1991, Field *et al.* 1992, Potter *et al.* 1993, Randerson *et al.* 1996, Hunt *et al.* 1996).

The fundamental assumption underlying my treatment of litter and soil processes is that the process of decomposition can be treated as an exponential decay, with recognizable transformations between components, and where each component has a characteristic decay rate (van Veen *et al.* 1984). Plant litter chemical composition is quite complex and variable, and the most accurate description of decay would treat the process along a continuum of substrate quality and decomposition rate (Carpenter 1981, Bosatta and Ågren 1985). It has been recognized for some time, however, that a realistic representation of nitrogen dynamics during the decay process is possible by considering only a small number of chemically distinct substrate materials (Melin 1930, Richards and Norman 1931, van Veen and Paul 1981, Hunt *et al.* 1988). Many studies have related rates of mass loss and nitrogen cycling to the chemical qualities of litter by considering the fractionation of heterogeneous litter materials into chemically homogeneous components (Minderman 1967, Hunt 1977, Aber and Melillo 1980, Sørensen 1981, Edmonds 1984, Parton *et al.* 1987, Berg and Tamm 1994). One common classification of the chemical fractions distinguishes between the litter component soluble in water and alcohol-benzene (typically proteins, starches and sugars), the part of the remaining component soluble in acid (typically cellulose and hemicellulose), and the non-soluble fraction (mostly lignin) (e.g. Taylor *et al.* 1989). These fractions are relatively easily determined and so have been measured for many different types of litter, providing a strong empirical base for simulations across a variety of vegetation types (e.g. Taylor *et al.* 1989, Aber *et al.* 1990). Just as important, from the point of view of numerical simulations of the decomposition process, is the fact that these classes have very distinct rates of decomposition under the same conditions of temperature, and water and nitrogen availability (Minderman 1967, Sørensen 1981, Entry *et al.* 1987, Aber *et al.* 1990). It is also recognized that each of these broad classes of litter material is acted upon by a characteristic subset of the soil taxa (Stevenson 1986), and so there is both a chemical and a biological basis for supposing that these litter fractions should have distinct decomposition dynamics.

Decomposition of a fresh litter substrate begins with its colonization by soil microorganisms (mostly bacteria and fungi) which metabolize the substrate through various enzymatic reactions in order to

obtain energy, carbon, and nutrients for growth, reproduction, and maintenance (Bolin 1983, Stevenson 1986). The resulting mass of new microbial cells is less than the mass of the metabolized litter due to respiration required to produce enzymes, perform reproductive functions, and build and maintain cellular structure. These respiration losses occur as carbon dioxide released to the atmosphere, and they can be quantified as a fraction of the total mass loss of the substrate (van Veen and Paul 1981, van Veen *et al.* 1984). The non-respired fraction is assimilated into the microbial biomass, initiating a process of secondary decay as this biomass is itself metabolized by other microbes (Waksman and Gerretsen 1931, Hunt 1977). This process of biomass recycling operates over multiple generations, usually through a complex series of taxa (Stevenson 1986), and results in progressive degradation in substrate quality for each successive generation, with concomitant losses in biomass due to respiration at each step.

This thought-model of the decomposition process can be translated to a simple compartmentalized flux model, as illustrated in Figure 3a, where compartment L is the unmetabolized litter material and compartment S1 is the soil organic matter, which consists in this simple case of the entire soil microbial biomass as well as any metabolic waste products that are accumulating as the result of microbial activity. This is a discrete (as opposed to continuous) time model, and assuming a timestep of one day, the flux out of L and out of S1 on each timestep (F_L, F_{S1}) is given as:

$$F_L = M_L\, k'_L$$

$$F_{S1} = M_{S1}\, k'_{S1}$$

where M_L and M_{S1} are the masses in L and S1, and k'_{S1} and k'_L are the discrete-time model rate constants for fluxes from L and S1, with units of (day^{-1}). Respiration fluxes are indicated by the curved arrows leaving L and S1. For L, the respiration flux (F_{Lr}) is calculated as a function of the F_L and the respiration fraction (f_L), as:

$$F_{Lr} = f_L\, F_L$$

For S1, in this simple formulation, it is assumed that the entire flux leaving the compartment occurs as respiration, since all other fluxes result in new microbial biomass or new microbial byproducts assumed to be a part of S1.

Decomposition Model Structures

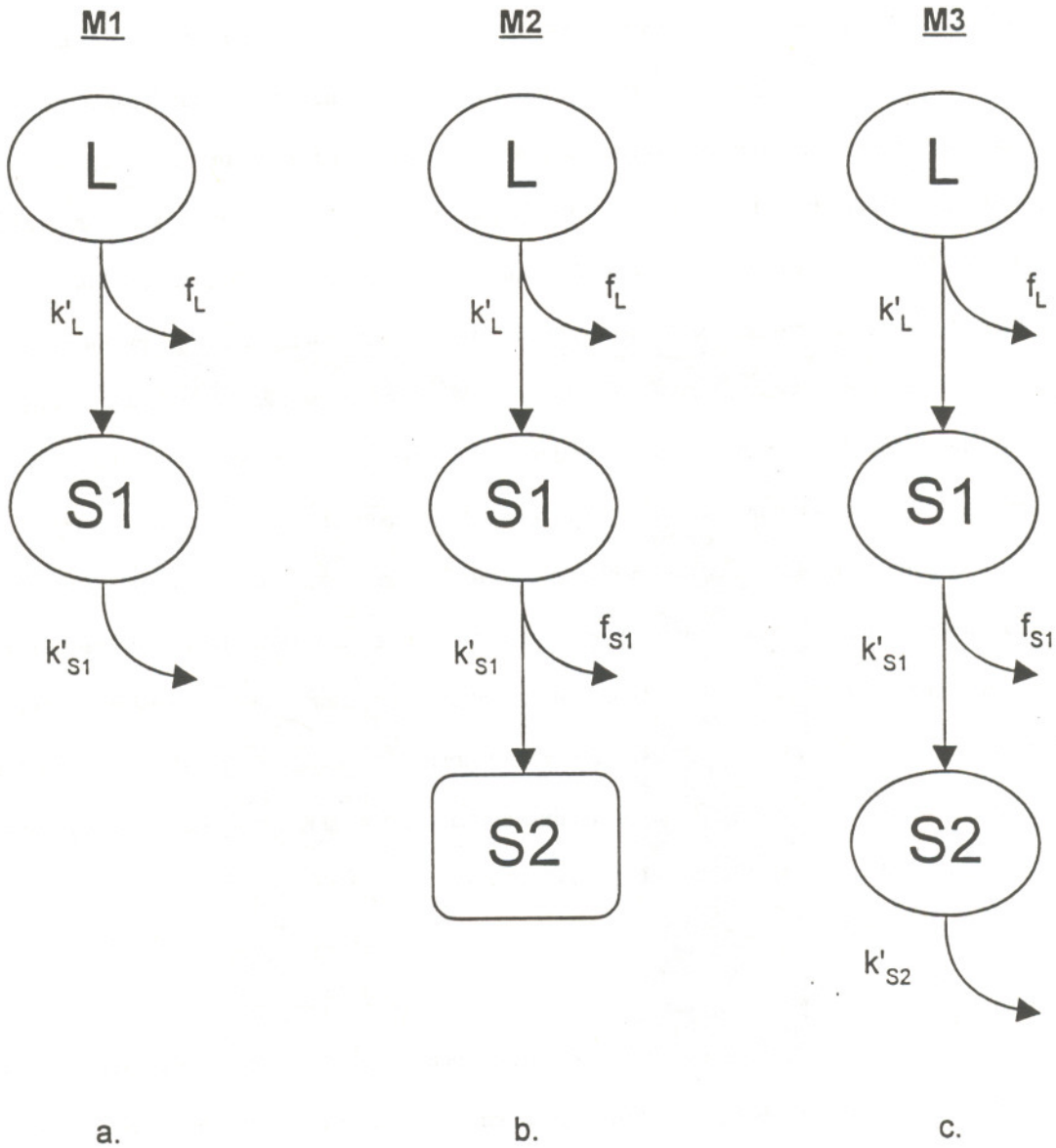


Figure 3

This configuration of compartments and fluxes provides the basis for a powerful numerical tool by which the dynamics of homogeneous substrate decomposition can be explored. For the case of an experiment in which a known quantity of fresh homogeneous substrate is allowed to decompose under controlled conditions while the release of CO_2 is either continuously of periodically monitored, and under the assumption that the rate of metabolism of L is substantially higher than the rate of subsequent respiration from S1, the fraction of original mass remaining (y) at time t can be approximated by a double exponential model, hereafter referred to as M1, and defined as:

M1 $y = a_L \exp(-k_L t) + a_{S1} \exp(-k_{S1} t)$

where k_L and k_{S1} are the continuous-time counterparts of k'_L and k'_{S1}, given in general as:

$$k = - \ln(1-k')$$

(Olson 1963). The assumption of $k_L >> k_{S1}$ is necessary because M1 assumes that L and S1 are initialized at time t=0 with fractions of the total input mass equal to a_L and a_{S1}, when in fact the experiment begins with all of the mass in L and nothing in S1. Given $k_L >> k_{S1}$, however, the transfer of the unrespired proportion of the input substrate mass from L to S1 (implicitly undergoing transformation to microbial biomass and metabolic byproduct along the way) is rapid with respect to its subsequent disappearance from S1, and the errors in estimation of a_L will be small.

An obvious extension of M1 is to include more than one microbial biomass recycling compartment. These might be in parallel with each other, representative of two or more distinct microbial communities metabolizing the same substrate, or they might be in series, which amounts to a finer resolution of the microbial recycling process, in which the metabolism of first-generation microbes and their byproducts by other members of the soil microfauna is treated explicitly. The case of two soil organic matter compartments in series with the litter substrate is illustrated in Figure 3c, and can be represented as a triple exponential model (M3). An intermediate stage between M1 and M3 can also be entertained, in which the last compartment has no decomposition rate, and represents a permanent storage pool (M2, Figure 3b). The exponential models which approximate M2 and M3 are given as:

M2 $y = a_L \exp(-k_L t) + a_{S1} \exp(-k_{S1} t) + y_0$

M3 $y = a_L \exp(-k_L t) + a_{S1} \exp(-k_{S1} t) + a_{S2} \exp(-k_{S2} t)$

where k_L, k_{S1}, and k_{S2} are the continuous-time model decomposition rates for the compartments L, S1, and S2 in Figure 3, and f_L and f_{S1} in Figure 3 are assumed for the moment to be related to a_L and a_{S1} as:

$$f_L = a_L$$

$$f_{S1} = a_{S1} / (1-a_L)$$

These numerical tools can be difficult to apply in practice. Because 1D-BGC is designed to estimate processes occurring under field conditions, the relevant decomposition rates will be for homogeneous litter substrate in association with natural assemblages of soil microorganisms. In the case of a soil sample collected in the field for laboratory testing of litter decomposition under controlled environmental conditions, native soil organic matter will be abundant relative to the new biomass and microbial byproducts generated as the result of experimental additions. Control experiments are possible, but there are many confounding errors related to differences in native soil organic matter respiration rates under varying conditions of substrate input (van Veen and Paul 1981). Tracer experiments using organic matter labeled with the ^{14}C isotope avoid many of the problems associated with non-labeled decomposition experiments (Voroney and Paul 1984). Many such experiments have been reported, and the methods lend themselves to the illumination of organic matter transformations along particular pathways in the litter-soil system (Jenkinson 1971).

In order to assess the differences between the decomposition dynamics of the major homogeneous litter fractions, I collected studies reporting the release of ^{14}C-CO_2 from labeled glucose, cellulose, and lignin. To be included in my assessment, these studies had to use pure glucose, cellulose, or lignin, and not, for example, mixtures of cellulose and lignin, as in the case of many studies on straw or other labeled plant parts. Because the rates of decomposition depend on temperature and soil water content, all included studies were conducted with temperature fixed between 20 and 30 °C, and with soil moisture maintained at or above 60% of saturated capacity. Using Sigmaplot 4.0 (SPSS Inc., © 1997), numerical methods were used to fit each of the three models discussed above (M1, M2, and M3) to observations of mass loss due to respiration from single-component decomposition experiments. In addition, for each homogeneous substrate component in question, a fit was performed for each model on the pooled data from all relevant experiments. Least-squares fitting results from these analyses are shown in Table 3.

Table 3. Regression results for ^{14}C decomposition experiments

Values in parentheses are standard errors.

Glucose experiments, fit to M1

a_L	k_L	k_{S1}	adj. r^2	ref
0.47 (0.02)	0.589 (0.072)	0.0058 (0.0009)	0.963	1
0.34 (0.01)	0.642 (0.087)[1]	0.0033 (0.0005)[1]	0.997	2
0.44 (0.02)	0.581 (0.089)[1]	0.0062 (0.0010)[1]	0.996	2
0.43 (0.02)	1.545 (0.184)	0.0020 (0.0002)	0.973	3
0.43 (0.03)	0.711 (0.157)	0.0104 (0.0024)	0.979	4
0.35 (0.01)	0.742 (0.050)	0.0025 (0.0020)	0.946	5
0.56 (0.03)	0.622 (0.067)	0.0038 (0.0008)	0.988	5
--- pooled data ---				
0.48 (0.02)	0.600 (0.057)	0.002 (0.001)	0.905	

Glucose experiments, fit to M2

a_L	k_L	a_{S1}	k_{S1}	adj r^2	ref
0.40 (0.02)	0.748 (0.116)	0.21 (0.02)	0.044 (0.012)	0.988	1
0.39 (0.01)[1]	0.761 (0.043)[1]	0.24 (0.01)[1]	0.036 (0.003)[2]	1.000	2
0.31 (0.01)	0.906 (0.066)	0.20 (0.01)	0.041 (0.009)	0.999	5
--- pooled data ---					
0.37 (0.03)	1.170 (0.270)	0.20 (0.02)	0.064 (0.018)	0.920	

Glucose experiments, fit to M3

a_L	k_L	a_{S1}	k_{S1}	k_{S2}	adj r^2	ref
0.39 (0.06)	0.785 (0.229)	0.16 (0.13)[1]	0.066 (0.101)[2]	0.0021 (0.0058)[2]	0.988	1
0.39 (0.01)	2.299 (0.314)	0.11 (0.02)	0.060 (0.021)	0.0011 (0.0002)	0.991	3
--- pooled data ---						
0.37 (0.03)	1.190 (0.290)	0.19 (0.03)	0.070 (0.030)[1]	0.0002 (0.0007)[2]	0.920	

Cellulose experiments, fit to M1

a_L	k_L	k_{S1}	adj. r^2	ref
0.74 (0.03)	0.050 (0.005)	0.0005 (0.0002)	0.979	6
0.66 (0.03)	0.049 (0.007)	0.0005 (0.0001)	0.972	6
0.68 (0.02)	0.045 (0.004)	0.0005 (0.0001)	0.987	6
0.57 (0.02)	0.060 (0.006)	0.0005 (0.0001)	0.986	6
--- pooled data ---				
0.66 (0.02)	0.051 (0.004)	0.0005 (0.0001)	0.960	

Cellulose experiments, fit to M2

a_L	k_L	a_{S1}	k_{S1}	adj r^2	ref
0.70 (0.05)	0.056 (0.008)	0.18 (0.04)	0.0030 (0.0020)[2]	0.981	6
0.37 (0.10)	0.125 (0.050)[1]	0.44 (0.10)	0.0140 (0.0040)	0.980	6
0.63 (0.03)	0.053 (0.006)	0.22 (0.03)	0.0029 (0.0010)[1]	0.990	6
0.53 (0.02)	0.070 (0.006)	0.26 (0.02)	0.0023 (0.0006)	0.991	6
--- pooled data ---					
0.61 (0.03)	0.060 (0.006)	0.23 (0.03)	0.0030 (0.0009)	0.964	

Cellulose experiments, fit to M3

a_L	k_L	a_{S1}	k_{S1}	k_{S2}	adj r^2	ref
0.42 (0.36)[2]	0.085 (0.057)[2]	0.37 (0.35)[2]	0.0230 (0.0180)[2]	0.0003 (0.0002)	0.980	6
0.29 (0.11)[1]	0.166 (0.100)[2]	0.46 (0.10)	0.0211 (0.0064)	0.0003 (0.0001)[1]	0.984	6
0.59 (0.19)	0.057 (0.018)	0.17 (0.13)[2]	0.0082 (0.0182)[2]	0.0003 (0.0002)[2]	0.990	6
--- pooled data ---						
0.48 (0.14)	0.077 (0.023)	0.25 (0.14)[2]	0.0145 (0.0091)[2]	0.0003 (0.0001)	0.964	

Lignin experiments, fit to M1

a_L	k_L	k_{S1}	adj. r^2	ref
0.44 (0.04)	0.0135 (0.0026)	0.0006 (0.0002)	0.965	7
0.25 (0.02)	0.0123 (0.0018)	0.0004 (0.0000)	0.986	7
0.26 (0.04)	0.0135 (0.0034)	0.0007 (0.0001)	0.972	7
0.15 (0.07)[1]	0.0079 (0.0052)[2]	0.0005 (0.0001)	0.953	7
--- pooled data ---				
0.27 (0.05)	0.0126 (0.0044)[1]	0.0005 (0.0001)	0.780	

Lignin experiments, fit to M2

a_L	k_L	a_{S1}	k_{S1}	adj r^2	ref
0.16 (0.08)[2]	0.0229 (0.0143)[2]	0.47 (0.05)	0.0024 (0.0013)[2]	0.972	7
0.05 (0.07)[2]	0.0241 (0.0460)[2]	0.48 (0.08)	0.0019 (0.0012)[2]	0.972	7
--- pooled data ---					
0.24 (0.19)[2]	0.0141 (0.0111)[2]	0.45 (0.52)[2]	0.0012 (0.0036)[2]	0.786	

Lignin experiments, fit to M3

a_L	k_L	a_{S1}	k_{S1}	k_{S2}	adj r^2	ref
--- pooled data ---						
0.24 (0.73)[2]	0.0141 (0.0271)[2]	0.38 (31.26)[2]	0.0014 (0.0810)[2]	0.0001 (0.0395)[2]	0.784	7

[1] not significant at $P < 0.01$
[2] not significant at $P < 0.05$
References: 1: Degens and Sparling 1996, 2: vanVeen *et al.* 1984, 3: Mary *et al.* 1993, 4: Sagaar *et al.* 1994, 5: Ladd *et al.* 1992, 6: Sørensen 1981, 7: Martin *et al.* 1980.

The purpose of this analysis is to identify the simplest model structure that accounts for the observed variation in respiration fractions and rates of transfer between compartments. Of the models included here, M3 is likely to be the most realistic representation of the true litter and soil organic matter dynamics, since it is recognized that litter material experiences successive stages of metabolism as it ages, with different stages having characteristic respiration fractions and rates (van Veen and Paul 1981). On the other hand, the number of observations in this analysis is too small to get highly significant estimates of the parameters for M3, especially for the lignin experiments. One possible interpretation of the results in Table 3 is that M1 parameters were most consistently significant, and so the true dynamics might be best represented using that model. Another possible interpretation is that M3 is in fact the better representation, and the low significance of many parameters is the result of too few observations, variations in the actual parameters between experiments, or experimental error.

Because of the approximate nature of the exponential models, owing to the difference between the assumption of instantaneous distribution into the various compartments and the reality of lagged distribution as biomass moves through progressive stages of decomposition, a theoretical assessment of the expected responses to changes in the underlying model was not attempted. Instead, I performed an empirical numerical experiment to test the behavior of the least-squares fitting procedure for cases in which the true underlying model behavior was known. I tested two hypotheses: first, that the true model was M1, and second, that the true model was M3, with parameters in both cases specified as in the regressions for pooled data in Table 3. In all I performed six experiments, by testing the M1 and M3 models parameterized from the regression results for pooled data on glucose, cellulose, and lignin decomposition.

For each experiment I prepared a data file by sampling the output of a numerical simulation model that operated on the basis of either M1 or M3, according to their structure as illustrated in Figure 3. Using parameters taken from Table 3 for the pooled data analysis for each substrate type and for fits to M1 and M3, I generated a 500-day sequence of total mass simulations, beginning, as in the ^{14}C experiments, with all of the mass in L and correcting the total mass on each timestep for the amounts respired from all compartments. From this 500-day sequence, I sampled 19 days to represent a sampling regime similar to that in the ^{14}C experiments, with more frequent sampling at the beginning than at the end of the experiment. I sampled on days 0, 1, 2, 4, 6, 9, 12, 16, 20, 30, 40, 60, 80, 120, 160, 240, 320, 400, and 499. In each experiment I performed the same numerical least-squares fitting procedure used for the analysis in Table 3, generating fits for M1, M2, and M3 to the test data. By comparing changes in fitted model parameters for M1, M2, and M3 between the ^{14}C experimental data and the numerical experiment data, I hoped to demonstrate that the patterns in the isotope data were more likely to have resulted from one or the other underlying model. Results of the numerical experiments are shown in Table 4, together with the results of pooled data regressions replicated from Table 3 for comparison.

Table 4. Results of numerical experiments for tests of decomposition models

Glucose

Fitted parameter	Fit to	Fitted values for M1 expt.	Fitted values for M3 expt.	Fitted values for ^{14}C expts.
a_L	M1	0.48	0.47	0.48
a_L	M2	0.48	0.37	0.37
a_L	M3	0.48	0.35	0.37
k_L	M1	0.5978	0.47	0.60
k_L	M2	0.5978	1.10	1.17
k_L	M3	0.5978	1.19	1.19
a_{S1}	M1	0.52	0.50	0.52
a_{S1}	M2	0.52	0.21	0.20
a_{S1}	M3	0.0	0.21	0.19
k_{S1}	M1	0.0020	0.0006	0.002
k_{S1}	M2	0.0020	0.052	0.064
k_{S1}	M3	0.1453	0.070	0.070

Cellulose

Fitted parameter	Fit to	Fitted values for M1 expt.	Fitted values for M3 expt.	Fitted values for ^{14}C expts.
a_L	M1	0.55	0.61	0.66
a_L	M2	0.54	0.45	0.61

a_L	M3	0.29	0.42	0.48
k_L	M1	0.0513	0.0496	0.0510
k_L	M2	0.0513	0.0722	0.0600
k_L	M3	0.0513	0.0770	0.0770
a_{S1}	M1	0.45	0.38	0.34
a_{S1}	M2	0.45	0.30	0.23
a_{S1}	M3	0.25	0.30	0.25
k_{S1}	M1	0.0005	0.0011	0.0005
k_{S1}	M2	0.0005	0.0111	0.0030
k_{S1}	M3	0.0513	0.0145	0.0145

Lignin

Fitted parameter	Fit to	Fitted values for M1 expt.	Fitted values for M3 expt.	Fitted values for ^{14}C expts.
a_L	M1	0.24	0.22	0.27
a_L	M2	0.24	0.20	0.24
a_L	M3	0.24	0.20	0.24
k_L	M1	0.0126	0.0130	0.0126
k_L	M2	0.0126	0.0141	0.0141
k_L	M3	0.0126	0.0141	0.0141
a_{S1}	M1	0.76	0.78	0.73
a_{S1}	M2	0.76	0.46	0.45
a_{S1}	M3	0.37	0.39	0.38
k_{S1}	M1	0.0005	0.0006	0.0005
k_{S1}	M2	0.0005	0.0013	0.0012
k_{S1}	M3	0.0005	0.0014	0.0014

The clear result from Table 4 is that the variation in parameters for the ^{14}C experimental data fitted to progressively more detailed models (from M1 to M2 to M3) is consistent with the variation that would be expected if the real system were more like M3 than M1. The pattern observed in the numerical experiments with M1 as the underlying model shows that the more complicated models can be fit just as well as M1, but that these fits result in redundant parameters. For example, in the fit of the cellulose M1 model results to the M3 model, a_L and a_{S1} have simply split the mass that was really in the L compartment, and k_L and k_{S1} have the same value, making the dynamics of these two compartments indistinguishable. This pattern of redundant parameters is not apparent in the models fitted to ^{14}C experimental data. On the other hand, for the numerical experiments using M3 as the underlying model, the variation in parameters with fitting to progressively less complex models results in a pattern that is very similar to the fitted

parameters for the ^{14}C experimental data. In the case of every parameter, the qualitative pattern in the ^{14}C results is observed also in the M3 numerical results, and, with a few exceptions, the quantitative information of the ^{14}C results is preserved as well.

Based on my analysis of the results summarized in Table 4, I selected M3, of the two tested models, as the more likely to represent the decomposition dynamics of these homogeneous substrates. The basis for selecting M3 over M1 was equally strong for all three substrates. It appeared that errors due to the assumption of simultaneous decomposition were manifested mostly in a_L and a_{S1}, since even though the underlying M3 model used the parameters derived from the ^{14}C data, the fitted parameters for the numerical experiments deviated from the input parameters for a_L and a_{S1}, but were very accurate for k_L and k_{S1}. Fitting to M3 always underestimated the value for a_L (average error -11.5%) and overestimated the value for a_{S1} (average error +4.4%). This result suggests that values for f_L and f_{S1} should be corrected, and lacking a more substantial theoretical basis, I made such corrections based on differences between parameters for the M3 numerical data fitted to the M3 model, and parameters derived from fitting the M3 model to ^{14}C experimental data under the assumption of instantaneous partitioning. The M3 model parameters for the three substrates after correction are shown in Figure 4.

Dashed lines in Figure 4 connect soil compartments between the models for glucose, cellulose, and lignin that have identical rate constants controlling the fluxes into them, as derived from completely independent data sources. Given that decomposition is principally an enzymatic process and that soil taxa are specialized to metabolize substrate of a certain chemistry on the basis of their enzyme production (Stevenson 1986), and given that successive generations of microbes recycling the same biomass will progressively lose quality as substrate materials for further biomass production as free energy is extracted for growth and maintenance (Hunt 1977), I propose that the compartments connected by dashed lines in Figure 4 are functional equivalents, representing sub-populations of the soil taxa with enzymatic production specialized to metabolize substrate of a certain chemical quality. This implies, for example, that the material resulting from a first generation of glucose metabolism is itself metabolized to form a second generation of microbial biomass and microbial byproducts which would at that point be indistinguishable from the primary products of cellulose metabolism. The same conclusions are drawn for the case of the

Homogeneous Substrate Decomposition Data Fitted to M3 Model

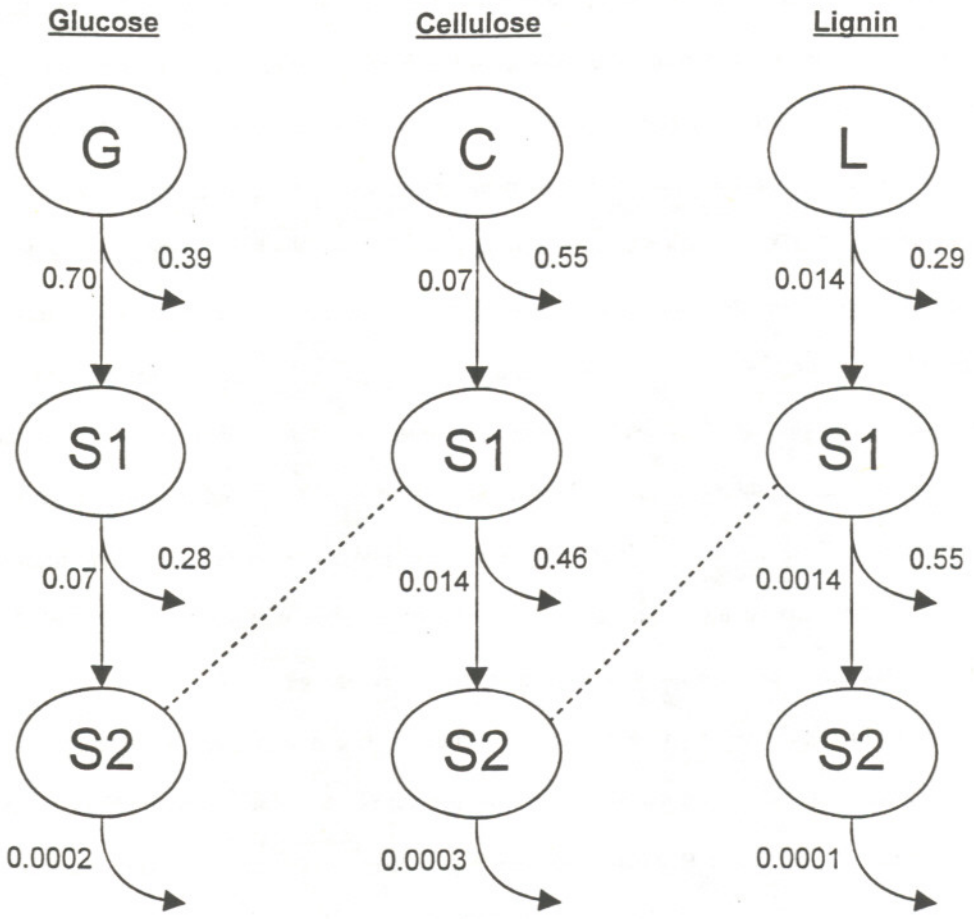


Figure 4

second generation of microbes and byproducts from cellulose metabolism and the first generation from the metabolism of lignin.

The final structure of the litter and soil decomposition model is shown in Figure 5. This structure reflects the arguments presented above, and stands as a testable hypothesis explaining decomposition dynamics of heterogeneous substrates as coupled decomposition pathways for three homogeneous components. In the generalization of the results from Tables 3 and 4, I have considered the glucose decomposition experiments to be representative of the labile, or water and alcohol soluble, litter fraction in general (Minderman 1967). I have also considered the S2 compartment from the M3 model fit to lignin decomposition data to represent the most recalcitrant fraction of soil organic matter labeled R in Figure 5. This treatment is suggested by Hunt (1977), who concluded from model comparisons with decomposition data that humic (recalcitrant) material accumulation resulted from the death and decay of soil microorganisms. The same conclusion was reached by Wieder and Lang (1982) after testing a number of different numerical formulations for litter decay against observations. There is debate over whether the processing of lignin by soil microbes, leading eventually to the accumulation of recalcitrant soil organic matter, proceeds mostly extracellularly or intracellularly (Martin *et al.* 1980). The low respiration fraction for lignin decomposition suggests that much of the lignin material proceeds through the first stage of processing without undergoing substantial metabolism. I avoid any explicit dependence on the localization of lignin metabolism as either internal or external to the microbes, since my treatment ignores the mechanisms of microbial metabolism and concentrates instead on inputs to and outputs from various stages of microbial processing. This treatment is consistent with the notion of soil microbial biomass acting as a transformation station, through which all litter material must pass (van Veen *et al.* 1984).

Based on the structure in Figure 5, I performed simulations of the decomposition of single homogeneous substrate components. The simulation results for each component are plotted in Figures 6a, 6b, and 6c, together with each component's pooled data and M3 regression curve. As I demonstrated through the numerical experiments in which data generated using M3 was subsequently fit to the M1 model, the rate from the final compartment in the series arrangement is underestimated, presumably because it is an amalgamated rate of progressively slower processes. By expanding the series of

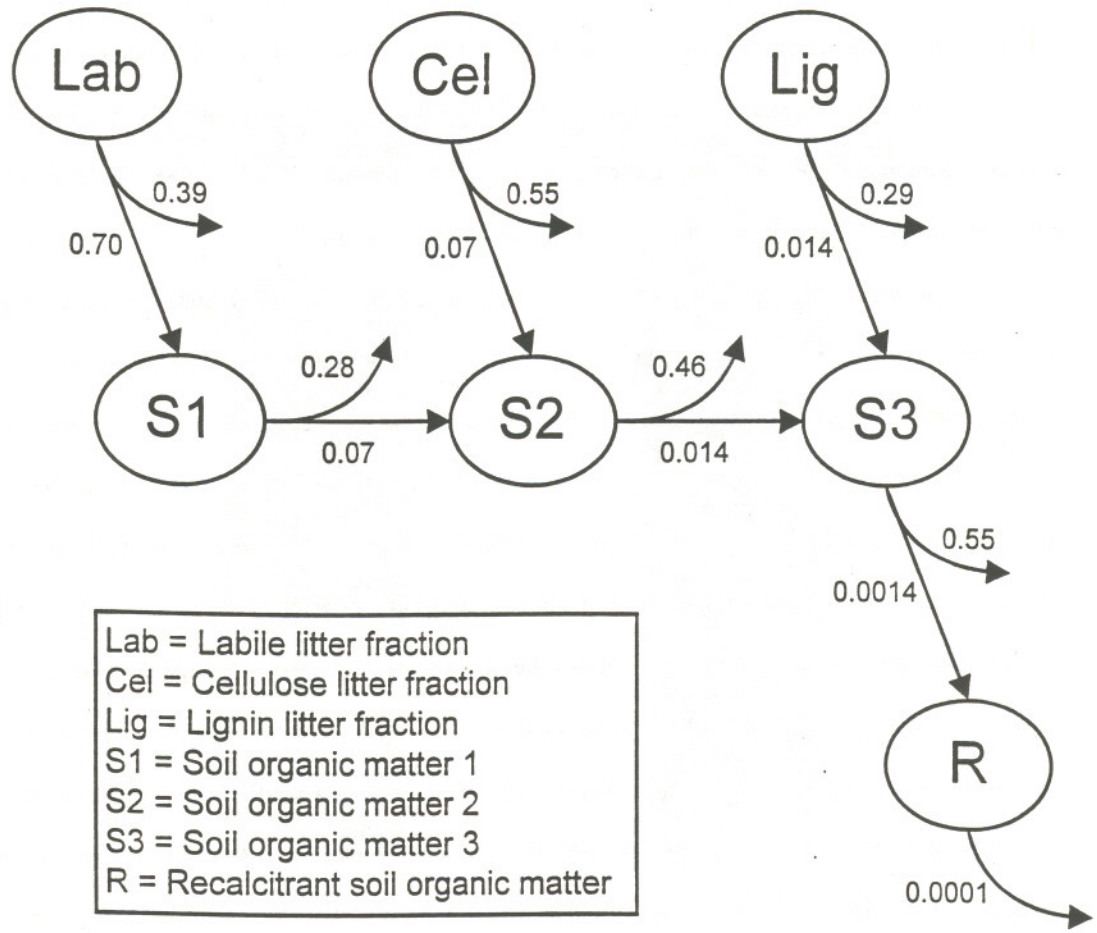

Litter and Soil Carbon Dynamics for 1D-BGC
Lab
Cel
Lig
0.39
0.70
0.55
0.07
0.29
0.014
S1
0.28
0.07
S2
0.46
0.014
S3
0.55
0.0014
R
0.0001
Lab = Labile litter fraction
Cel = Cellulose litter fraction
Lig = Lignin litter fraction
S1 = Soil organic matter 1
S2 = Soil organic matter 2
S3 = Soil organic matter 3
R = Recalcitrant soil organic matter

Figure 5

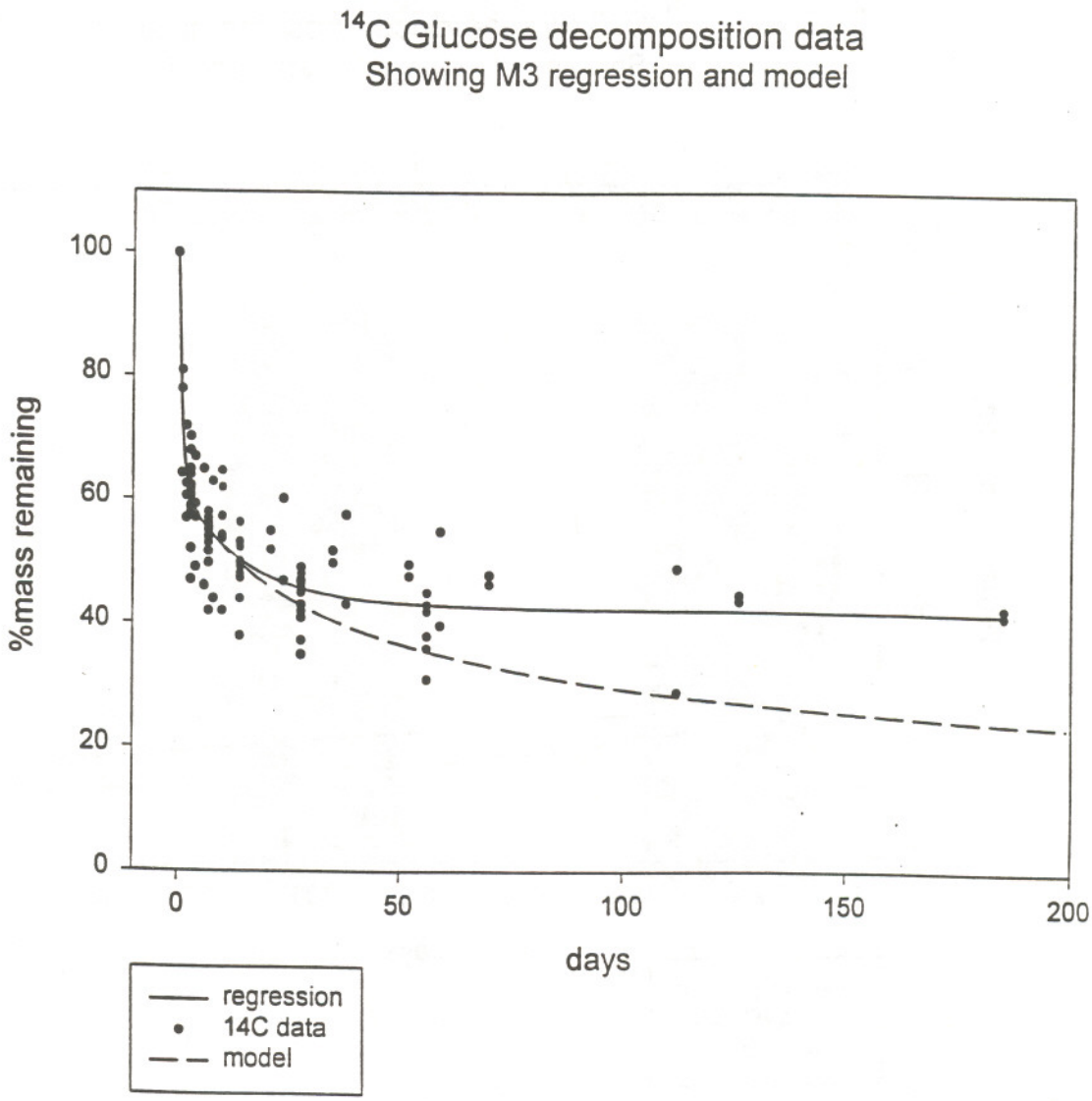

^{14}C Glucose decomposition data
Showing M3 regression and model
%mass remaining
100
80
60
40
20
0
0
50
100
150
200
days
regression
14C data
model

Figure 6a

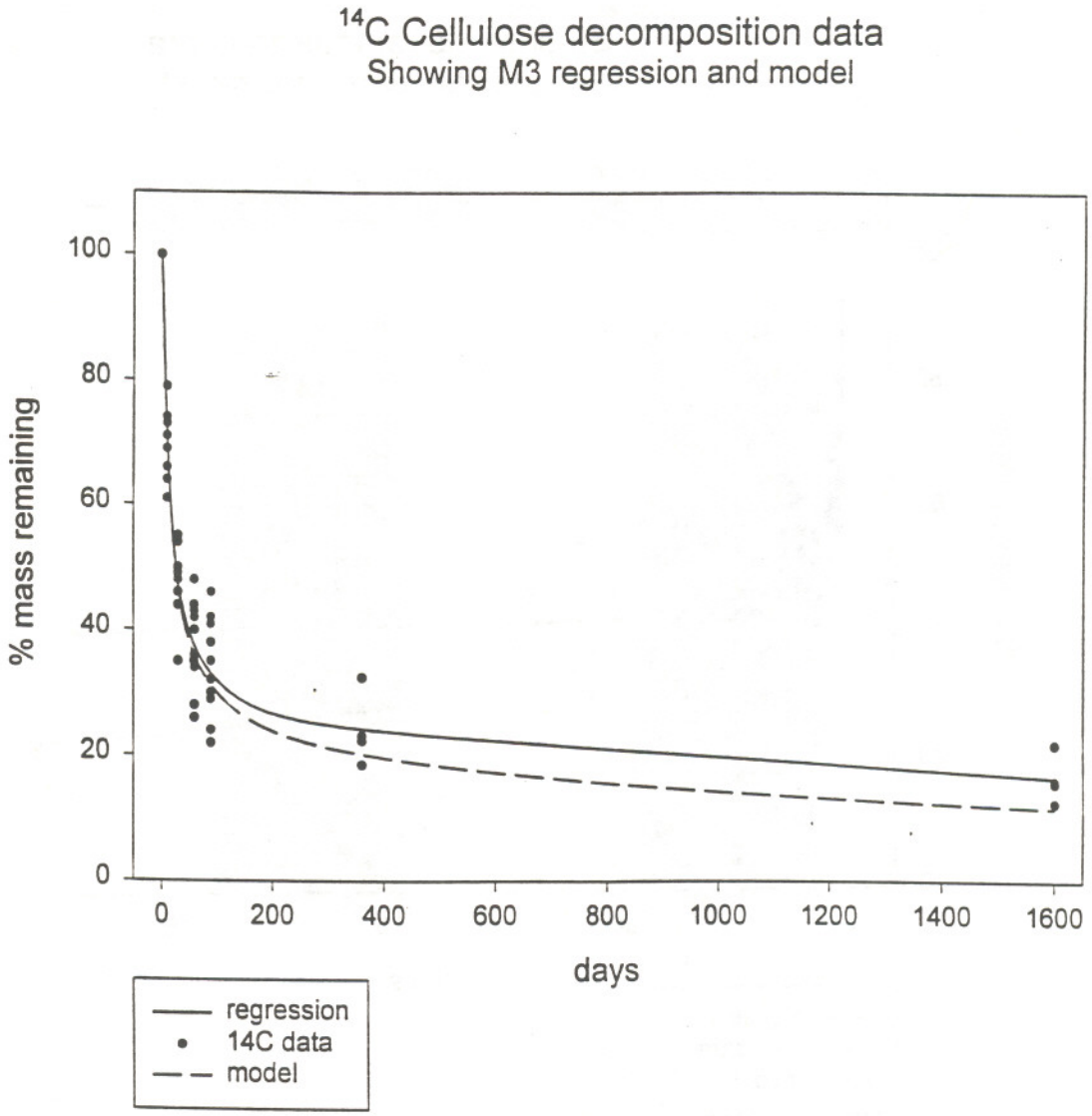

^{14}C Cellulose decomposition data
Showing M3 regression and model
% mass remaining
100
80
60
40
20
0
0
200
400
600
800
1000
1200
1400
1600
days
regression
14C data
model

Figure 6b

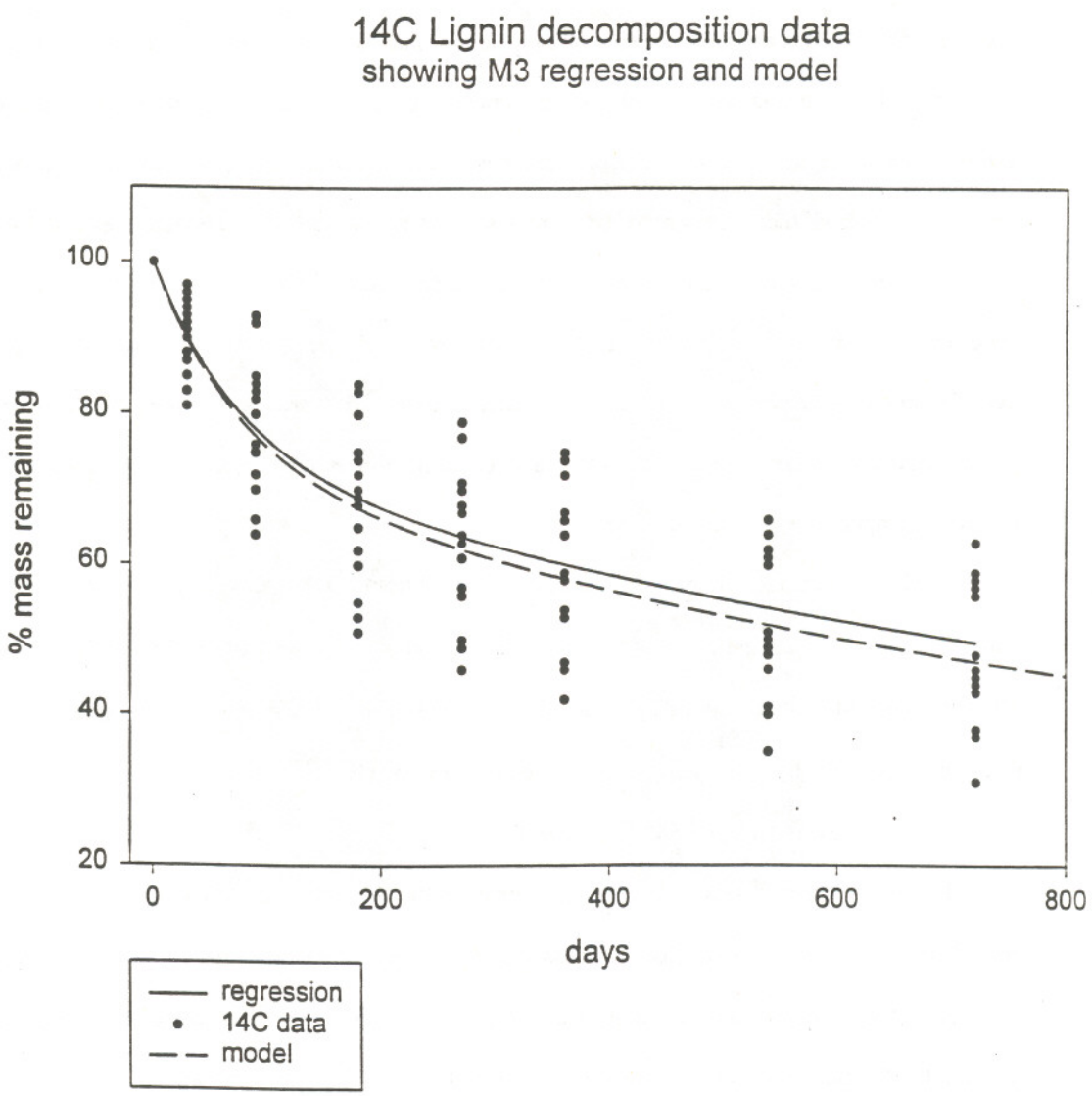

14C Lignin decomposition data
showing M3 regression and model
% mass remaining
100
80
60
40
20
0
200
400
600
800
days
regression
14C data
model

Figure 6c

compartments, these rates can be resolved. Given rates lower than those resolved for S3 in Figure 5, the resolution of compartments beyond R would require multiple-year sampling, but sampling periods longer than two years are rare. Other authors have resorted to radioactive dating methods to determine the half-lives for the slowest soil organic matter pools, yielding values between 200 and several thousand years (van Veen and Paul 1981, Parton *et al.* 1987). It may be essential for some purposes to include a representation of this very slowly decomposing pool, but a treatment of the problem is beyond the scope of this work. Ignoring the influence of suboptimal temperature and soil moisture on decomposition rates, the half-life for decomposition from the slowest pool in Figure 5 is about 20 years,. Cooler temperatures and dryer soils could easily push this half-life into the observed range for slow decomposition pools. Stevenson (1986) states that respiration losses from the recalcitrant soil organic matter pool are typically between 2 and 5% $year^{-1}$, with the higher values for warmer and wetter climates. This gives a daily rate of release of 0.0001 for the upper end of that range, which is the value determined from the decomposition experiments under favorable temperature and moisture conditions.

The rate constants shown in Figure 5 are derived from experiments carried out at uniform temperature between 20 and 30 °C with moist soils. It is known that temperature and moisture conditions affect decomposition rates. From Lloyd and Taylor (1994), I include a correction for the decomposition rate constants on the basis of temperature above or below 25 °C:

$$k_{corr} = k \exp(4.34 - 308.56/(T_{soil}+46.02))$$

where k is any of the rate constants, and T_{soil} is the estimated daily average soil temperature in the upper 10 cm (Zheng *et al.* 1993). Decomposition is assumed to be negligible below -10 °C. From studies of the effect of soil moisture on decomposition rates (Orchard and Cook 1983, Andrén and Paustian 1987), I derived the following rate constant correction for variation in soil water potential:

$$k_{corr} = k \ln\left(\frac{-10.0/\psi}{\ln(-10.0/\psi_{sat})}\right)$$

The availability of soil mineral nitrogen acts as a final control on decomposition rates. As litter material is metabolized to become soil microbial biomass and microbial byproducts, some carbon is lost through respiration, but nitrogen from the original material is retained. Nitrogen (as well as other nutrients)

is required for microbial biosynthesis and growth, and the ratio of carbon to nitrogen (C:N) in microbial tissue is relatively conservative, ranging from about 8:1 to 15:1 (Aber *et al.* 1990, Smolander *et al.* 1994). When the C:N of a microbial growth substrate is higher than about 12:1, after accounting for loss of C due to respiration, mineral nitrogen (ammonium and nitrate) in the soil may be incorporated in the new microbial biomass during substrate metabolism (immobilization) to make up the nitrogen deficit and bring the final microbial C:N down to around 12:1 (Stevenson 1986). For substrates with higher N concentration or for high respiration fractions, more N may be present in the metabolized substrate than is required for microbial growth, and the extra N is released to the soil in a mineral form (mineralization). Immobilization and mineralization occur simultaneously in soils having organic matter in varying stages of decay, and over some period of time the net result may be an increase in soil mineral N (net mineralization), or a decrease in soil mineral N (net immobilization).

The N requirements for the decomposition of a particular substrate depend on the substrate C:N, the respiration fraction, the microbial C:N, and on possible N losses during metabolism. It is commonly observed that low N availability with respect to N demand results in an increased respiration fraction for the same quality of substrate (Melin 1930, Waksman and Gerretsen 1931, Allison and Murphy 1962, Berg and Tamm 1994). Presumably this results from the additional respiration costs of active N transport, as opposed to passive diffusion which dominates under non-limiting N conditions (Taylor *et al.* 1991). Reduction of N limitation by low-level fertilization increases decomposition rates for whole-plant materials (Waksman and Gerretsen 1931, Berg and Tamm 1994) and for homogeneous cellulose and lignin fractions (Melin 1930). As shown by McClaugherty *et al.* (1985), additions of N under conditions for which the N immobilization demand is already met by the N mineralization rates do not result in increased decomposition rates.

Based on observations of reduced decomposition rates under N limitation, decomposition rates for the litter and soil compartments in 1D-BGC are limited by net mineralization rates in a two-step process. First, all decomposition processes are estimated with temperature and moisture corrected rates from Figure 5. Any flux that results in N mineralization is allowed to proceed at this daily potential rate, and the addition to soil mineral N from all such mineralizing fluxes is calculated. These fluxes are added to the soil

mineral N pool remaining from the previous day, and the total soil mineral N available for decomposition fluxes that result in N immobilization is determined as 90% of the mineral N pool. Next, if the estimated immobilization due to potential decomposition fluxes is less than this available soil mineral N, then these fluxes are allowed to operate at their daily potential rates, and mineral N is immobilized according to the initial estimates. If, on the other hand, available mineral N is less than required by the potential immobilizing fluxes, all the immobilizing fluxes are scaled uniformly to adjust the demand to exactly meet the supply, resulting in a maximum of 90% of the soil mineral N pool being immobilized by decomposition fluxes on each day.

After all mineralization and immobilization fluxes have been reconciled, any remaining soil mineral N is made available for uptake by plants. When the phenological signals are such that fine roots are present, the plant takes up 90% of the available soil mineral N. No N uptake occurs on days when fine roots are not present. In earlier BGC versions, plant N uptake was treated as a function of the fine root biomass and a maximum rate of N uptake per unit of root mass (Hunt and Running 1992). This approach is also used by Rastetter et al (1991) who make N uptake a function of total root N content. Compared to other processes in terrestrial ecosystems, belowground production and the division of total production between above and belowground components are poorly understood (Nadelhoffer *et al.* 1985, Raich and Nadelhoffer 1989, Holmes and Zak 1994, Zak *et al.* 1994). For this reason, even though it may be a realistic approach, I have rejected the connection between fine root mass or N concentration in the determination of N uptake. Another approach (McMurtrie 1991, Potter *et al.* 1993) is to specify a range for foliar nitrogen concentration and allow the plant to take up as much nitrogen as is available until the maximum foliar concentration is reached, after which surplus mineral N remains in the soil. I argued in Section 5.1 that leaf N concentration can remain constant if leaf area is assumed to be the dominant mechanism for plant response to N limitation. Potter *et al.* (1993) find that their model generally reaches its specified maximum foliar N concentration, and they report that variation in leaf area rather than variation in the photosynthetic capacity of leaves is the dominant mode of variation within a lifeform class for their global simulations.

In my treatment of plant N uptake, I need to accommodate both daily and annual carbon and nitrogen allocation models (see Section 5.7). The implication for N uptake is that part of the N demand is unknown until the end of the growth season. The first part of the solution to this problem is to allow "luxury" N uptake, as just discussed. Some of this N is used in new growth specified by the daily carbon and nitrogen allocation model, part of it is used for woody tissue growth at the end of the growing season, and part of it is stored to initiate new growth of leaves and fine roots in the following growth season. After all these demands are met, there may be additional N that was taken up but not allocated. My solution to this problem is to release this mineral N back to the soil mineral N pool in equal increments over the course of the next year. The result of luxury N uptake, daily and annual allocation, and slow return of unused N from the plant to the mineral soil pool is that net plant N uptake resembles N demand without having to diagnose the true N demand for annual allocation in advance.

Figure 7 illustrates the dominant pools and processes and the linkages between C and N in the plant-litter-soil system. For purposes of clarity, only one generic pathway through the litter and soil system is included in this diagram. Note that although litter substrate decomposition is shown to immobilize N and that decomposition from the first soil organic matter compartment is shown to mineralize N, both immobilization and mineralization are possible at both these points, depending on the C:N of litter substrate, the specified respiration fraction, and the specified C:N of the soil organic matter pools.

In addition to the active labile, cellulose, and lignin pools illustrated in Figure 5, there is a fourth litter pool which represents cellulose that is not available for decomposition as the result of it being associated with lignin which is shielding it from the microbial community (Donelly *et al.* 1990, Taylor *et al.* 1991). Data from Taylor *et al.* (1991) show that decomposition rates for substrates containing both cellulose and lignin decrease as lignin content increases from 10% to 28% of total litter mass, but that for larger fractions of lignin the decomposition rate is relatively constant. This implies that lignin completely dominates the decomposition dynamics at a lignin content of about 30%. Given a typical range for total cellulose content of non-woody litters between 36 and 50% (Berg *et al.* 1984, Berg and McClaugherty 1989, Stump and Binkley 1993), and assuming an average value of about 40%, I estimate the maximum level of lignin shielding of cellulose to occur at a ratio of lignin to cellulose of 0.7. For ratios greater than

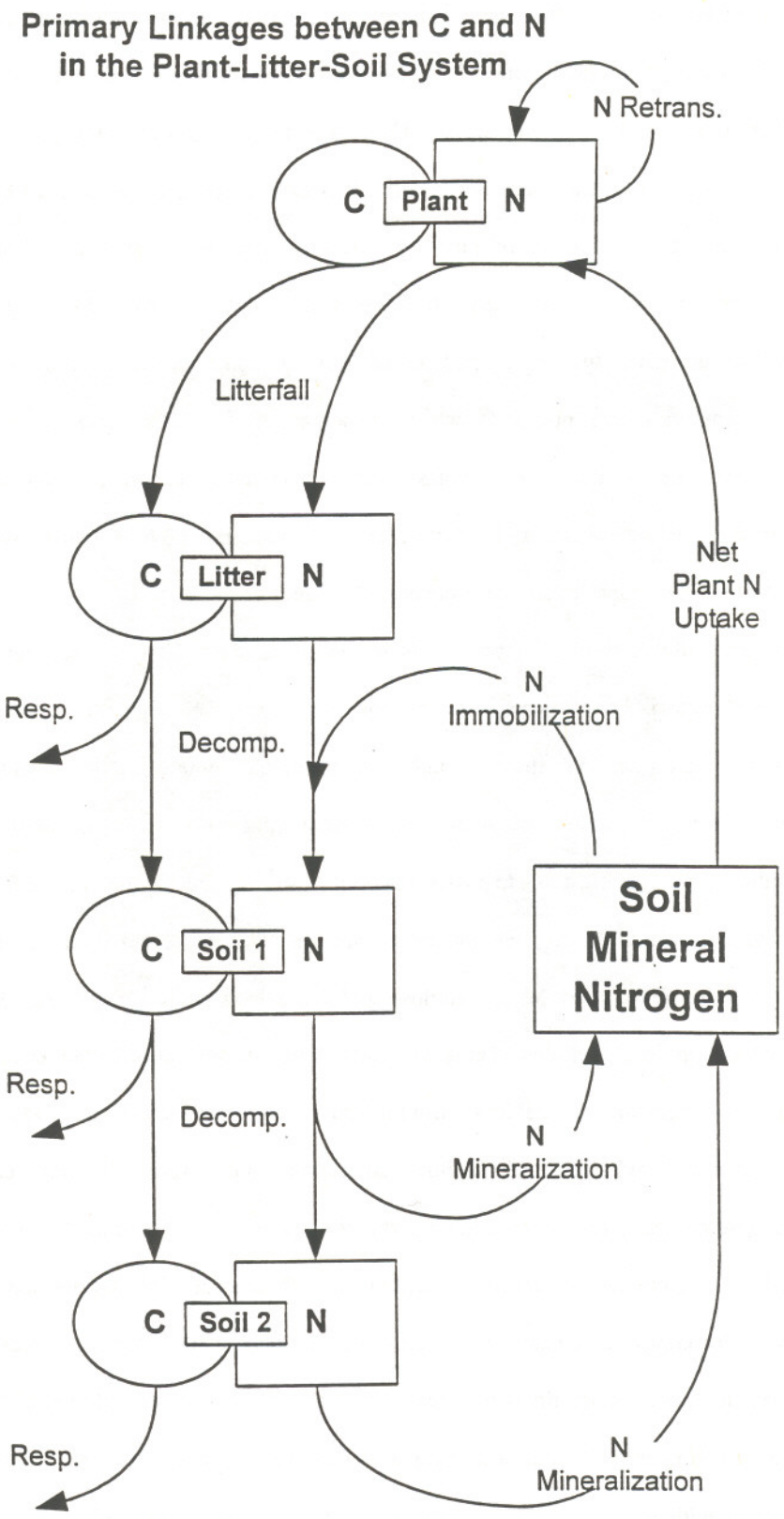

Primary Linkages between C and N
in the Plant-Litter-Soil System
N Retrans.
C
Plant
N
Litterfall
C
Litter
N
Net
Plant N
Uptake
Resp.
Decomp.
N
Immobilization
Soil
Mineral
Nitrogen
C
Soil 1
N
Resp.
Decomp.
N
Mineralization
C
Soil 2
N
Resp.
N
Mineralization

Figure 7

or equal to 0.7, 80% of the litter cellulose occurs in the shielded fraction. For ratios below 0.7 and greater than 0.25, the amount of cellulose in the shielded fraction drops from 80% to 0%, with no cellulose shielding for ratios lower than 0.25. The shielded cellulose pool becomes available for active decomposition, according to the ratio between shielded cellulose and total lignin, as the lignin litter pool is metabolized. Cellulose released from shielding simply enters the homogeneous cellulose litter pool and is then treated as part of the total active cellulose pool.

The final component of the plant-litter-soil model is a storage compartment for coarse woody debris, defined here to mean the cellulose and lignin fractions of the live and dead stem and coarse root compartments that are being lost from the plant compartments due to whole-plant mortality (Section 5.5). Coarse woody debris is included as a separate compartment because it must undergo a process of physical fragmentation before decomposition can begin (Harmon *et al.* 1986). Grier (1978) found a fragmentation rate constant of 0.015 yr^{-1} for Hemlock logs in Oregon, under conditions of high precipitation and mild temperature. Sollins (1982) obtained similar results for bole fragmentation in western Oregon and Washington, with an average fragmentation rate constant of 0.018 yr^{-1}. The dependence of fragmentation on variation in moisture and temperature is poorly understood (Harmon *et al.* 1986). I have assumed the same dependence on temperature and moisture as for the metabolic decomposition rates, and have furthermore assumed that the rates obtained in the Oregon and Washington studies are on the medium to fast end of the fragmentation gradient due to climate. Therefore I assigned a base fragmentation rate of 0.0001 day^{-1} for all coarse woody debris. Lignin and cellulose are not discriminated within the coarse woody debris compartment, but their relative proportions are maintained, with the result that material leaving that compartment due to fragmentation fluxes is partitioned between the cellulose and lignin litter pools on the basis of average coarse woody debris cellulose and lignin contents. Cellulose leaving the coarse woody debris compartment is assumed to be shielded or not on the basis of the lignin to cellulose ratio for the entire compartment.

Atmospheric N deposition occurs at a specified daily rate, with all atmospheric inputs entering the soil mineral N pool. N leaching occurring in conjunction with soil water outflow is the only N removal pathway. The rate of leaching loss is calculated by assuming that during outflow the water soluble fraction

of the soil mineral N pool is being transported out of the system according to its concentration in the soil water. I assume that 10% of the soil mineral N pool is water soluble, and calculate the N leaching flux as the soluble mineral N pool multiplied by daily outflow and divided by soil water content. Volatilization fluxes from the soil mineral N pool to the atmosphere are not considered at present. A very simple treatment of the volatilization fluxes is proposed by Parton *et al.* (1993), who estimate volatilization as 1% of gross N mineralization. This treatment could easily be incorporated into 1D-BGC if it were deemed necessary.

The plant-litter-soil model in 1D-BGC is fully defined once the user has supplied the following parameters: the C:N of live foliage(CN_{leaf}), the C:N of foliage litter after retranslocation has occurred ($CN_{leaflitter}$), the C:N of fine roots (CN_{froot}), the C:N of live wood (CN_{lwood}), the labile proportion of leaf litter and fine root litter ($fLab_{leaf}$, $fLab_{froot}$), the cellulose proportion of leaf litter, fine root litter, and dead wood ($fCel_{leaf}$, $fCel_{froot}$,$fCel_{dwood}$), and the lignin proportion of leaf litter, fine root litter, and dead wood ($fLig_{leaf}$, $fLig_{froot}$, $fLig_{dwood}$).

5.7 Carbon and nitrogen allocation

On a daily basis, carbon is added to an abstract storage pool (C_{pool}) as the result of photosynthesis, and is subtracted from the same pool as the result of maintenance and growth respiration. Similarly, nitrogen is accumulated in an abstract storage pool (N_{pool}) as the result of daily plant N uptake. This section treats the simulation of new growth resulting from the allocation of C and N from C_{pool} and N_{pool} to other plant compartments. This allocation can take place by two very different mechanisms; both can operate within the same simulation. The first mechanism, which I will refer to as annual allocation, occurs once per year, on the final day of litterfall (see Section 5.4). The second mechanism, referred to as daily allocation, can take place on any day between the first day of the new leaf growth period in the spring until the first day of the litterfall period. Of the two, annual allocation is more involved and more sensitive to seasonal integrals of growth conditions. I first provide outlines for annual and daily allocation, and then present more detailed descriptions

Annual allocation is a major point of reconciliation between the plant carbon and nitrogen cycles, as well as an opportunity for the plant to respond to various signals from the growth environment which are

not easily integrated in the more mechanistic treatments of daily biophysical and biochemical processes. The fundamental assumption in the annual allocation mechanism is that all distributions of available carbon and nitrogen between different plant compartments can be derived allometrically having specified only the allocation to new leaf growth. Other ecosystem process models have used a similar logic (e.g. McMurtrie 1991, Rastetter *et al.* 1991, Running and Gower 1991, Hunt and Running 1992, Potter *et al.* 1993). All plant growth controls can then be estimated through their impact on new leaf growth. This is logically consistent with the hierarchy of growth precedence suggested by Waring and Pitman (1985) and Waring and Schlesinger (1985), giving the highest precedence to new leaf growth and bud formation (preparation for next growth season's new leaf growth). Four resource availability factors control new leaf growth: the availability of carbon, the availability of nitrogen, the integrated seasonal water stress, and the availability of unintercepted light. A secondary limitation is established by a user-prescribed value that sets an upper limit on the percent difference between years in the carbon allocated to new leaf growth. These potential limitations are considered together to determine the dominant control on new leaf growth, after which all of the mandatory allometric relationships are applied to determine the growth for fine roots and the live components of stem and coarse root. It is possible that carbon remains after all mandatory allocation is complete, and for trees this remaining carbon is allocated to new growth of stem and coarse root wood, which is assumed to enter the dead wood compartments (i.e., new xylem tissue). For nonwoody plants, this excess carbon is sent to the labile litter compartment over the course of the next growing season, emulating the production of root exudate. Excess nitrogen, for both woody and nonwoody plants, is returned to the soil over the course of the next year, as described above (Section 5.6).

The daily allocation mechanism is optional, and in that it uses carbon and nitrogen from the C_{pool} and N_{pool}, it competes with the annual allocation mechanism. The user specifies the degree of dominance between the daily and annual allocation by setting a linear function that determines, on each day between the beginning of new leaf growth and the beginning of litterfall, the proportions of daily net photosynthesis and plant N uptake available for daily allocation. The remaining proportions are retained in the C_{pool} and N_{pool} and are available at the end of the growing season for the annual allocation mechanism. Daily allocation is much simpler than annual allocation, in that only controls due to carbon and nitrogen

availability are considered, without any explicit treatment of water stress or light limitation. Because all influences on resipiration, photosynthesis, and decomposition are manifest in the daily estimates of net photosynthesis and N uptake, factors such as soil water depletion, high VPD, N mineralization rates, or canopy light interception are acting on the daily allocation mechanism implicitly through many other model processes. Given a fraction of the daily additions to the C_{pool} and N_{pool}, new daily leaf growth is estimated as the dominant limitation, and allocation to all other plant compartments is based on the same mandatory growth allometry used in the annual mechanism. There is no explicit treatment of excess C or N for daily allocation, since it is simply assumed to remain in the storage pools until the end of the growth season, at which point it becomes available to the annual allocation mechanism.

A fundamental difference between the annual and daily allocation mechanisms is that the annual allocation to new leaf growth is made at the end of the growth season, and enters a temporary storage pool that will be moved into actual displayed leaf material only at the beginning of the next growth season (controlled by the phenological model, Section 5.4), while new leaf growth allocated by the daily mechanism is immediately displayed. Thus, the annual allocation mechanism is responsible for the establishment of a plant architecture that is balanced with respect to long-term patterns in the growth environment, while the daily allocation mechanism allows for a more immediate response of the vegetation to the current year's growth conditions.

For the annual allocation mechanism, the user specifies the maximum percent change between years in the amount of carbon allocated to new leaves (lgf_{max}). This sets an upper limit on the amount of new leaf growth when all other resources are adequate, and is intended to restrain growth rates in accordance with the requirements for structural growth to support new leaf tissue. The maximum annual amount of carbon allocated to new leaf growth ($MAX_{newleafC}$) is then given as:

$$MAX_{newleafC} = PREV_{newleafC} (1.0 + lgf_{max})$$

where $PREV_{newleafC}$ is the amount of new leaf carbon allocated in the previous growth season. The parameter lgf_{max} also sets the maximum amount by which new leaf growth can be reduced from the previous year due to water stress in the current year, as discussed below.

The allometric constants defining the relationship of new leaf growth to mandatory new growth in the other plant compartments are specified by the user. They are: the ratio of new fine root growth to new leaf growth (f_1); the ratio of new live coarse root growth to new live stem growth (f_2); the ratio of new live stem growth to new leaf growth (f_3); the proportion of C_{pool} to retain for the next growing season (f_4); and the proportion of N_{pool} to retain for the next growing season (f_5). Other relevant parameters include the growth respiration constants for nonwoody and woody components (G_n and G_w, see Section 5.3), and the C:N ratios for live leaf, fine root, and live wood (CN_{leaf}, CN_{froot}, and CN_{lwood}, see Section 5.6). Based on these parameters, the estimates of annual allocation of carbon to new leaf growth as controlled by C and N availability ($C_{newleafC}$ and $N_{newleafC}$, respectively) are defined for woody plants as:

$$C_{newleafC} = C_{pool}\,(1.0-f_4) / (1.0+G_n+f_1+f_1G_n+f_3+f_3G_w+f_2f_3+f_2f_3G_w)$$

$$N_{newleafC} = N_{pool}\,(1.0-f_5) / (1.0/CN_{leaf} + f_1/CN_{froot} + (f_3+f_2f_3)/CN_{lwood})$$

and for nonwoody plants as:

$$C_{newleafC} = C_{pool}\,(1.0-f_4) / (1.0+G_n+f_1+f_1G_n)$$

$$N_{newleafC} = N_{pool}\,(1.0-f_5) / (1.0/CN_{leaf} + f_1/CN_{froot})$$

The availability of water is known to limit leaf growth (e.g., Grier and Running 1977), and I include a mechanism that makes the annual allocation to new leaf growth sensitive to an index of integrated water stress over the growing season. This mechanism is designed to maximize long-term productivity under constant climatic conditions by reducing new leaf growth in years when complete stomatal closure due to low soil water potential is reached. I have found, for a constant climate simulation (i.e. a multiple year simulation which uses a single year of meteorological data repeatedly), that production is maximized when leaf area is just sufficient to bring the soil water content to the point at which complete stomatal closure occurs on one day during the growing season. Production falls off for leaf area resulting in more than one day of complete stomatal closure due to water limitation. On this basis, I define an annual integrated water stress index as the number of days with complete stomatal closure due to water limitation (n_{dry}), and when n_{dry} is greater than one day a control on new leaf growth due to water limitation ($W_{newleafC}$) is implemented, defined as:

$$W_{newleafC} = PREV_{newleafC}\,(1.0 - (lgf_{max}\,(n_{dry}-1)/45)) \qquad \text{for } n_{dry} <= 45 \text{ days, and}$$

$$W_{newleafC} = PREV_{newleafC}\ (1.0 - lgf_{max}) \qquad \text{for } n_{dry} > 45 \text{ days}$$

New leaf growth is also controlled by the availability of radiation, such that new growth is inhibited when the fractional absorption of incident PAR approaches 100% (Larcher 1995). I define a linear control between 80 and 100% fractional PAR absorption ($R_{PAR,C}/R_{PAR}$, see Section 3.2), and set a control on allocation of carbon to new leaf growth due to light limitation ($L_{newleafC}$) as:

$$L_{newleafC} = P_{newleafC}\ [1.0 + lgf_{max}(1.0 - ((R_{PAR,C}/R_{PAR})-0.8)/0.2)] \qquad \text{for } R_{PAR,C}/R_{PAR} > 0.8$$

The dominant control on carbon allocation to new leaf growth ($POT_{newleafC}$) is determined as the minimum of five potential controls: $MAX_{newleafC}$, $C_{newleafC}$, $N_{newleafC}$, $W_{newleafC}$, and $L_{newleafC}$. Because new leaf growth controls due to water and nitrogen limitations given above are implemented only when they are the most limiting control, these controls by themselves can lead to abrupt growth declines when water or nitrogen resource availability is changing rapidly. For example, in the case of an evergreen canopy which is accumulating nitrogen in its leaves as canopy leaf area increases, the onset of nitrogen limitation may occur abruptly, causing large differences between growth in two successive years. It seems plausible that plants could have advance warning of such approaching limitations, resulting in reduced growth rates before a particular resource limitation becomes dominant. Regardless of the existence of such a mechanism in real plants, I find that the implementation of such an preemptive growth limitation in 1D-BGC helps to reduce interannual oscillations of leaf area in the vicinity of a point of co-limitation by multiple resources. I implement two such controls, causing a restriction of new leaf growth for either an approaching water limitation or an approaching nitrogen limitation. These controls are implemented only in years when $POT_{newleafC} > PREV_{newleafC}$. The secondary water control can only be implemented when water is not the primary limitation, and likewise for the secondary nitrogen control. The secondary water and nitrogen controls are implemented on the basis of water and nitrogen availability indices (W_{AI} and N_{AI}, respectively), which are unitless scalars, defined as:

$$W_{AI} = \min\{1.0,\ 2(\theta_{min} - \theta_{crit})/(\theta_{sat} - \theta_{crit})\}$$

$$N_{AI} = \min\{1.0,\ (N_{newleafC}/POT_{newleafC}) - 1.0\}$$

where θ_{min} is the growing season minimum value of the volumetric water content, θ_{crit} is the volumetric water content at the soil water potential specified to cause complete stomatal closure, θ_{sat} is the saturated volumetric water content (Section 4.1), and min{a,b} indicates "the lesser of a and b".

Only one of these controls can operate in a given year, and in order for either to be operative, its value must be less than 1.0. When both W_{AI} and N_{AI} are less than 1.0, the lesser of them is selected as the secondary growth control, and the actual carbon allocation for new leaf growth ($ACT_{newleafC}$) is calculated as:

$$ACT_{newleafC} = PREV_{newleafC} + \min\{W_{AI}, N_{AI}\}(POT_{newleafC} - PREV_{newleafC})$$

Otherwise, for the case in which a secondary growth limitation is not invoked:

$$ACT_{newleafC} = POT_{newleafC}$$

The effect of the secondary water limitation is to begin to reduce new leaf growth from its potential rate when half of the available soil water is depleted at any time during the growing season, and to stabilize the annual new leaf growth at its value from the previous year when all the available soil water is used at any time during the growth season. The effect of the secondary nitrogen limitation is to begin a reduction on new growth rate when the primary nitrogen limit to new leaf growth is twice as large as the dominant primary limitation, ending in no additional increase in new growth from the previous year just before N limitation becomes dominant.

With the final prediction of carbon allocation to new leaf growth, the other mandatory growth fluxes of carbon and nitrogen are given as:

C allocation to new leaf C storage = $ACT_{newleafC}$

C allocation to new fine root C storage = $f_1\ ACT_{newleafC}$

C allocation to new live stem C = $f_3\ ACT_{newleafC}$

C allocation to new coarse root C = $f_2\ f_3\ ACT_{newleafC}$

N allocation to new leaf N storage = $ACT_{newleafC} / CN_{leaf}$

N allocation to new fine root N storage = $f_1 ACT_{newleafC} / CN_{froot}$

N allocation to new live stem N = $f_3\ ACT_{newleafC} / CN_{lwood}$

N allocation to new coarse root N = $f_2\ f_3\ ACT_{newleafC} / CN_{lwood}$

where woody allocation is only applied for woody plants.

For the case of woody plants, any carbon remaining in the C_{pool} in excess of the prescribed retention fraction (f_4) is used to produce new dead woody tissue. This mechanism is intended to represent the luxury growth of xylem tissue in woody plants in years when growth conditions are favorable. My purpose in distinguishing between live and dead woody components is to be able specify as mandatory growth the "thin sheathing lateral meristem" of phloem which is alive, respiring, and an integral part of the carbohydrate transport system between leaves and other plant parts (Kramer and Kozlowski 1979), while allowing increased stem wood production under favorable growth conditions to enter a dead stem pool representative of xylem tissue. This logic is consistent with observations of a decrease in the xylem to phloem mass ratio under stressed conditions (Kramer and Kozlowski 1979), and with studies indicating that this decreasing ratio is the result of reduced xylem production as opposed to increased phloem production (Bannan 1955, Evert 1960). Nitrogen accompanies the luxury growth of new dead wood according to the specified ratio of lignin to cellulose in dead wood and constant values for lignin and cellulose C:N ratios, specified as $CN_{lignin} = 500$ and $CN_{cellulose} = 250$, estimated from data on wood lignin and cellulose content and nitrogen concentration (Berg *et al.* 1984, Berg and McClaugherty 1989, Aber *et al.* 1990, Donelly *et al.* 1990, Taylor *et al.* 1991, Stump and Binkley 1993). If not enough N is available to meet these fixed C:N ratios, wood production occurs at lower values of C:N, using the same amount of C and all available N.

My treatment of mandatory new live stem and new live coarse root growth with respect to new leaf growth (parameters f_2 and f_3, above) requires some comment. Earlier BGC versions treated live and dead woody material as a single component, one for stem and another for coarse root material. This treatment presents a problem for the estimation of maintenance respiration, since it is only the live cells which respire, and a large fraction of the stem and coarse root material is dead. With the exception of axial and ray parenchyma cells which extend into the sapwood, all the living woody material is in the cambium and the phloem. Differences in nitrogen content between the phloem (~2% N), the sapwood (~0.3% N), and the heartwood (~0.1% N) reflect the differences in live material between these tissues (Kramer and Kozlowski 1979). I hypothesize that respiration from cambial and phloem tissue dominates the total woody tissue maintenance respiration budget, and that live phloem is required as a transport pathway

between fine roots and leaves. Based on a logic analogous to that presented by Waring and Schlesinger (1985) for the linear relationship between conducting xylem cross-sectional area and leaf area, I propose that a linear relationship should exist between phloem cross-sectional area and leaf area. The total phloem mass, required in estimating maintenance respiration and mandatory live stem and live coarse root growth, is then a function of the leaf area and the average stand height. By this logic two stands with equal leaf area but unequal average height will have unequal phloem respiration costs, with the ratio of their respiration costs varying as the ratio of their average heights. I have made the additional assumption that live woody material turns over at a rate equal to the leaf turnover rate.

This treatment depends on an estimate of average stand height. From tables of growth, yield, and stand volume for various conifer forest types in the western United States and Alaska, I estimated the total stem C and average stand height for stands of different site index at various ages. In my calculations I assumed a constant green wood density of 400 kg m^{-3}, a constant water fraction of green wood weight of 0.5, a constant carbon fraction of dry weight of 0.5, and a conical growth form, with the reported diameters at breast height taken as the cone base diameter. Results of these calculations using tables for unmanaged stands of douglas fir (McArdle and Meyer 1930), western white pine (Haig 1932), Alaskan white spruce, and ponderosa pine (Meyer 1938) are shown in Figure 8. Douglas fir and western white pine types are grouped together for a single regression, and ponderosa pine and white spruce are grouped for another regression, predicting height as a function of total stem C using a power relationship. The regressions shown in Figure 8 are highly significant, with standard errors of estimation less than 4% for all parameters. These results suggest that similar relationships could be derived for other woody vegetation types. Regressions based on tabular data for managed stands were also significant, but had higher values for both regression parameters and lower values of R^2 (data not shown).

In non-woody vegetation, any excess carbon is assumed to be sent belowground, without an associated nitrogen flux. This belowground carbon return enters the labile litter pool, and like the luxury N return discussed in Section 5.6, this carbon return occurs in equal amounts on each day of the following year. The return of excess carbon is more a matter of maintaining a plant mass balance than of representing any particular ecosystem process. A modeled surplus of photosynthate which is not needed in new leaf or

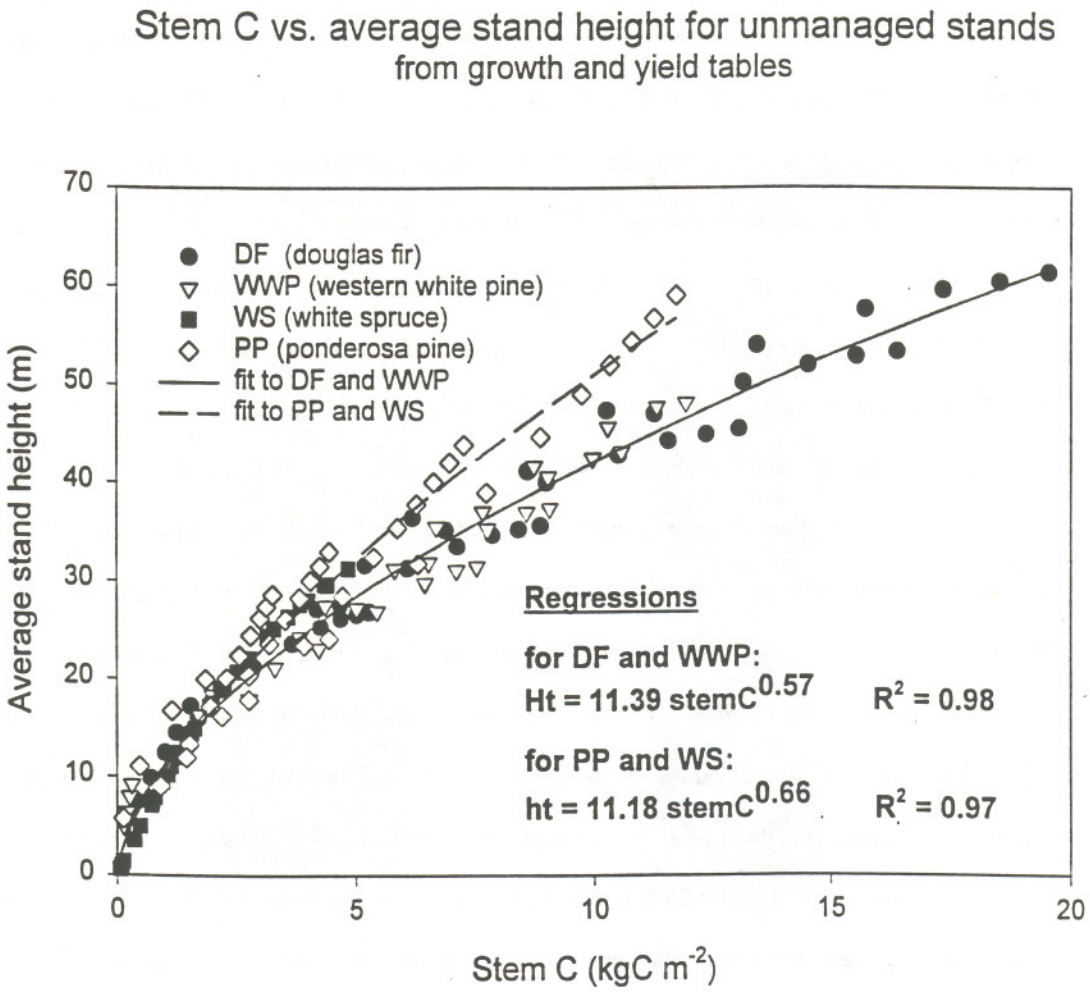

Stem C vs. average stand height for unmanaged stands
from growth and yield tables
Average stand height (m)
Stem C (kgC m-2)
DF (douglas fir)
WWP (western white pine)
WS (white spruce)
PP (ponderosa pine)
fit to DF and WWP
fit to PP and WS
Regressions
for DF and WWP:
Ht = 11.39 stemC0.57
R2 = 0.98
for PP and WS:
ht = 11.18 stemC0.66
R2 = 0.97
0
10
20
30
40
50
60
70
0
5
10
15
20

Figure 8

root growth can be allowed to accumulate in the C_{pool} or it can be sent to another compartment. Exudation of carbon-rich solution through the roots of non-woody plants is well documented (e.g., Schwenke and Wagner 1992, Hoffland *et al.* 1992, Xu and Juma 1994), and careful study with ^{14}C labeling has shown that up to 10% of annual NPP in wild grasses is released as organic root exudate, the respiration of which makes up greater than 30% of the total rhizospheric respiration in some cases (Johansson 1992, Cheng *et al.* 1993). These studies do not relate root exudation directly to photosynthate surplus, but they do indicate that exudation is common and can represent a large fraction of total production.

As discussed in Section 5.6, any N remaining in the N_{pool} after annual allocation is complete, in excess of the specified retention fraction (f_5), is returned to the soil at a constant rate over the course of the next year.

Annual allocation of C and N for new leaf and new fine root growth is made to storage compartments that are inactive until the first day of new leaf growth at the beginning of the next growing season. At this point, under the control of the phenological signals discussed in Section 5.4, C and N are moved from storage compartments into active leaf and fine root compartments. Maintenance respiration costs are only incurred once leaves and fine roots are displayed. One important implication of this storage in the fall for display in the spring is that the benefits of a good growing season in the current year are not realized as additional leaf growth until the following year. This sort of storage is important for the development of a positive photosynthetic balance in deciduous species during early new leaf growth (Kramer and Kozlowski 1979). Although net production is quite responsive to current growth conditions, the annual allocation mechanism does not allow for additional leaf growth during a favorable growth season. The purpose of the daily allocation mechanism is to allow an option for immediate response in growth to current conditions.

In comparison to the annual allocation mechanism, the treatment of daily allocation is very simple. Using the same formulations for the C limitation and N limitation to new leaf growth given above for annual allocation, but replacing the C_{pool} and N_{pool} with the fraction of daily net C fixation and daily plant N uptake available for daily allocation, daily C and N limits to new growth are calculated. The lesser of these

is the actual new daily growth, and all the mandatory growth fluxes discussed above are applied. There is no treatment of luxury C or N in the daily allocation mechanism.

6. Example simulations

I have concentrated my discussion so far on the logic and reasoning behind the treatment of individual processes in the 1D-BGC model, and have presented justification for my decisions at what I perceive as the most crucial junctions of these processes in the model structure. In order to expand this discussion to the integrative functioning of the whole model, I present here an analysis of some example simulations. The purpose of this analysis is to provide an overview of the integrated model behavior, and not to present a rigorous validation of any particular simulation component. I will, however, present various lines of evidence arguing that the overall model performance is not unreasonable, and will comment on the integrative aspects of the model which suggest possible future improvements.

The first example consists of two 2000-year simulations begun with very low L_P and no organic matter in the litter or soil compartments. The only source of N is from atmospheric deposition (0.4 gN m^{-2} yr^{-1}), and I refer to this scenario as a case of severe N-limitation, because in the early part of the simulations N-mineralization is very low due to lack of soil organic matter. The simulations are performed using one repeated year (1984) of meteorological data from Missoula, MT, and are intended to represent the establishment of an evergreen canopy starting from very low leaf area ($L_P = 0.1$). The sole difference between the two simulations is that one uses both annual and daily allocation mechanisms (referred to as dual allocation, and indicated by solid lines in Figures 9, 10, and 11), and the other uses only the annual mechanism (dashed lines).

Leaf C on yearday 250 (prior to litterfall) for the first 200 years of the simulations is shown in Figure 9a, illustrating that dual allocation and annual allocation both follow the same initial trend in canopy development, but that dual allocation takes a stable trajectory while annual allocation exhibits damped oscillations with a frequency of around 25 years. These oscillations are the result of fluctuating soil mineral N levels, which are exaggerated by the one-year time lag in the response of new leaf growth to N

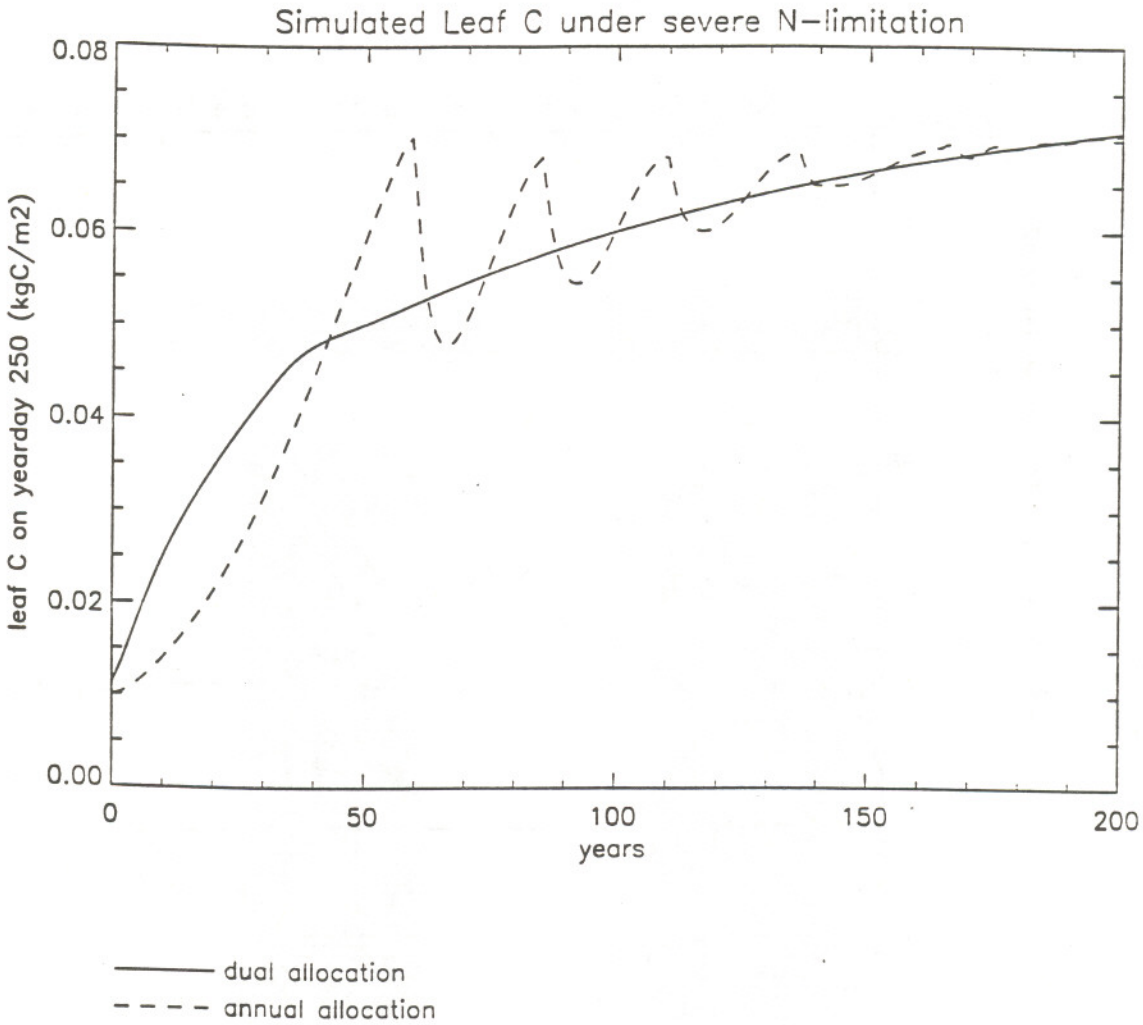

Simulated Leaf C under severe N-limitation
leaf C on yearday 250 (kgC/m2)
0.08
0.06
0.04
0.02
0.00
0
50
100
150
200
years
dual allocation
annual allocation

Figure 9a

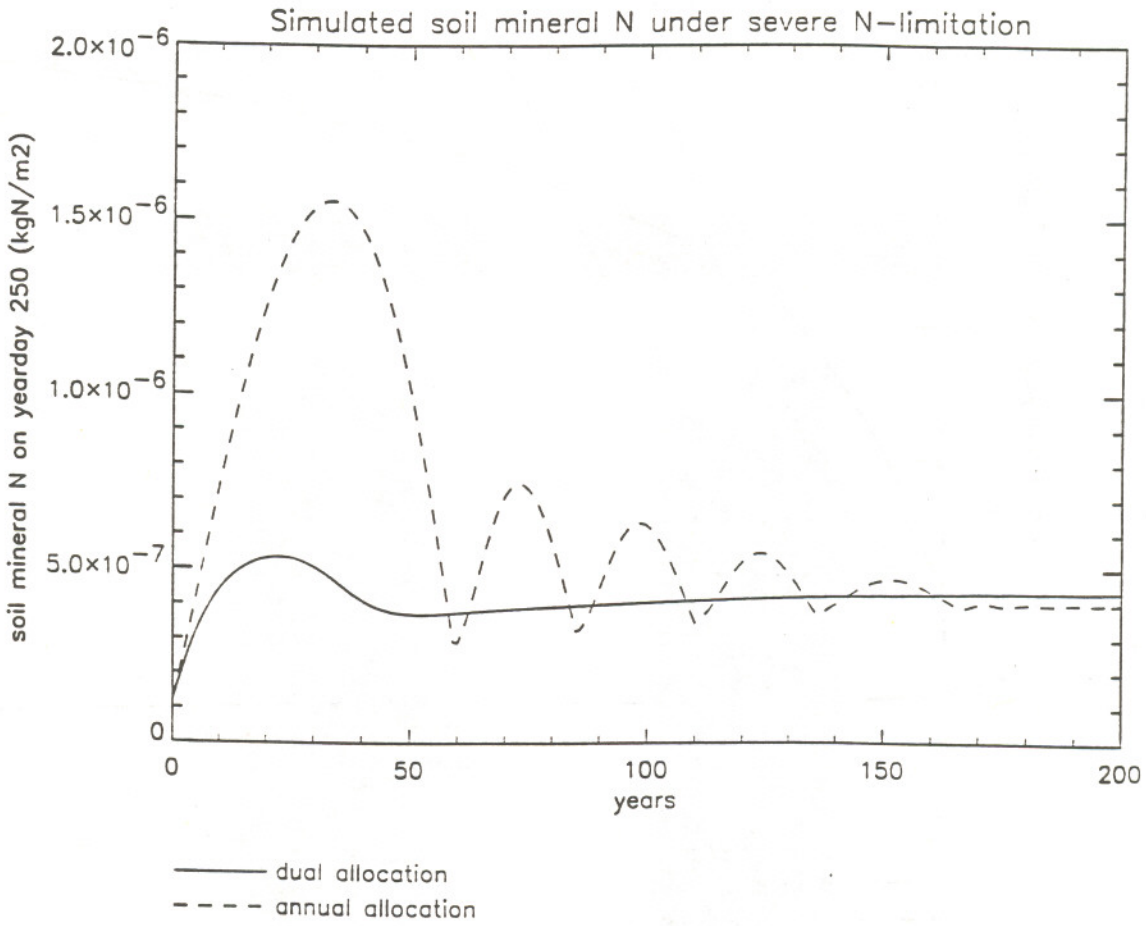

Simulated soil mineral N under severe N-limitation
soil mineral N on yearday 250 (kgN/m2)
2.0x10^-6
1.5x10^-6
1.0x10^-6
5.0x10^-7
0
0
50
100
150
200
years
dual allocation
annual allocation

Figure 9b

availability for annual allocation, but which are damped by the sensitivity of current-year growth to available N with the dual allocation mechanism (Figure 9b).

At around year 900 for both allocation schemes, the dominant limitation switches from available N to available soil water. The long initial period of N-limitation ends when coarse woody debris and litter inputs have increased the soil organic matter content to levels at which annual N mineralization is sufficient to supply the N-immobilization demands imposed by new litter material (Figure 10b). At this point the canopy predictions from the two schemes diverge as a result of the water limitation in the annual allocation mechanism (Figure 10a). For annual allocation, the canopy development is curtailed when the annual minimum soil water content reaches the critical value causing stomatal closure. For dual allocation, the same is true for the part of each year's leaf growth coming out of storage from the previous year, but as long as the daily carbon balance is positive and mineral N is being taken up, the daily growth mechanism can increase leaf area above the level required to just reach critical water content. In this case, water is used earlier in the season and the period of late-season water stress is extended. These dynamics can be observed clearly in the water budget. The annual minimum soil water content over the entire simulation is shown for both schemes in Figure 11a, where the minimum for dual allocation is seen to fall below the minimum for annual allocation after the annual scheme has reached its critical water content. The annual courses of soil water content at years 1, 200, and 2000 are shown in Figure 11b. For years 1 and 200, the two schemes have nearly identical soil water dynamics, and only one curve is shown for each year for the two simulations. In year 2000, however, the soil water dynamics have diverged, as shown by the lowest two curves in Figure 10b. Annual net primary production (ANPP) for the dual scheme is greater than for the annual scheme after the divergence (Figure 12). The reason for a benefit in ANPP at a higher leaf area is not immediately obvious, since a similar amount of water is being used in both schemes, but at different times. The answer, in this case, is that high evaporative demand and reductions in stomatal conduction due to vapor pressure deficit and high temperatures occur in the late summer period more than in the early summer, so that for annual allocation, water saved until the late summer is being used more rapidly for a given stomatal conductance and stomatal conductances are being driven down, while dual allocation results in the same water being used earlier and under more optimal growth conditions.

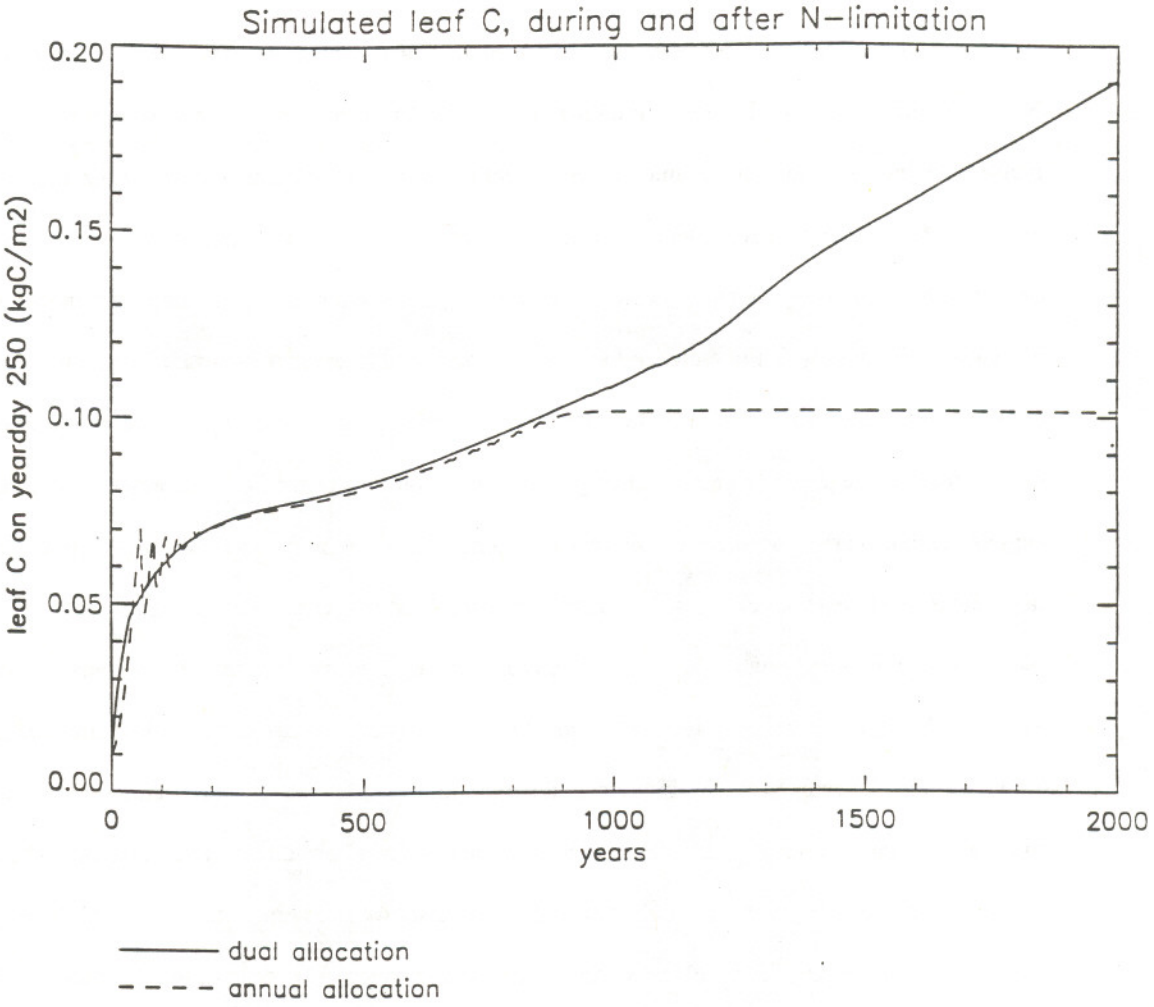

Simulated leaf C, during and after N-limitation
leaf C on yearday 250 (kgC/m2)
0.20
0.15
0.10
0.05
0.00
0
500
1000
1500
2000
years
dual allocation
annual allocation

Figure 10a

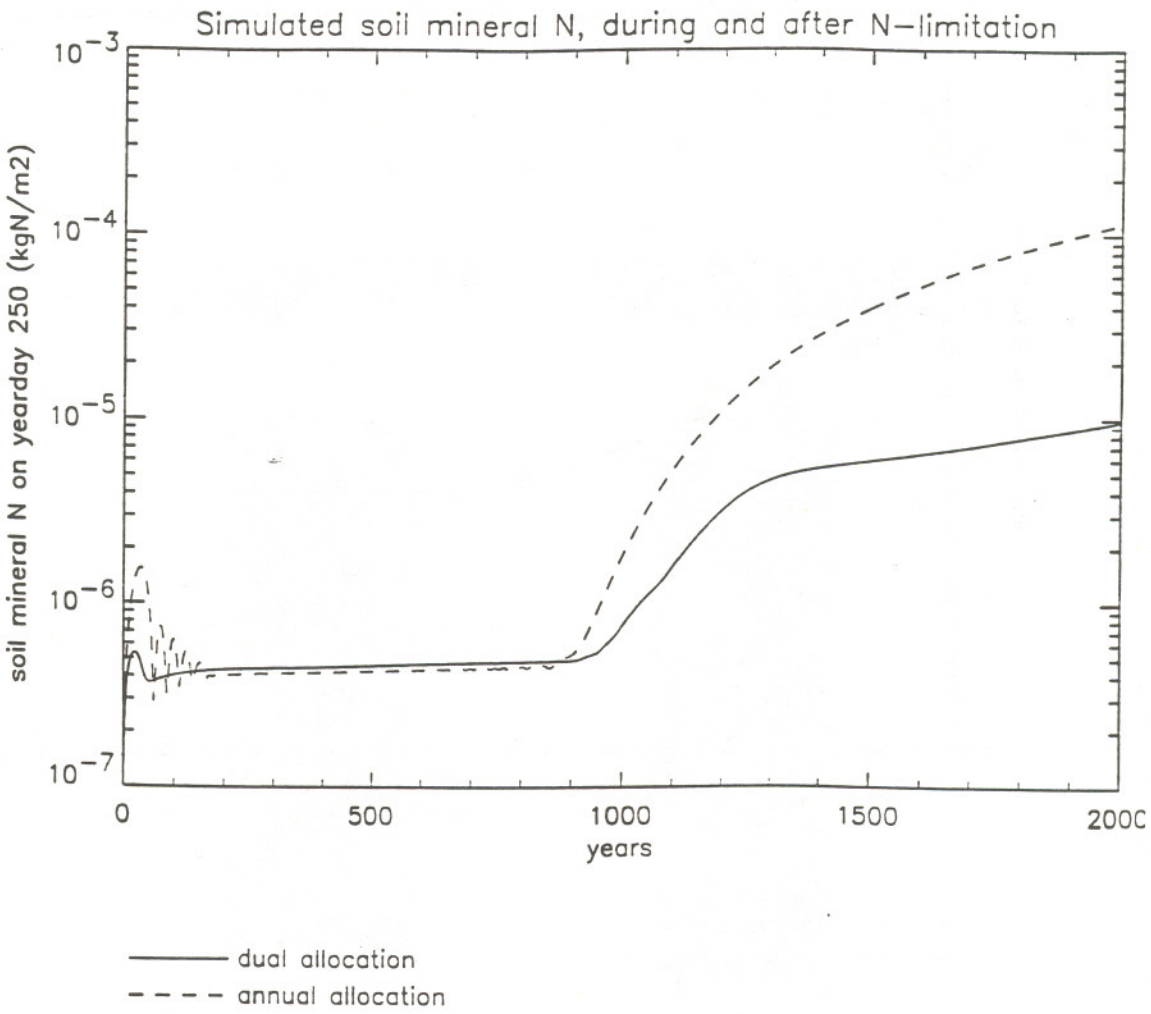

Simulated soil mineral N, during and after N−limitation
soil mineral N on yearday 250 (kgN/m2)
10^-3
10^-4
10^-5
10^-6
10^-7
0
500
1000
1500
2000
years
dual allocation
annual allocation

Figure 10b

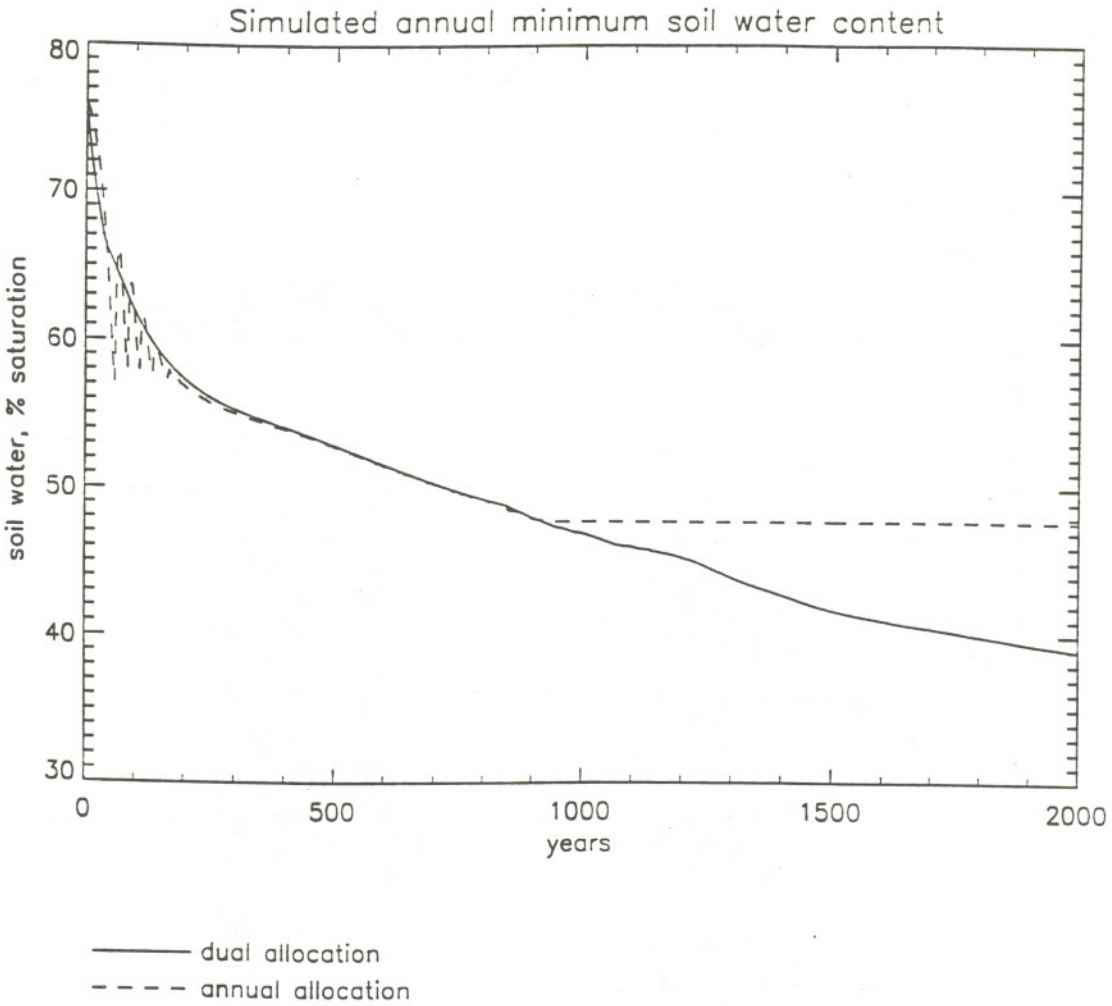

Simulated annual minimum soil water content
soil water, % saturation
80
70
60
50
40
30
0
500
1000
1500
2000
years
dual allocation
annual allocation

Figure 11a

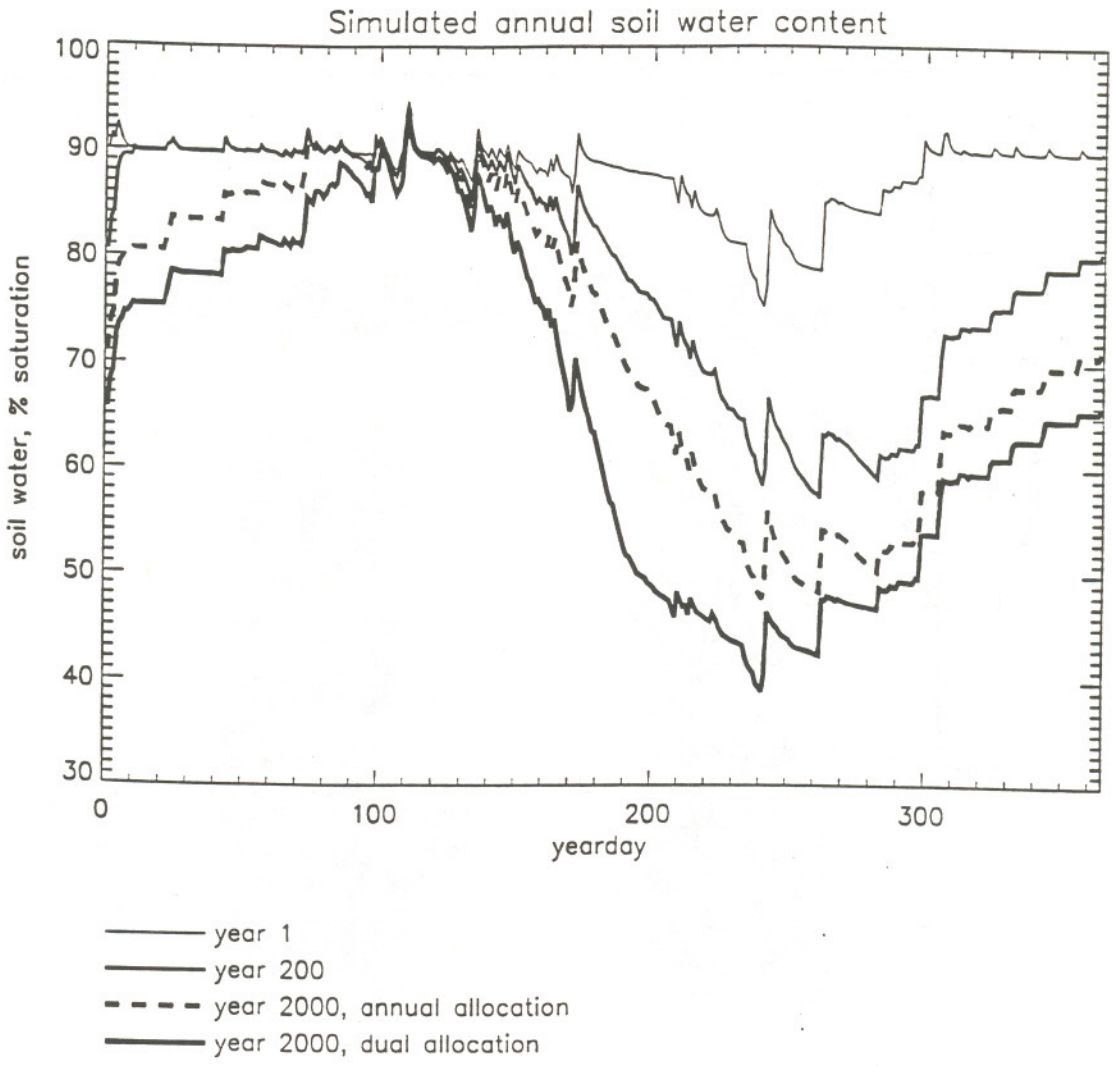

Simulated annual soil water content
soil water, % saturation
100
90
80
70
60
50
40
30
0
100
200
300
yearday
year 1
year 200
year 2000, annual allocation
year 2000, dual allocation

Figure 11b

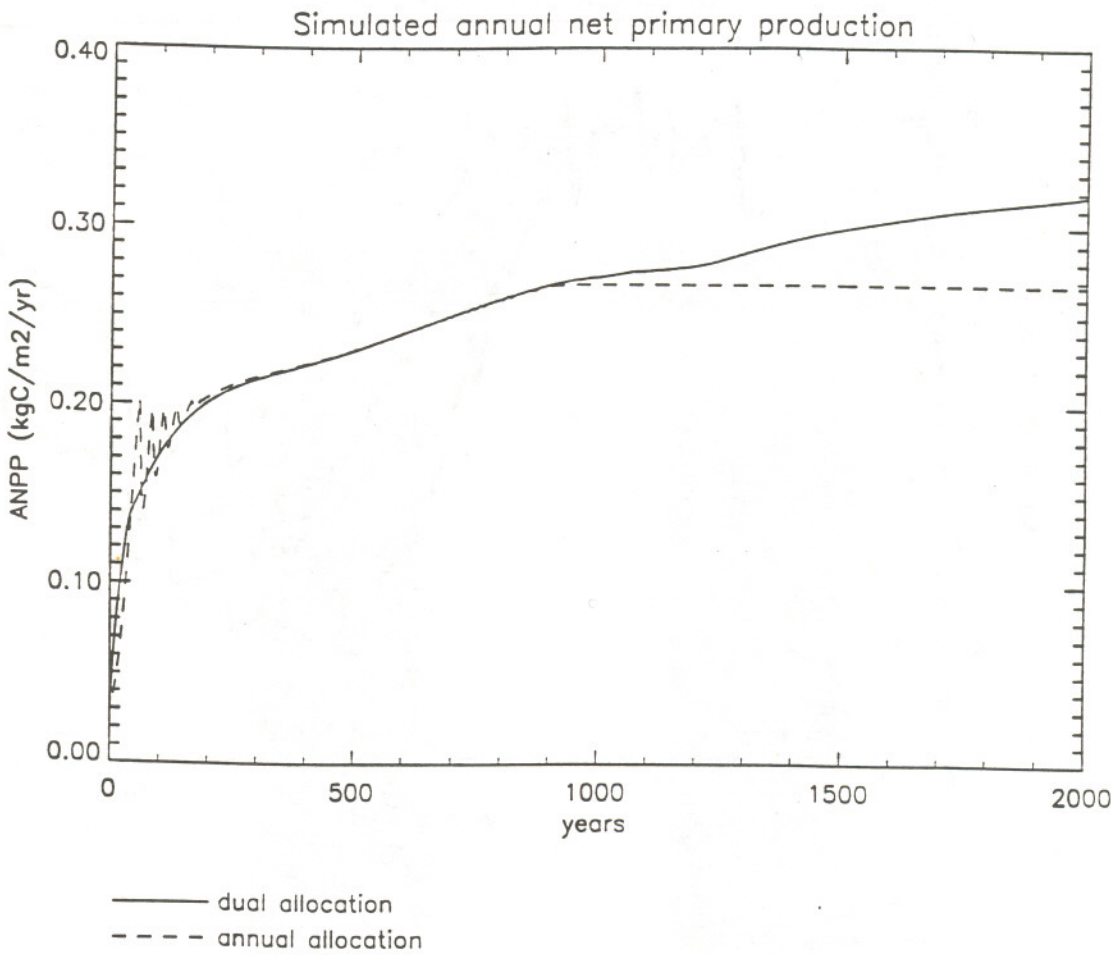

Simulated annual net primary production
ANPP (kgC/m2/yr)
0.40
0.30
0.20
0.10
0.00
0
500
1000
1500
2000
years
dual allocation
annual allocation

Figure 12

For comparison with the ANPP results in Figure 12, ANPP for ponderosa pine stands along an elevation gradient in the Santa Catalina Mountains, Arizona (a dry-summer climate with annual average temperature similar to the Missoula 1984 record) was observed to be 0.31 (kgC m^{-2}) for a stand with L_P = 1.9, and 0.35 for a stand with L_P = 2.4, after estimating root production as 30% of total production (Whittaker and Niering 1968, Whittaker and Niering 1975). ANPP at the end of the annual allocation simulation is 0.27, with L_P = 1.2, and at the end of the dual allocation simulation ANPP = 0.32, with L_P = 2.3.

As an illustration of the sensitivity of the model to interannual variation in environmental conditions influencing growth, I performed two experiments analogous to those just described, but using a 44-year (1950-1993) meteorological data record for Missoula in place of the 1984 record. The 44-year record was repeated as needed for the 2000-year simulations. Leaf C from simulations using the dual allocation scheme are shown for the constant climate and the variable climate in Figure 13a. Note that although the interannual climate variation results in large differences between years in leaf C, the trend due to long-term changes in resource limitation is preserved. For the Missoula climate record, 1984 was milder than average (average maximum temperature, minimum temperature, and vapor pressure deficit were 13.1 °C, 0.8 °C, and 838 Pa, respectively, compared with 13.5 °C, 0.1°C, and 900 Pa for the 44-year average conditions), with average precipitation (34.5 cm in 1984 vs. 35.0 cm for the 44-year average), resulting in predictions of leaf C for under the constant climate scenarios that are higher than the average predictions from the variable climate simulations. A similar pattern is observed in comparing leaf C from the annual allocation simulations for constant and variable climates (Figure 13b). Interannual variation in leaf C is more variable for dual allocation than for annual allocation.

The responses shown in Figures 9-13 represent the initiation of a canopy from what could be considered a "scorched earth" scenario: the very slow response of increasing leaf area over the first 1000 years is the result of a slowly accumulating nitrogen resource, and does not represent the expected behavior of canopy reinitiation after a mature canopy with an equilibrated soil organic matter budget has been removed by rapid disturbance. In order to illustrate the model's behavior under these more typical growth conditions, I used coarse woody debris and soil organic matter values obtained at the end of the 2000-year

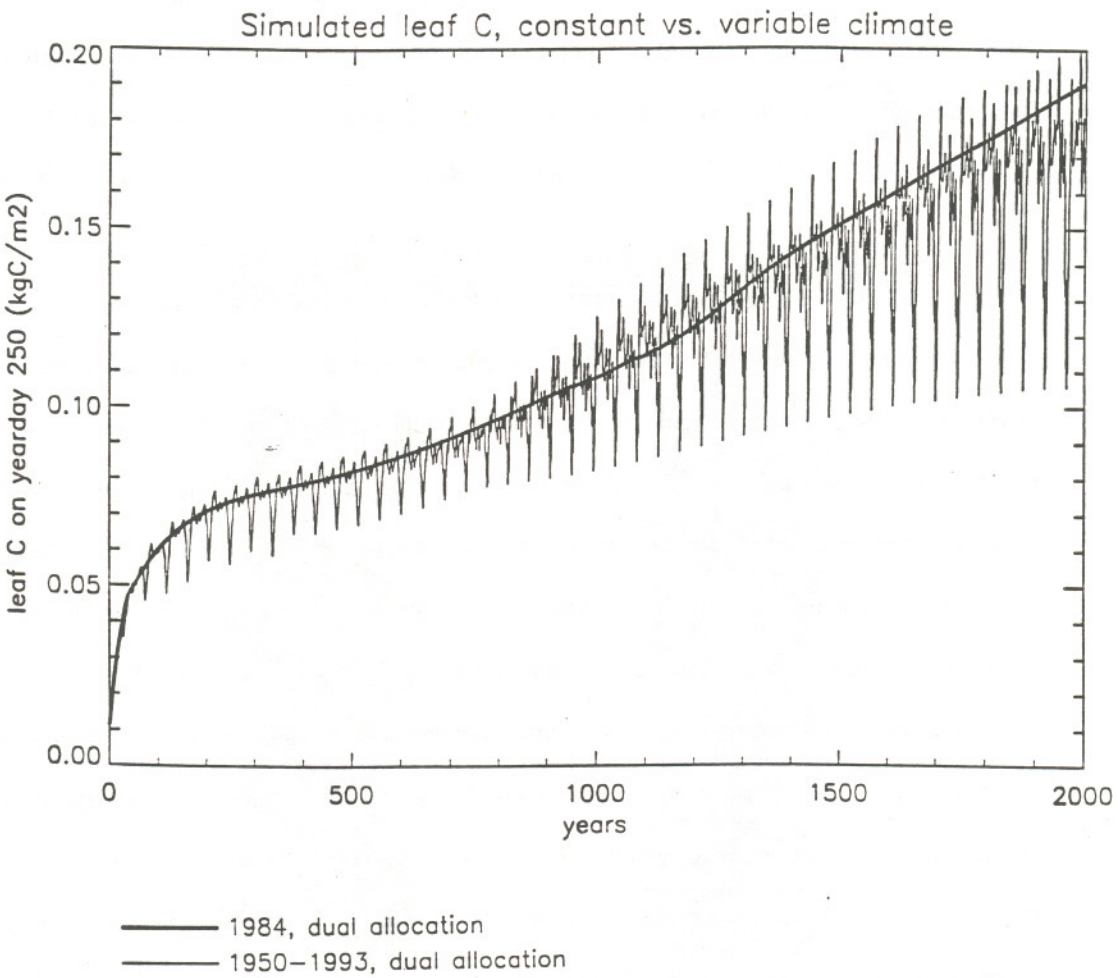

Simulated leaf C, constant vs. variable climate
leaf C on yearday 250 (kgC/m2)
0.20
0.15
0.10
0.05
0.00
0
500
1000
1500
2000
years
1984, dual allocation
1950–1993, dual allocation

Figure 13a

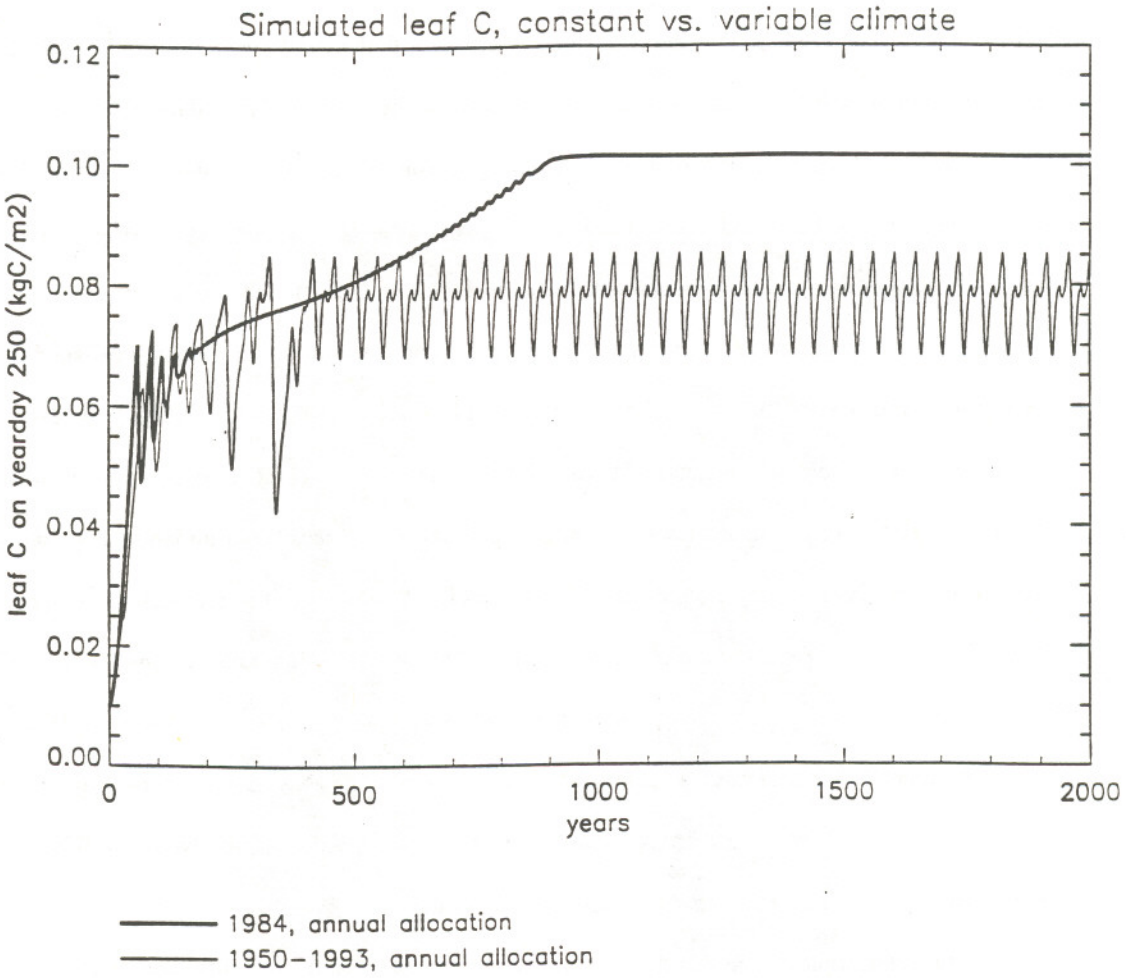

Simulated leaf C, constant vs. variable climate
leaf C on yearday 250 (kgC/m2)
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
500
1000
1500
2000
years
1984, annual allocation
1950–1993, annual allocation

Figure 13b

simulations as the initial conditions for a 100-year simulation, but set all the plant carbon and nitrogen compartments to very low initial values. I performed four such 100-year simulations, with the same variation between simulations as described for the 2000-year experiments. Two experiments were performed with the 1984 climate and two were performed with the 1950-1993 climate. For each climate, one simulation used both annual and daily allocation, while the other used only daily allocation.

Predicted leaf C at the middle of the growing season for these four scenarios over the 100-year simulation period is shown in Figure 14a. The same general patterns noted above are observed here as well, with respect to the differences in equilibrated leaf area between the dual allocation and the annual allocation simulations, and with respect to the relationship between the single climate-year simulations and the variable climate simulations. Notice, however, that both allocation scenarios in the 100-year simulations rise to their equilibrium leaf C values much more rapidly, due to a reduction in N-limitation. The dual allocation mechanism allows for a more rapid leaf area increase after initiation, approaching equilibrium after 20-30 years as opposed to 70-80 years for the annual allocation mechanism operating alone. The rate of increase in total stem wood for the 1984-climate simulations averaged 0.106 kgC m^{-2} yr^{-1} for the final ten years of the 100-year simulations (0.103 kgC m^{-2} yr^{-1} for dual allocation and 0.110 kgC m^{-2} yr^{-1} for annual allocation), which compares well with an average aboveground woody growth rate of 0.117 kgC m^{-2} yr^{-1} observed by Whittaker and Niering (1968) and Whittaker and Niering (1975) for their plots in Arizona, where the average stand age was 121 years.

It is interesting to note that the woody growth rates between the two different allocation scenarios are quite similar, although the scenarios differ by a factor of two in leaf area. As shown in Figure 14b, ANPP for the two allocation scenarios is more similar than leaf area, and since the dual allocation scenario is using more of its ANPP for new leaf growth each year (both scenarios are defined to have the same leaf turnover rate of 0.33 yr^{-1}), the production remaining for "luxury growth" of new dead woody material is about the same in either case. Also note that interannual variation in ANPP is much greater than for leaf C, in part because the 3-year leaf lifespan buffers changes in leaf C, while ANPP is more responsive to current-year growth conditions. The accumulation of carbon in the dead stem compartment is illustrated in Figure 14c, where again the similarity in growth rates for all scenarios by the end of the 100-year

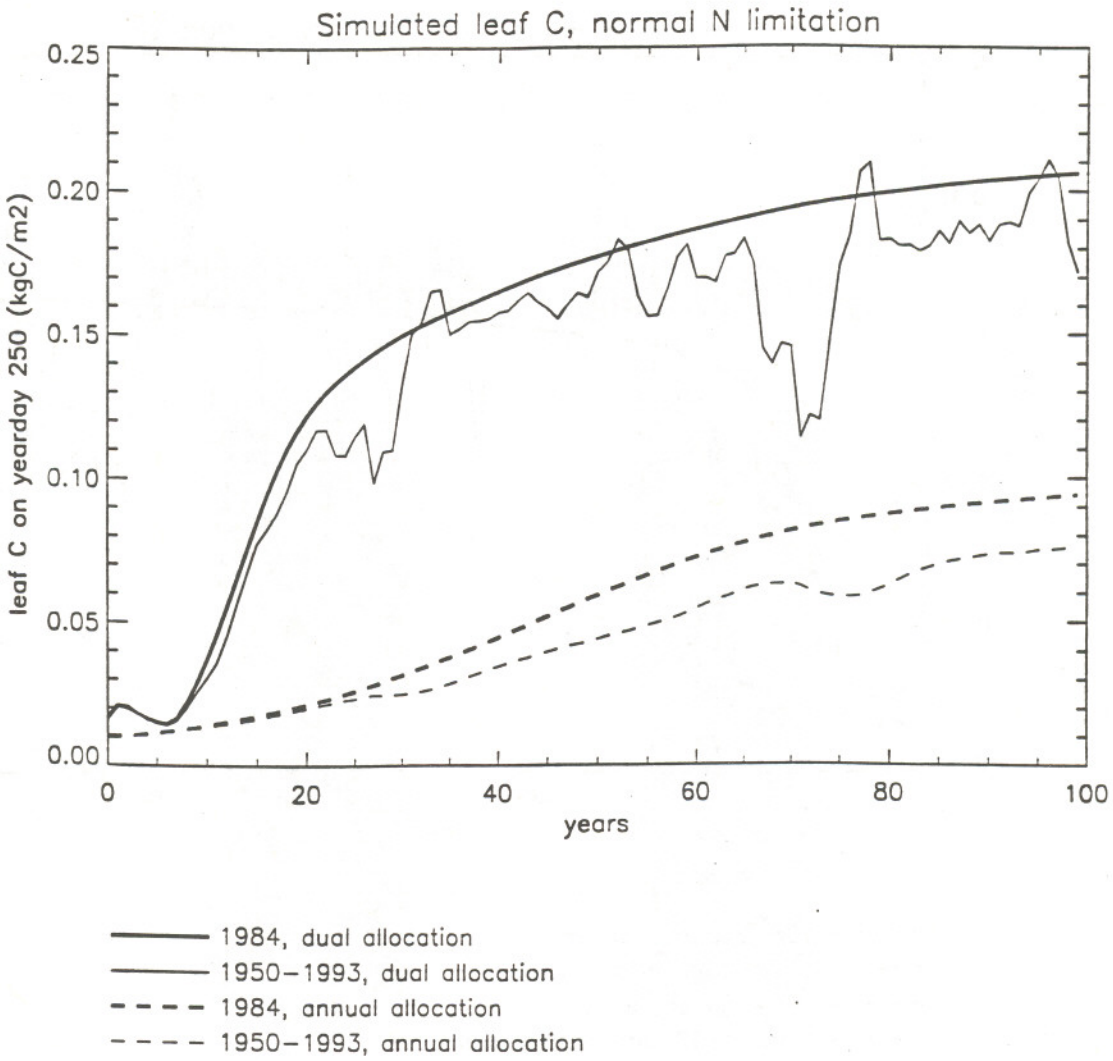

Simulated leaf C, normal N limitation
leaf C on yearday 250 (kgC/m2)
0.25
0.20
0.15
0.10
0.05
0.00
0
20
40
60
80
100
years
1984, dual allocation
1950–1993, dual allocation
1984, annual allocation
1950–1993, annual allocation

Figure 14a

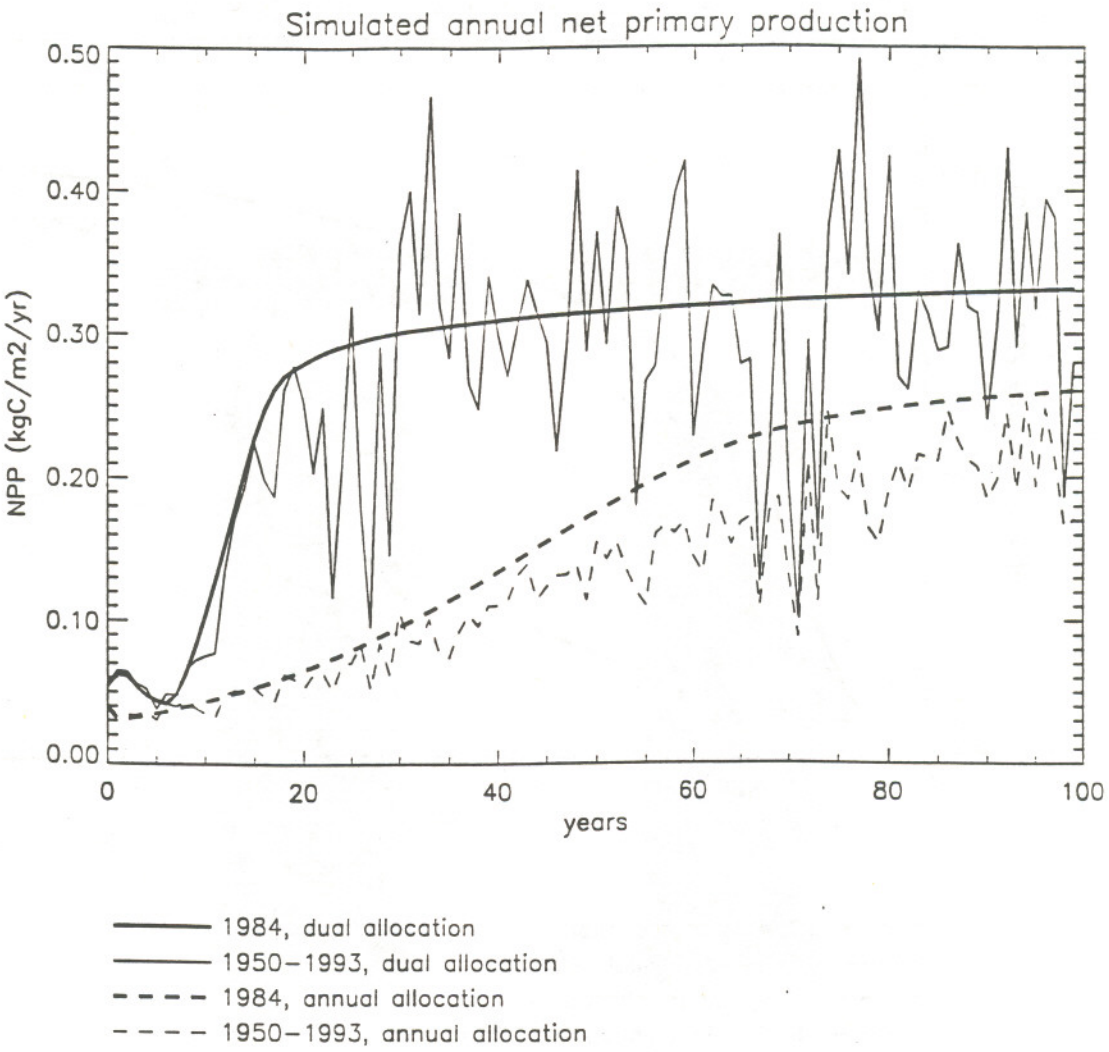

Simulated annual net primary production
NPP (kgC/m2/yr)
0.50
0.40
0.30
0.20
0.10
0.00
0
20
40
60
80
100
years
1984, dual allocation
1950–1993, dual allocation
1984, annual allocation
1950–1993, annual allocation

Figure 14b

simulation is obvious, but it can also be seen that the annual allocation scenarios have accumulated less dead stem C than the dual allocation scenarios by the time rates have equilibrated (around year 90), and that the variable climate scenarios accumulate less than their constant climate counterparts. Simulated annual stem growth increments are shown in Figure 14d, where it can be seen more clearly that stem growth rates for the four scenarios are very similar during the last ten years of the simulation.

Interannual variation in annual stem growth increment is observed to be well-correlated with climatic variation, including variation in precipitation, temperature, and radiation (Fritts and Swetnam 1989). A potentially valuable model validation tool would be a comparison between observed and simulated dendroecological relationships. For example, the simulated annual growth increment for the last 44 years of these simulations, when plotted against annual total precipitation for the same years, shows a reasonable correlation for the dual allocation scenario, and a weak correlation for the annual allocation (Figure 14e). A comparison with observed dendroecological relationships could provide a basis for selecting one or another of these scenarios as more realistic (e.g., Keane *et al.* 1996b).

An exploration of predicted and observed responses to interannual climate variation resulting from various levels of dominance between the daily and annual allocation mechanisms would be an interesting exercise, and of relevance to studies of the global carbon budget under changing climatic conditions (Field *et al.* 1995, VEMAP Members 1995). Urban *et al.* (1993) suggest that simulations of forest productivity are sensitive both to internal model parameters and to the patterns of external climate forcing. My results here indicate that the long-term average response of predictions from 1D-BGC may be less sensitive to the interannual pattern of climate forcing than to the combination of long-term average climatic conditions and model representation of key processes (e.g. allocation). Botkin *et al.* (1992) reach the same conclusion in testing the sensitivity of a forest gap-dynamics model to variation in model parameters and input climate records. Recent model-based analyses of the responses of regional ecosystems to interannual climate variation have concluded that human-induced changes (Burke *et al.* 1991) and periodic climate signals (Yeakley *et al.* 1994) are at least as important as short-term climate variations in predictions of ecosystem productivity. Model predictions of annual net ecosystem production (ANEP) are of particular relevance to the question of global terrestrial ecosystem response to increasing atmospheric CO_2 and possible associated

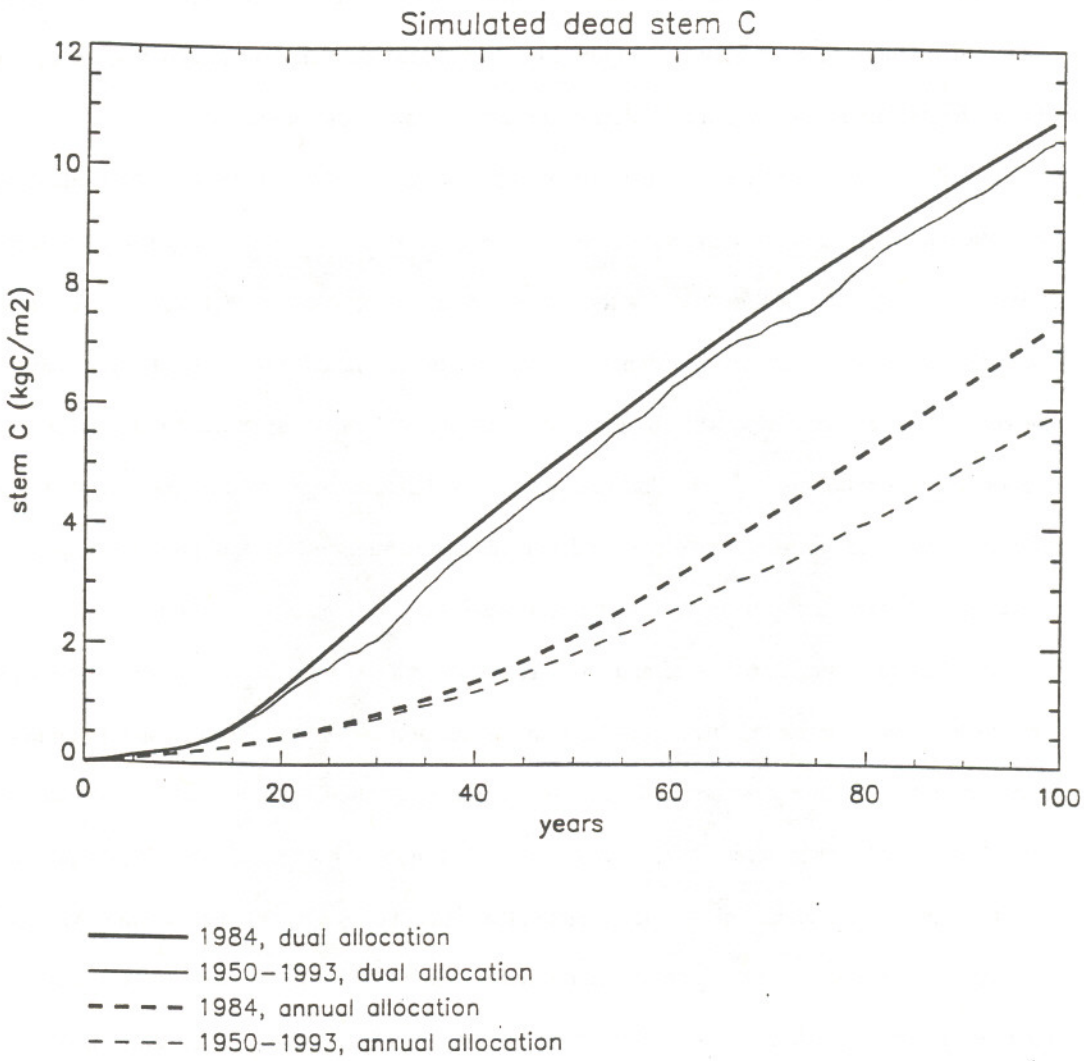

Simulated dead stem C
stem C (kgC/m2)
years
0
2
4
6
8
10
12
0
20
40
60
80
100
1984, dual allocation
1950–1993, dual allocation
1984, annual allocation
1950–1993, annual allocation

Figure 14c

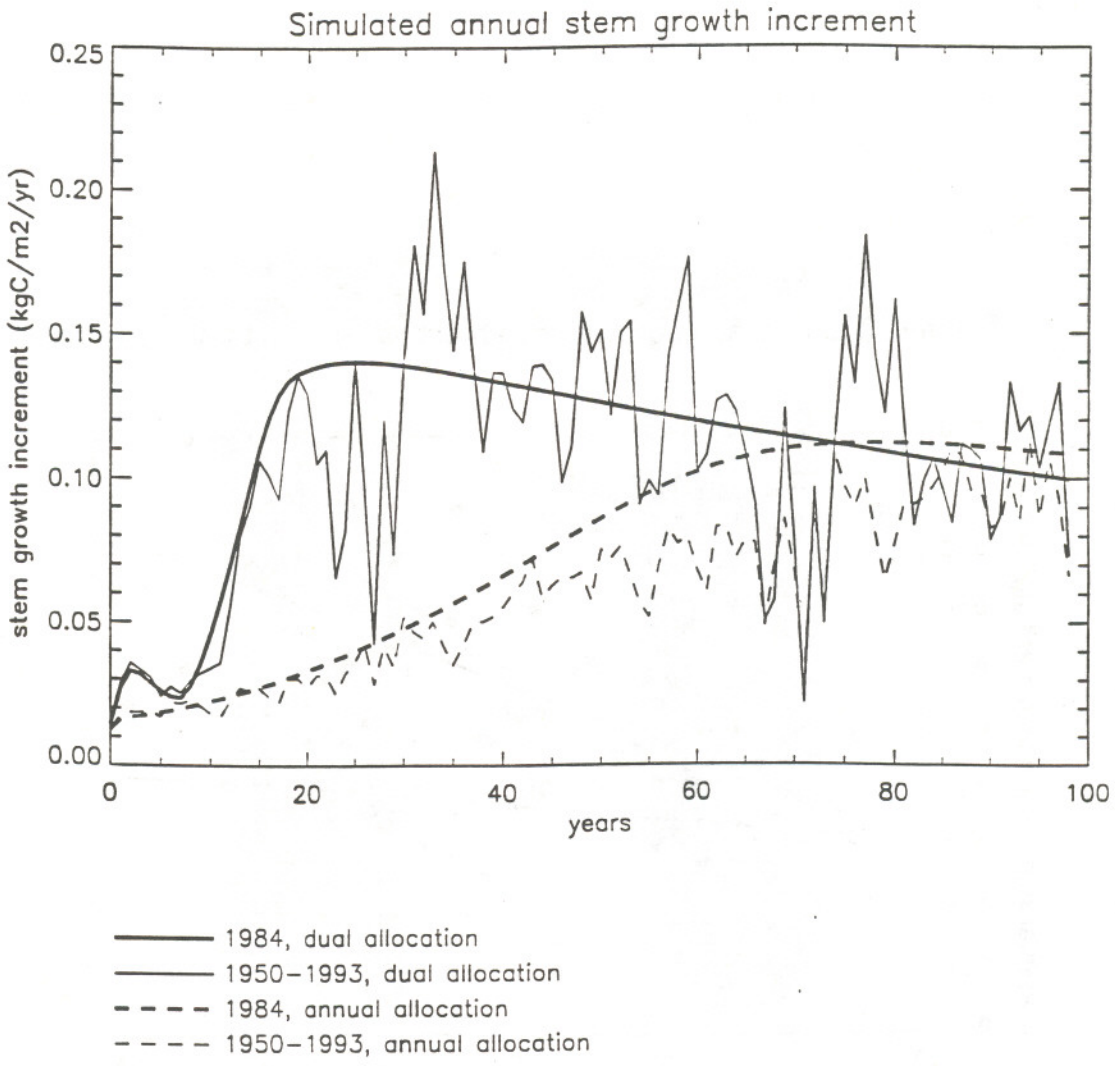

Simulated annual stem growth increment
stem growth increment (kgC/m2/yr)
0.25
0.20
0.15
0.10
0.05
0.00
0
20
40
60
80
100
years
1984, dual allocation
1950–1993, dual allocation
1984, annual allocation
1950–1993, annual allocation

Figure 14d

Simulated annual stem growth increment vs. annual total precipitation

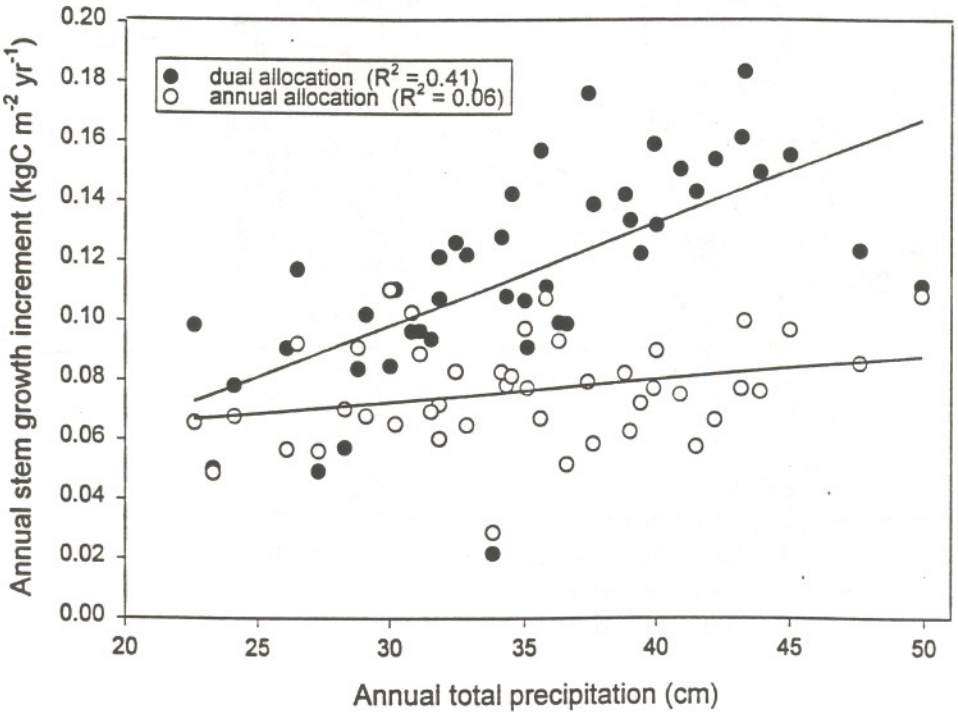


Figure 14e

changes in climatic means or degrees of variation. Estimation of ANEP includes heterotrophic respiration (the primary sources of which are litter and soil respiration) in addition to autotrophic (maintenance and growth) respiration as losses in calculating the carbon budget.

Schimel (1995) suggests that the global deviation of ANEP from zero, when averaged over a period of several years, should be no greater than 1-2% of global ANPP. ANEP for the 100-year simulations is shown in Figure 14f. All scenarios show a negative ANEP (net loss of C to the atmosphere) during the canopy initiation phase, due to respiration from recalcitrant soil organic matter and coarse woody debris pools that were initialized using values from the end of the 2000-year simulations. With a period of low litter input as the canopy and woody biomass increase, heterotrophic respiration losses are unbalanced. All scenarios show a positive ANEP at the end of the 100-year period, due to a steady increase in dead stem and coarse root wood (Figure 14c). Plant wood production continues to increase over multiple centuries for these simulations (results not shown), with long-term positive ANEP averaging 54% of ANPP, suggesting that these simulations exhibit an unrealistically high accumulation of woody biomass.

One very likely explanation for excessive wood accumulation is the negligence of the influence of fire, since it is the primary disturbance mechanism acting to eliminate woody biomass and litter in the climate zone represented by the Missoula climate records (Agee 1993, Keane *et al.* 1990, Ryan 1991). A fire disturbance mechanism will likely be especially important for accurate estimates of ANEP in climates where dry or cold conditions limit decomposition and physical fragmentation of coarse woody debris. The incorporation of fire processes and effects in an ecosystem model such as 1D-BGC presents many problems, including the stochastic nature of the disturbance, the estimation of disturbance frequency, and the simulation of fire intensity and effects on plant, litter, and soil once an event has been predicted to occur. These difficulties have been addressed by Keane *et al.* (1990, 1996a, 1996b) in their development of the FIRESUM and FIRE-BGC models. These models operate simultaneously at multiple spatial scales, integrating stand-level simulations, such as have been discussed here, with a representation of growth for individual trees within subsamples of the stand. As a result of this detailed treatment, the models are computationally intensive, and the spatial extent of their application has so far been relatively limited. It is conceivable, however, that the detailed treatment of fire processes and effects in FIRE-BGC could be

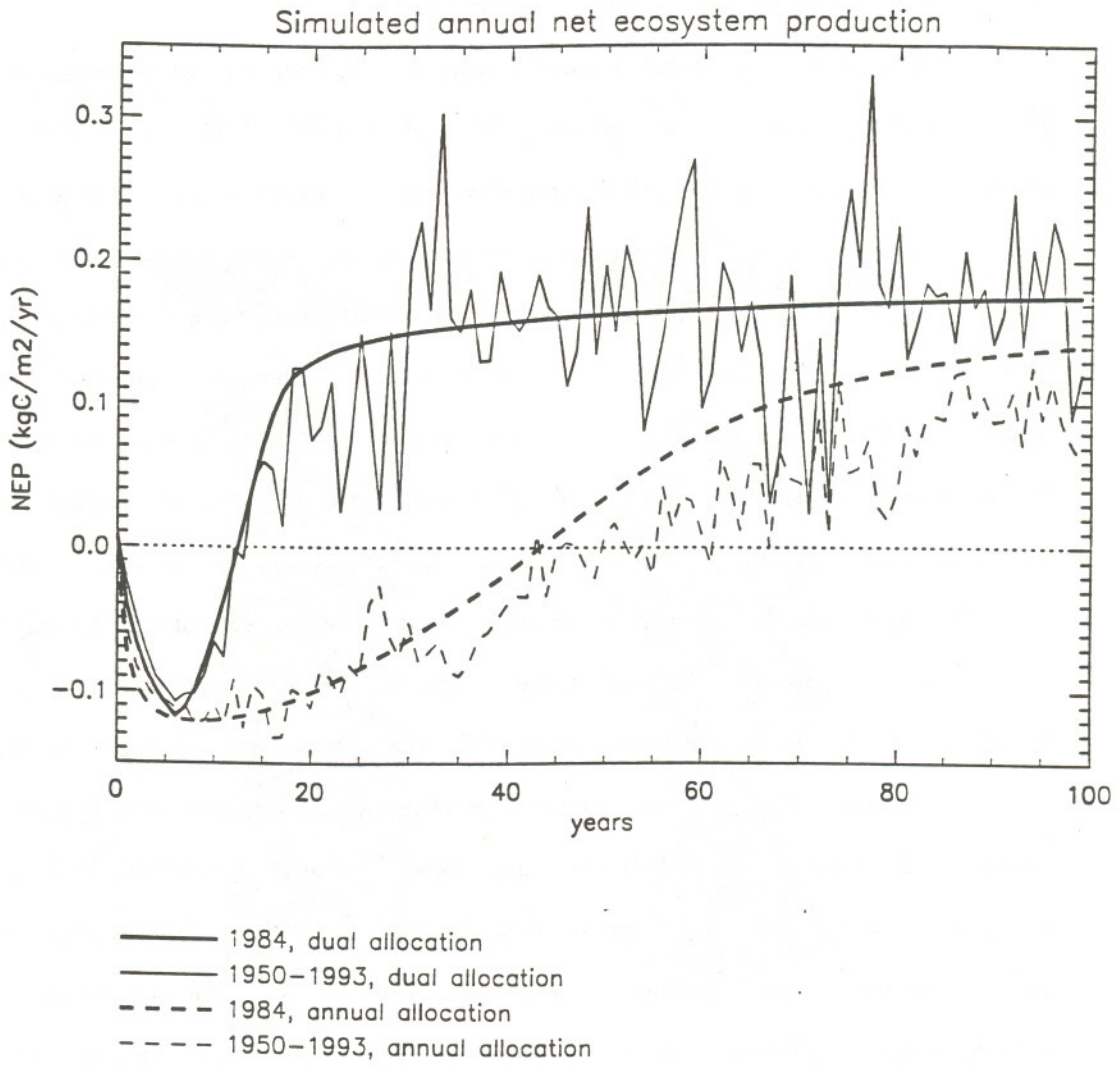

Simulated annual net ecosystem production
NEP (kgC/m2/yr)
0.3
0.2
0.1
0.0
−0.1
0
20
40
60
80
100
years
1984, dual allocation
1950−1993, dual allocation
1984, annual allocation
1950−1993, annual allocation

Figure 14f

exploited to derive a simpler parameterization of some of the most significant influences for incorporation into the 1D-BGC logic, for the purpose of improving long-term estimates of ANEP. I see this as a promising avenue for future model development.

The extension of ecological understanding derived from intensive study at the plot level to large spatial and long temporal scales is necessary if interactions between projected alterations in the global climate and integrative ecological processes such as net production are to be understood (Dale and Rauscher 1994, Chase *et al.* 1996). The application of numerical simulation models is one powerful approach to the problems of regional and global-scale ecosystem analysis (Nemani *et al.* 1996), but it is difficult to validate regional and global extrapolations of ecological knowledge gained at much finer scales (Running *et al.* 1989, McGuire *et al.* 1992, Wessman 1992, Running 1994). In Chapter 4 an extension of the 1D-BGC logic is presented which allows a comparison between model predictions of the surface energy budget and satellite observations of radiometric temperature. The study in Chapter 4, in conjunction with the methods presented in Chapter 2, lays the foundation for a large-scale validation of the energy balance components of the 1D-BGC model.

7. Literature cited

Aber, J.D. and Melillo, J.M., 1980. Litter decomposition: measuring relative contributions of organic matter and nitrogen to forest soils. Canadian Journal of Botany, 58: 416-421.

Aber, J.D., Melillo, J.M. and McClaugherty, C.A., 1990. Predicting long-term patterns of mass loss, nitrogen dynamics, and soil organic matter formation from initial fine litter chemistry in temperate forest ecosystems. Canadian Journal of Botany, 68: 2201-2208.

Agee, J.K., 1993. Fire Ecology of Pacific Northwest Forests. Island Press, Washington, D.C., Covelo, CA.

Allison, F. and Murphy, R.M., 1962. Comparitive rates of decomposition in soil of wood and bark particles of several hardwood species. Soil Science Society Proceedings, 26(5): 463-466.

Andrén, O. and Paustian, K., 1987. Barley straw decomposition in the field: A comparison of models. Ecology, 68(5): 1190-1200.

Bannan, M.W., 1955. The vascular cambium and radial growth of *Thuja occidentalis* L. Canadian Journal of Botany, 33: 113-138.

Barradas, V.L., Jones, H.G. and Clark, J.A., 1994. Stomatal responses to changing irradiance in *Phaseolus vulgaris* L. Journal of Experimental Botany, 45(276): 931-936.

Beets, P.N. and Whitehead, D., 1996. Carbon partitioning in *Pinus radiata* stands in relation to foliage nitrogen status. Tree Physiology, 16: 131-138.

Berg, B., Ekbohm, G. and McClaugherty, C., 1984. Lignin and holocellulose relations during long-term decomposition of some forest litters. Long-term decomposition in a Scots pine forest. IV. Canadian Journal of Botany, 62: 2540-2550.

Berg, B. and McClaugherty, C., 1989. Nitrogen and phosphorus release from decomposing litter in relation to the disappearance of lignin. Canadian Journal of Botany, 67: 1148-1156.

Berg, B. and Tamm, C.O., 1994. Decomposition and nutrient dynamics of litter in long-term optimum nutrition experiments. II. Nutrient concentrations in decomposing *Picea abies* needle litter. Scandinavian Journal of Forest Research, 9: 99-105.

Berninger, F., Mäkelä, A. and Hari, P., 1996. Optimal control of gas exchange during drought: empirical evidence. Annals of Botany, 77: 469-476.

Bolin, B., 1983. The carbon cycle. In: B. Bolin and R.B. Cook (Editors), The Major Biogeochemical Cycles and Their Interactions. Wiley, New York, pp. 41-45.

Bosatta, E. and Ågren, G.I., 1985. Theoretical analysis of decomposition of heterogeneous substrates. Soil Biology and Biochemistry, 17(5): 601-610.

Botkin, D.B. and Nisbet, R.A., 1992. Forest response to climatic change: Effects of parameter estimation and choice of weather patterns on the reliability of projections. Climatic Change, 20: 87-111.

Burke, I.C., Kittel, T.G.F., Lauenroth, W.K., Snook, P., Yonker, C.M. and Parton, W.J., 1991. Regional analysis of the central great plains. BioScience, 41(10): 685-692.

Camillo, P.J. and Gurney, R.J., 1986. A resistance parameter for bare soil evaporation models. Soil Science, 141: 95-105.

Campbell, G.S., 1977. An Introduction to Environmental Biophysics. Springer-Verlag, New York, New York.

Campbell, G.S., 1986. Extinction coefficients for radiation in plant canopies calculated using an ellipsoidal inclination angle distribution. Agricultural and Forest Meteorology, 36: 317-321.

Carpenter, S.R., 1981. Decay of heterogenous detritus: A general model. Journal of Theoretical Biology, 89: 539-547.

Chase, T.N., Pielke, R.A., Kittel, T.G.F., Nemani, R. and Running, S.W., 1996. Sensitivity of a general circulation model to global changes in leaf area index. Journal of Geophysical Research, 101(D3): 7393-7408.

Cheng, W., Coleman, D.C., Carroll, C.R. and Hoffman, C.A., 1993. *In situ* measurement of root, respiration and soluble carbon concentrations in the rhizosphere. Soil Biology and Biochemistry, 25(9): 1189-1196.

Clapp, R.B. and Hornberger, G.M., 1978. Empirical equations for some soil hydraulic properties. Water Resources Research, 14: 601-604.

Cosby, B.J., Hornberger, G.M., Clapp, R.B. and Ginn, T.R., 1984. A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. Water Resources Research, 20: 682-690.

Dale, V.H. and Rauscher, H.M., 1994. Assessing impacts of climate change on forests: The state of biological modeling. Climatic Change, 28: 65-90.

de Pury, D.G.G. and Farquhar, G.D., 1997. Simple scaling of photosynthesis from leaves to canopies without the errors of big-leaf models. Plant, Cell and Environment, 20: 537-557.

Degens, B. and Sparling, G., 1996. Changes in aggregation do not correspond with changes in labile organic C fractions in soil amended with ^{14}C-glucose. Soil Biology and Biochemistry, 28(4/5): 453-462.

Dewar, R.C., 1996. The correlation between plant growth and intercepted radiation: An interpretation in terms of optimal plant nitrogen content. Annals of Botany, 78: 125-136.

Dolman, A.J., 1987. Summer and winter rainfall interception in an oak forest. Predictions with an analytical and a numerical simulation model. Journal of Hydrology, 90: 1-9.

Donnelly, P.K., Entry, J.A., Crawford, D.L. and Cromack, K., Jr., 1990. Cellulose and lignin degradation in forest soils: Response to moisture, temperature, and acidity. Microbial Ecology, 20: 289-295.

Edmonds, R.L., 1984. Long-term decomposition and nutrient dynamics in Pacific silver fir needles in western Washington. Canadian Journal of Forest Research, 14: 395-400.

Edmonds, R.L., 1987. Decomposition rates and nutrient dynamics in small-diameter woody litter in four forest ecosystems in Washington, U.S.A. Canadian Journal of Forest Research, 17: 499-509.

Ellsworth, D.S. and Reich, P.B., 1993. Canopy structure and vertical patterns of photosynthesis and related leaf traits in a deciduous forest. Oecologia, 96: 169-178.

Entry, J.A., Stark, N.M. and Loewenstein, H., 1987. Timber harvesting: Effects on degradation of cellulose and lignin. Forest Ecology and Management, 22: 79-88.

Evans, J.R., 1989. Photosynthesis and nitrogen relationships in leaves in C_3 plants. Oecologia, 78: 9-19.

Evert, R.F., 1960. Phloem structure in *Pyrus communis* L. and its seasonal changes. University of California, Publications in Botany, 32: 127-194.

Farquhar, G.D., von Caemmerer, S. and Berry, J.A., 1980. A biochemical model of photosynthetic CO_2 assimilation in leaves of C_3 species. Planta, 149: 78-90.

Fasman, G.D. (Editor), 1976. Handbook of Biochemistry and Molecular Biology. Proteins, III. CRC Press, Cleveland, OH.

Fassnacht, K.S., Gower, S.T., Norman, J.M. and McMurtrie, R.E., 1994. A comparison of optical and direct methods for estimating foliage surface area index in forests. Agricultural and Forest Meteorology, 71: 183-207.

Field, C., 1983. Allocating leaf nitrogen for the maximization of carbon gain: Leaf age as a control on the allocation program. Oecologia, 56: 341-347.

Field, C. and Mooney, H.A., 1983. Leaf age and seasonal effects on light, water, and nitrogen use efficiency in a California shrub. Oecologia, 56: 348-355.

Field, C. and Mooney, H.A., 1986. The photosynthesis - nitrogen relationship in wild plants, Sixth Maria Moors Cabot Symposium. Evolutionary Constraints on Primary Productivity: Adaptive Patterns of Energy Capture in Plants. Press Syndicate of the University of Cambridge, pp. 25-55.

Field, C.B., Chapin, F.S., III, Matson, P.A. and Mooney, H.A., 1992. Responses of terrestrial ecosystems to the changing atmosphere: a resource-based approach. Annual Reviews of Ecology and Systematics, 23: 201-235.

Field, C.B., Randerson, J.T. and Malmström, C.M., 1995. Global net primary production: Combining ecology and remote sensing. Remote Sensing of Environment, 51: 74-88.

Fritts, H.C. and Swetnam, T.W., 1989. Dendroecology: A tool for evaluating variations in past and present forest environments. Advances in Ecological Research, 19: 111-175.

Gash, J.H.C., 1979. An analytical model of rainfall interception in forests. Quarterly Journal of the Royal Meteorological Society, 105: 43-55.

Gash, J.H.C., Lloyd, C.R. and Lachaud, G., 1995. Estimating sparse forest rainfall interception with an analytical model. Journal of Hydrology, 170: 79-86.

Gates, D.M., 1980. Biophysical Ecology. Springer-Verlag, New York, New York.

Giacomin, A. and Trucchi, P., 1992. Rainfall interception in a beech coppice (Acquerino, Italy). Journal of Hydrology, 137: 141-147.

Graham, J.S. and Running, S.W., 1984. Relative control of air temperature and water status on seasonal transpiration of *Pinus contorta*. Canadian Journal of Forest Research, 14: 833-838.

Grier, C.C., 1978. A *Tsuga heterophylla - Picea sitchensis* ecosystem of coastal Oregon: Decomposition and nutrient balances of fallen logs. Canadian Journal of Forest Research, 8: 198-206.

Grier, C.C. and Running, S.W., 1977. Leaf area of mature northwestern coniferous forests: relation to site water balance. Ecology, 58: 893-899.

Gutschick, V.P., 1995. Physiological control of evapotranspiration by shrubs: scaling measurements from leaf to stand with the aid of comprehensive models, Proceedings: Shrubland ecosystem dynamics in a changing environment. Ogden, UT: U.S. Dept. of Agriculture, Forest Service, Intermountain Research Station (General technical report INT;GTR-338), Las Cruses, NM, pp. 214-219.

Haig, I.t., 1932. Second-growth yield, stand, and volume tables for the western white pine type. Technical Bulletin #323. USDA, Washington, D.C.

Harmon, M.E., Franklin, J.F., Swanson, F.J., Sollins, P., Gregory, S.V., Lattin, J.D., Anderson, N.H., Cline, G.W., Aumen, N.G., Sedell, J.R., Lienkaemper, G.W., Cromack, K., Jr. and Cummins, K.W., 1986. Ecology of Coarse Woody Debris in Temperate Ecosystems. Advances in Ecological Research, 15: 133-302.

Haxeltine, A. and Prentice, I.C., 1996a. BIOME3: An equilibrium terrestrial biosphere model based on ecophysiological constraints, resource availability, and competition among plant functional types. Global Biogeochemical Cycles, 10(4): 693-709.

Haxeltine, A. and Prentice, I.C., 1996b. A general model for the light-use efficiency of primary production. Functional Ecology, 10: 551-561.

Haxeltine, A., Prentice, I.C. and Creswell, I.D., 1996. A coupled carbon and water flux model to predict vegetation structure. Journal of Vegetation Science, 7: 651-666.

Hikosaka, K., Terashima, I. and Katoh, S., 1994. Effects of leaf age, nitrogen nutrition and photon flux density on the distribution of nitrogen among leaves of a vine (*Ipomoea tricolor* Cav.) grown horizontally to avoid mutual shading of leaves. Oecologia, 97: 451-457.

Hirose, T. and Werger, M.J.A., 1994. Photosynthetic capacity and nitrogen partitioning among species in the canopy of a herbaceous plant community. Oecologia, 100: 203-212.

Hirose, T. and Werger, M.J.A., 1995. Canopy structure and photon flux partitioning among species in a herbaceous plant community. Ecology, 76(2): 466-474.

Hoffland, E., van den Boogaard, R., Nelemans, J. and Findenegg, G., 1992. Biosynthesis and root exudation of citric and malic acids in phosphate-starved rape plants. The New Phytologist, 122(4): 675-680.

Hollinger, D.Y., 1996. Optimality and nitrogen allocation in a tree canopy. Tree Physiology, 16: 627-634.

Holmes, W.E. and Zak, D.R., 1994. Soil microbial biomass dynamics and net nitrogen mineralization in northern hardwood ecosystems. Soil Science Society of America Journal, 58: 238-243.

Holmgren, B., Ovhed, M. and Karlsson, P.S., 1996. Measuring and modeling stomatal and aerodynamic conductances of mountain birch: implication for treeline dynamics. Artic and Alpine Research, 28(4): 425-434.

Hunt, E.R., Jr., Piper, S.C., Nemani, R., Keeling, C.D., Otto, R.D. and Running, S.W., 1996. Global net carbon exchange and intra-annual atmospheric CO_2 concentrations predicted by an ecosystem process model and three-dimensional atmospheric transpost model. Global Biogeochemical Cycles, 10(3): 431-456.

Hunt, E.R., Jr. and Running, S.W., 1992. Simulated dry matter yields for aspen and spruce stands in the north american boreal forest. Canadian Journal of Remote Sensing, 18(3): 126-133.

Hunt, E.R., Jr., Weber, J.A. and Gates, D.M., 1985. Effects of nitrate application on *Amaranthus* powellii Wats. III. Optimal allocation of leaf nitrogen for photosynthesis and stomatal conductance. Plant Physiology, 79: 619-624.

Hunt, H.W., 1977. A simulation model for decomposition in grasslands. Ecology, 58: 469-484.

Hunt, H.W., Ingham, E.R., Coleman, D.C., Elliott, E.T. and Reid, C.P.P., 1988. Nitrogen limitation of production and decomposition in prairie, mountain meadow, and pine forest. Ecology, 69(4): 1009-1016.

Irbane, J.V. and Godson, W.L., 1981. Atmospheric Thermodynamics. D. Reidel Publishing Company, Dordrecht, The Netherlands.

Jarvis, P.G. and McNaughton, 1986. Stomatal control of transpiration: Scaling up from leaf to region. Advances in Ecological Research, 15: 1-49.

Jenkinson, D.S., 1971. Studies on the decomposition of ^{14}C labelled organic matter in soils. Soil Science, 111: 64-70.

Johannson, G., 1992. Release of organic C from growing roots of organic fescue (*Festuca pratensis* L.). Soil Biology and Biochemistry, 24(5): 427-433.

Jones, H.G., 1992. Plants and Microclimate. University Press, Cambridge, 428 pp.

Keane, R.E., Arno, S.F., Brown, J.K. and Tomback, D.F., 1990. Modelling stand dynamics in whitebark pine (*Pinus albicaulis*) forests. Ecological Modelling, 51: 73-95.

Keane, R.E., Morgan, P. and Running, S.W., 1996a. FIRE-BGC - A mechanistic ecological process model for simulating fire succession on coniferous forest landscapes of the Northern Rocky Mountains. Res. Pap. INT-RP-484. U.S. Department of Agriculture, Forest Service, Intermountain Research Station, Ogden, UT:, 122 pp.

Keane, R.E., Ryan, K.C. and Running, S.W., 1996b. Simulating effects of fire on northern Rocky Mountain landscapes with the ecological process model FIRE-BGC. Tree Physiology, 16: 319-331.

Kelliher, F.M., Leuning, R., Raupach, M.R. and Schulze, E.-D., 1995. Maximum conductances for evaporation from global vegetation types. Agricultural and Forest Meteorology, 73: 1-16.

Kelliher, F.M., Whitehead, D. and Pollock, D.S., 1992. Rainfall interception by trees and slash in young *Pinus radiata* D. Don stands. Journal of Hydrology, 131: 187-204.

Klaassen, W., Lankreijer, H.J.M. and Veen, A.W.L., 1996. Rainfall interception near a forest edge. Journal of Hydrology, 185: 349-361.

Körner, C., 1995. Leaf diffusive conductances in the major vegetation types of the globe. In: E.-D. Schulze and M.M. Caldwell (Editors), Ecophysiology of Photosynthesis. Springer-Verlag, Berlin Heidelberg New York, pp. 463-490.

Kramer, P.J. and Kozlowski, T.T., 1979. Physiology of Woody Plants. Academic Press, New York San Francisco London.

Kuehn, G.D. and McFadden, B.A., 1969a. Ribulose 1,5-diphosphate carboxylase from *Hydrogenomonas eutropha* and *Hydrogenomonas facilis*. I. Purification, metallic ion requirements, inhibition, and kinetic constants. Biochemistry, 8(6): 2394-2402.

Kuehn, G.D. and McFadden, B.A., 1969b. Ribulose 1,5-diphosphate carboxylase from *Hydrogenomonas eutropha* and *Hydrogenomonas facilis*. II. Molecular weight, subunits, composition, and sulfhydryl groups. Biochemistry, 8(6): 2403-2408.

Ladd, J.N., Jocteur-Monrozier, L. and Amato, M., 1992. Carbon turnover and nitrogen transformations in an alfisol and vertisol amended with [U-^{14}C] glucose and [^{15}N] ammonium sulfate. Soil Biology and Biochemistry, 24(4): 359-371.

Lankreijer, H.J.M., Hendriks, M.J. and Klaassen, W., 1993. A comparison of models simulating rainfall interception of forests. Agricultural and Forest Meteorology, 64: 187-199.

Larcher, W., 1995. Physiological Plant Ecology. Springer-Verlag, Berlin Heidelberg.

Leuning, R., 1990. Modeling stomatal behavior and photosynthesis of *Eucalyptus grandis*. Australian Journal of Plant Physiology, 17: 159-175.

Lloyd, J. and Taylor, J.A., 1994. On the temperature dependence of soil respiration. Functional Ecology, 8: 315-323.

Mahfouf, J.F. and Noilhan, J., 1991. Comparative study of various formulations of evaporation from bare soil using *in situ* data. Journal of Climate and Applied Meteorology, 30(9): 1354-1365.

Mahfouf, J.-F., Ciret, C., Ducharne, A., Irannejad, P., Noilhan, J., Shao, Y., Thornton, P., Xue, Y. and Yang, Z.-L., 1996. Analysis of transpiration results from the RICE and PILPS workshop. Global and Planetary Change, 13: 73-88.

Mäkelä, A., Berninger, F. and Hari, P., 1996. Optimal control of gas exchange during drought: theoretical analysis. Annals of Botany, 77: 461-467.

Marks, D., Dozier, J. and Davis, R.E., 1992. Climate and energy exchange at the snow surface in the alpine region of the Sierrra Nevada. I. Meteorological measurements and monitoring. Water Resources Research, 28(11): 3029-3042.

Martin, J.P., Haider, K. and Kassim, G., 1980. Biodegradation and stabilization after 2 years of specific crop, lignin, and polysaccharide carbons in soils. Soil Science Society of America Journal, 44: 1250-1255.

Mary, B., Fresneau, C., Morel, J.L. and Mariotti, A., 1993. C and N cycling during decomposition of root mucilage, roots and glucose in soil. Soil Biology and Biochemistry, 25(8): 1005-1014.

McArdle, R.E. and Meyer, W.H., 1930. The yield of douglas fir in the Pacific Northwest. Technical Bulletin #201. USDA.

McCaughey, J.H. and Iacobelli, A., 1993. Modelling stomatal conductance in a northern deciduous forest, Chalk River, Ontario. Canadian Journal of Forest Research, 24: 904-910.

McClaugherty, C.A., Pastor, J. and Aber, J.D., 1985. Forest litter decomposition in relation to soil nitrogen dynamics and litter quality. Ecology, 66(1): 266-275.

McGuire, A.D., Melillo, J.M., Joyce, L.A., Kicklighter, D.W., Grace, A.L., Moore, B., III and Vorosmarty, C.J., 1992. Interactions between carbon and nitrogen dynamics in estimating net primary

productivity for potential vegetation in North America. Global Biogeochemical Cycles, 6(2): 101-124.

McMurtrie, R.E., 1991. Relationship of forest productivity to nutrient and carbon supply - a modeling analysis. Tree Physiology, 9: 87-99.

Medlyn, B.E., 1996. The optimal allocation of nitrogen within the C_3 photosynthetic system at elevated CO_2. Australian Journal of Plant Physiology, 23: 593-603.

Melin, E., 1930. Biological decomposition of some types of litter from North American forests. Ecology, 11(1): 72-101.

Meyer, W.H., 1938. Yield of even-aged stands of ponderosa pine. Technical Bulletin #630. USDA.

Minderman, G., 1967. Addition, decomposition and accumulation of organic matter in forests. Journal of Ecology, 56: 355-362.

Monteith, J.L., 1995. A reinterpretation of stomatal responses to humidity. Plant, Cell and Environment, 18: 357-364.

Nadelhoffer, K.J., Aber, J.D. and Melillo, J.M., 1985. Fine roots, net primary production, and soil nitrogen availability: A new hypothesis. Ecology, 66(4): 1377-1390.

Nadelhoffer, K.J. and Raich, J.W., 1992. Fine root production estimates and belowground carbon allocation in forest ecosystems. Ecology, 73(4): 1139-1147.

Neilson, R.P., 1993. Transient ecotone response to climatic change: Some conceptual and modelling approaches. Ecological Applications, 3(3): 385-395.

Neilson, R.P. and Marks, D., 1994. A global perspective of regional vegetation and hydrologic sensitivities from climatic change. Journal of Vegetation Science, 5: 715-730.

Nemani, R.R., Running, S.W., Pielke, R.A. and Chase, T.N., 1996. Global vegetation cover changes form coarse resolution satellite data. Journal of Geophysical Research, 101(D3): 7157-7162.

Nobel, P.S., 1991. Physiochemical and Environmental Plant Physiology. Academic Press, San Diego, CA.

Norman, J.M., 1980. Interfacing leaf and canopy irradiance interception models. In: J.D. Hesketh and J.W. Jones (Editors), Predicting Photosynthesis for Ecosystem Models. CRC Press, Inc., Boca Raton, FL, pp. 49-67.

Olson, J.S., 1963. Energy storage and the balance of producers and decomposers in ecological systems. Ecology, 44(2): 322-331.

Orchard, V.A. and Cook, F.J., 1983. Relationship between soil respiration and soil moisture. Soil Biology and Biochemistry, 15(4): 447-453.

Parton, W.J., Haxeltine, A., Thornton, P., Anne, R. and Hartman, M., 1996. Ecosystem sensitivity to land-surface models and leaf area index. Global and Planetary Change, 13: 89-98.

Parton, W.J., Schimel, D.S., Cole, C.V. and Ojima, D.S., 1987. Analysis of factors controlling soil organic matter levels in great plains grasslands. Soil Science Society of America Journal, 51: 1173-1179.

Parton, W.J., Scurlock, M.O., Ojima, D.S., Gilmanov, T.G., Scholes, R.J., Schimel, D.S., Kirchner, T., Menaut, J.-C., Seastedt, T., Garcia Moya, E., Kamnalrut, A. and Kinyamario, J.I., 1993. Observations and modeling of biomass and soil organic matter dynamics for the grassland biome worldwide. Global Biogeochemical Cycles, 7(4): 785-809.

Pereira, J.S., 1995. Gas exchange and growth. In: E.-D. Schulze and M.M. Caldwell (Editors), Ecophysiology of Photosynthesis. Springer-Verlag, Berlin Heidelberg New York, pp. 147-181.

Peterson, K.L., Moreshet, S. and Fuchs, M., 1991. Stomatal responses of field-growth cotton to radiation and soil moisture. Agronomy Journal, 83: 1059-1065.

Potter, C.S., Randerson, J.T., Field, C.B., Matson, P.A., Vitousek, P.M., Mooney, H.A. and Klooster, S.A., 1993. Terrestrial ecosystem production: a process model based on global satellite and surface data. Global Biogeochemical Cycles, 7(4): 811-841.

Pregitzer, K.S., Hendrick, R.L. and Fogel, R., 1993. The demography of fine roots in response to patches of water and nitrogen. The New Phytologist, 125: 575-580.

Prentice, I.C., Cramer, W., Harrison, S., Leemans, R., Monserud, R.A. and Solomon, A.M., 1992. A global biome model based on plant physiology and dominance, soil properties and climate. Journal of Biogeography, 19: 117-134.

Raich, J.W. and Nadelhoffer, K.J., 1989. Belowground carbon allocation in forest ecosystems: Global trends. Ecology, 70(5): 1346-1354.

Randerson, J.T., Thompson, M.V., Malmstrom, C.M., FIeld, C.B. and Fung, I.Y., 1996. Substrate limitations for heterotrophs: Implications for models that estimate the seasonal cycle of atmospheric CO_2. Global Biogeochemical Cycles, 10(4): 585-602.

Randlett, D.L., Zak, D.R., Pregitzer, K.S. and Curtis, P.S., 1996. Elevated atmospheric carbon dioxide and leaf litter chemistry: Influences on microbial respiration and net nitrogen mineralization. Soil Science Society of America Journal, 60: 1571-1577.

Rastetter, E.B., Ryan, M.G., Shaver, G.R., Melillo, J.M., Nadelhoffer, K.J., Hobbie, J.E. and Aber, J.D., 1991. A general biogeochemical model describing the responses of the C and N cycles in terrestrial ecosystems to changes in CO_2, climate and N deposition. Tree Physiology, 9: 101-126.

Reich, P.B., Kloeppel, B.D., Ellsworth, D.S. and Walters, M.B., 1995. Different photosynthesis-nitrogen relations in deciduous hardwood and evergreen coniferous tree species. Oecologia, 104: 24-30.

Reich, P.B. and Walters, M.B., 1994. Photosynthesis-nitrogen relations in Amazonian tree species. II. Variation in nitrogen vis-a-vis specific leaf area influences mass- and area-based expressions. Oecologia, 97: 73-81.

Richards, E.H. and Norman, A.G., 1931. The biological decomposition of plant materials. V. Some factors determining the quantity of nitrogen immobilised during decomposition. Biochemistry Journal, 25: 1769-1778.

Running, S.W., 1984. Microclimate control of forest productivity: analysis by computer simulation of annual photosynthesis/transpiration balace in different environments. Agricultural and Forest Meteorology, 32: 267-288.

Running, S.W., 1994. Testing FOREST-BGC ecosystem process simulations across a climatic gradient in Oregon. Ecological Applications, 4(2): 238-247.

Running, S.W. and Coughlan, J.C., 1988. A general model of forest ecosystem processes for regional applications. I. Hydrological balance, canopy gas exchange and primary production processes. Ecological Modelling, 42: 125-154.

Running, S.W. and Gower, S.T., 1991. FOREST BGC, A general model of forest ecosystem processes for regional applications. II. Dynamic carbon allocation and nitrogen budgets. Tree Physiology, 9: 147-160.

Running, S.W., Nemani, R.R., Peterson, D.L., Band, L.E., Potts, D.F., Pierce, L.L. and Spanner, M.A., 1989. Mapping regional forest evapotranspiration and photosynthesis by coupling satellite data with ecosystem simlation. Ecology, 70(4): 1090-1101.

Ryan, K.C., 1991a. Vegetation and wildland fire: implications of global climate change. Environment International, 17: 169-178.

Ryan, M.G., 1991b. A simple method for estimating gross carbon budgets for vegetation in forest ecosystems. Tree Physiology, 9: 255-266.

Saggar, S., Tate, K.R., Feltham, C.W., Childs, C.W. and Parshotam, A., 1994. Carbon turnover in a range of allophanic soils amended with ^{14}C-labelled glucose. Soil Biology and Biochemistry, 26(9): 1263-1271.

Sampson, D.A. and Smith, F.W., 1993. Influence of canopy architecture on light penetration in logepole pine (*Pinus contorta* var. *latifolia*) forests. Agricultural and Forest Meteorology, 64: 63-79.

Saxton, K.E., Rawls, W.J., Romberger, J.S. and Papendick, R.I., 1986. Estimating generalized soil-water characteristics from texture. Soil Science Society of America Journal, 50: 1031-1036.

Scatena, F.N., 1990. Watershed scale rainfall interception on two forested watersheds in the Luquillo Mountains of Puerto Rico. Journal of Hydrology, 113: 89-102.

Schimel, D.S., 1995. Terrsetrial ecosystems and the carbon cycle. Global Change Biology, 1: 77-91.

Schulze, E.-D., Lange, O.L., Evenori, M., Kappen, L. and Buschbom, U., 1980. Long-term effects of drought on wild and cultivated plants in the Negev Desert. II. Dirurnal patterns of net photosynthesis and daily carbon gain. Oecologia, 45: 19-25.

Schulze, E.-D. and Hall, A.E., 1982. Stomatal responses, water loss, and CO_2 assimilation rates of plants in contrasting environments. In: O.L. Lange, P.S. Nobel, C.B. Osmond and H. Ziegler (Editors), Physiological Plant Ecology. II. Encyclopedia of Plant Physiology, New Series. Springer-Verlag, Berlin, pp. 181-230.

Schwenke, H. and Wagner, E., 1992. A new concept of root exudation. Plant, Cell and Environment, 15(3): 289-299.

Sellers, P.J., Berry, J.A., Collatz, G.J., Field, C.B. and Hall, F.G., 1992. Canopy reflectance, photosynthesis, and transpiration. III. A reanalysis using improved leaf models and a new canopy integration scheme. Remote Sensing of Environment, 42: 187-216.

Shao, Y., Anne, R.D., Henderson-Sellers, A., Irannejad, P., Thornton, P.E., Liang, X., Chen, T.H., Ciret, C., Desborough, C., Balachove, O., Haxeltine, A. and Ducharne, A., 1994. Soil moisture simulation: A report of the RICE and PILPS workshop. IGPO Publication Series. World Climate Research Programme.

Shao, Y. and Henderson-Sellers, A., 1996. Validation of soil moisture simulation in landsurface parameterisation schemes with HAPEX data. Global and Planetary Change, 13: 11-46.

Sims, D.A. and Pearcy, R.W., 1989. Photosynthesis characteristics of a tropical forest understory herb, *Alocasia macrorrhiza*, and a related crop species, *Colocasia esculenta* grown in contrasting light environments. Oecologia, 79: 53-59.

Sinclair, T.R., Murphy, C.E. and Knoerr, K.R., 1976. Development and evaluation of simplified models for simulating canopy photosynthesis and transpiration. Journal of Applied Ecology, 13: 813-829.

Smith, W.K., Young, D.R., Carter, G.A., Hadley, J.L. and McNaughton, G.M., 1984. Autumn stomatal closure in six conifer species of the central Rocky Mountains. Oecologia, 63: 237-242.

Smolander, A., Kurka, A., Kitunen, V. and Mälkönen, E., 1994. Microbial biomass C and N, and repiratory activity in soil of repeatedly limed and N- and P-fertilized Norway Spruce stands. Soil Biology and Biochemistry, 26(8): 957-962.

Sollins, P., 1982. Input and decay of coarse woody debris in coniferous stands in western Oregon and Washington. Canadian Journal of Forest Research, 12: 18-28.

Sørensen, L.H., 1981. Carbon-nitrogen relationships during the humification of cellulose in soils containing different amounts of clay. Soil Biology and Biochemistry, 13: 313-321.

Sparks, T.H. and Carey, P.O., 1995. The response of species to climate over two centuries: An analysis of the Marsham phenological records, 1736-1947. Journal of Ecology, 83: 3321-329.

Stevenson, J.S., 1986. Cycles of Soil. John Wiley and Sons, Inc., New York Chichester Brisbane Toronto Singapore.

Stump, L.M. and Binkley, D., 1993. Relationships between litter quality and nitrogen availability in Rocky Mountain forests. Canadian Journal of Forest Research, 23: 492-502.

Taylor, B.R., Parkinson, D. and Parsons, W.F.J., 1989. Nitrogen and lignin content as predictors of litter decay rates: A microcosm test. Ecology, 70(1): 97-104.

Taylor, B.R., Prescott, C.E., Parsons, W.J.F. and Parkinson, D., 1991. Substrate control of litter decomposition in four Rocky Mountain coniferous forests. Canadian Journal of Botany, 69: 2242-2250.

Tenhunen, J.D., Beyschlag, W., Lange, O.L. and Harley, P.C., 1987. Changes during summer drought in leaf CO_2 uptake rates of macchia shrubs growing in Portugal: Limitations due to photosynthetic capacity, carboxylation efficiency, and stomatal conductance. In: J.D. Tenhunen, F.M. Catarino, O.L. Lange and W.C. Oechel (Editors), Plant Responses to Stress. NATO ASI Series G15. Springer, Berlin Heidelberg New York, pp. 305-327.

Thornton, P.E., Running, S.W. and White, M.A., 1997. Generating surfaces of daily meteorological variables over large regions of complex terrain. Journal of Hydrology, 190: 214-251.

Urban, D.L., Harmon, M.E. and Halpern, C.B., 1993. Potential response of Pacific northwestern forests to climatic change, effects of stand age and initial composition. Climatic Change, 23: 247-266.

van de Griend, A.A. and Owe, M., 1994. Bare soil surface resistance to evaporation by vapor diffusion under semiarid conditions. Water Resources Research, 30(2): 181-188.

van Veen, J.A., Ladd, J.N. and Frissel, M.J., 1984. Modelling C and N turnover through the microbial biomass in soil. Plant and Soil, 76: 257-274.

van Veen, J.A. and Paul, E.A., 1981. Organic carbon dynamics in grassland soils. 1. Background information and computer simulation. Canadian Journal of Soil Science, 61(2): 185-201.

VEMAP Members, 1995. Vegetation/ecosystem modeling and analysis project: Comparing biogeography and biogeochemistry models in a continental-scale study of terrestrial ecosystem responses to climatic change and CO_2 doubling. Global Biogeochemical Cycles, 9(4): 407-437.

Vogt, K.A., Grier, C.C. and Vogt, D.J., 1986. Production, turnover, and nutrient dynamics of above- and below-ground detritus of world forests. Advances in Ecological Research, 15: 303-377.

Voroney, R.P. and Paul, E.A., 1984. Determination of k_C and k_N in situ for calibration of the chloroform fumigation-incubation method. Soil Biology and Biochemistry, 16(1): 9-14.

Waksman, S.A. and Gerretsen, F.C., 1931. Influence of temperature and moisture upon the nature and extent of decomposition of plant residues by microorganisms. Ecology, 12(1): 33-60.

Waring, R.H. and Pitman, G.B., 1985. Modifying lodgepole pine stands to change susceptibility to mountain pine beetle attack. Ecology, 66: 889-897.

Waring, R.H. and Schlesinger, W.H., 1985. Forest Ecosystems. Academic Press, Inc., San Diego, CA.

Wessman, C.A., 1992. Spatial scales and global change: bridging the gap from plots to GCM grid cells. Annual Reviews of Ecology and Systematics, 23: 175-200.

White, M.A., Thornton, P.E. and Running, S.W., 1997. A continental phenology model for monitoring vegetation responses to interannual climatic variability. Global Biogeochemical Cycles, 11(2): 217-234.

Whittaker, R.H. and Niering, W.A., 1968. Vegetation of the Santa Catalina Mountains, Arizona. IV. Limestone and acid soils. Journal of Ecology, 56: 523-544.

Whittaker, R.H. and Niering, W.A., 1975. Vegetation of the Santa Catalina Mountains, Arizona. V. Biomass, production, and diversity along the elevational gradient. Ecology, 56: 771-790.

Wieder, R.K. and Lang, G.E., 1982. A critique of the analytical methods used in examining decomposition data obtained from litter bags. Ecology, 63(6): 1636-1642.

Woodrow, I.E. and Berry, J.A., 1988. Enzymatic regulation of photosynthetic CO_2 fixation on C_3 plants. Annual Reviews of Plant Physiology and Plant Molecular Biology, 39: 533-594.

Wullschleger, S.D., 1993. Biochemical limitations to carbon assimilation in C_3 plants - A retrospective analysis of the A/C_i curves from 109 species. Journal of Experimental Botany, 44(262): 907-920.

Xu, J.G. and Juma, N.G., 1994. Relations of shoot C, root C, and root length with root-released C of two barley cultivars and the decomposition of root-released C in soil. Canadian Journal of Soil Science, 74(1): 17-22.

Yang, X., Miller, D.R. and Montgomery, M.E., 1993. Vertical distribution of canopy foliage and biologically active radiation in a defoliated/refoliated hardwood forest. Agricultural and Forest Meteorology, 67: 129-146.

Yeakley, J.A., Moen, R.A., Breshears, D.D. and Nungesser, M.K., 1994. Responses of North American ecosystem models to multi-annual periodicities in temperature and precipitation. Landscape Ecology, 9(4): 249-260.

Zak, D.R., Tilman, D., Parmenter, R.R., Rice, C.W., Fisher, F.M., Vose, J., Milchunas, D. and Martin, C.W., 1994. Plant production and soil microorganisms in late-successional ecosystems: A continental-scale study. Ecology, 75(8): 2333-2347.

Zheng, D., Hunt, E.R., Jr. and Running, S.W., 1993. A daily soil temperature model based on air temperature and precipitation for continental applications. Climate Research, 2: 183-191.

Chapter 4

Estimation of surface resistance parameters from remote sensing and surface meteorological observations

1. Introduction

In the first section I present some basic relationships relevant to surface flux and resistance formulations. The purpose of this section is to establish a quantitative framework as a precursor to a review of relevant literature. Some of the relationships introduced there are expanded upon in later sections. Following this overview of theory I present a discussion of relevant literature. Following the literature review, I present an operational statement of my research problem. This consists of a condensation of the quantitative relationships predicted by the energy balance theory, and a schematic presentation of the logic by which I intend to test those relationships with the available data. I then present a detailed account of the methods used, followed by results and discussion.

2. Review of theory of mass and energy storage and transport

The first part of this discussion addresses the Earth's surface radiation budget, and the second part the surface energy balance, including convective transport terms and a discussion of micrometeorological methods.

2.1 Radiation budget

Most of the commonly observed processes at the Earth's surface (wind, waves, rain, cloud formation, plant growth, etc.) are driven by energy received from the Sun. This energy arrives at the exterior of the atmosphere in the form of electromagnetic radiation, photons, with an intensity and distribution of wavelengths that is characteristic of the temperature of the radiating surface of the Sun. An object's characteristic distribution of energy over a range of wavelengths is known as its emission

spectrum. The following equations describe the total amount of radiation (Eq. 1) and the dominant wavelength (Eq. 2) in the emission spectrum of an ideal radiator, or black body (an object that absorbs and emits perfectly at all wavelengths):

Eq. 1 $R_T = \sigma T^4$

where R_T = total radiant energy (Wm^{-2}), T = radiant surface temperature (K), and σ is the Stefan-Boltzmann constant ($5.67x10^{-8}$ $Wm^{-2}K^{-4}$).

Eq. 2 $\lambda_{max} = \frac{2.88x10^{-3}}{T}$

where λ_{max} is the peak wavelength in the emission spectrum (m).

For a surface temperature of 6000 K, representative of the Sun, the emission peak predicted by Eq. 2 is at about 0.48 μm, in the middle of the visible spectrum. For a surface temperature of 300 K, typical of the Earth's surface, the peak emission is at about 9.6 μm, in the thermal infrared range. Eqs. 1 and 2 refer to ideal radiators. In the case of real objects with non-ideal radiative properties a correction is applied; the emissivity, ε. Emissivity can vary with wavelength, in which case it is referred to as a narrowband emissivity. For typical Earth surface emission temperatures, it is more usual to consider the emissivity in a broad band, from 3-100 μm, and for the moment that is the sense in which I will use the term. The corrected form of Eq. 1 is:

Eq. 3 $R_T = \varepsilon\sigma T^4$

The radiant energy from the Sun passes through Earth's atmosphere relatively easily: of 1367 Wm^{-2} incident at the top of the atmosphere as much as 1000 Wm^{-2} (i.e., about 73%) can reach the surface in the middle of a perfectly clear, dry day at a few thousand meters above sea level. Solar energy transmitted through the atmosphere enters the Earth's surface radiation budget as shortwave radiation, and this component will be referred to below as incident shortwave radiation ($S\downarrow$). Some of this is reflected from the surface ($S\uparrow$), and the rest is eventually absorbed. Some of this energy is converted to heat by interacting with molecules of the surface and increasing their kinetic energy. The Earth's surface temperature is raised to its familiar levels primarily by this mechanism.

Given some temperature, the surface itself radiates according to Eq. 3, and loses energy towards the atmosphere primarily in the longwave or thermal infrared region of the spectrum ($L\uparrow$). At the same time, the atmosphere, especially that part of it closest to the surface, is radiating back down toward the surface as a function of its bulk temperature ($L\downarrow$). The net result of these radiative exchanges is called the surface net radiation (R_n) and is defined as:

Eq. 4 $$R_n = S\downarrow - S\uparrow + L\downarrow - L\uparrow$$

This net radiation defines the input to the surface energy balance; it is the amount left over to do work at the surface, for example by evaporating water, transporting heat, or participating in chemical reactions.

2.2 Energy balance

The energy associated with the net radiation from the radiation budget is conserved according to the following balance:

Eq. 5 $$R_n = H + \lambda E + G + e$$

where H is the flux of sensible heat between the surface and the atmosphere (sensible heat being the energy associated with molecular kinetic energy), λ is the latent heat of vaporization for water, E is the transport of evaporated water vapor between the surface and the atmosphere, G is the transport of heat between the surface and the interior of the surface (soil warming, for example), and e is an extra term associated with other minor energy components, such as the photochemical reactions taking place in leaves, which do not generally have much impact on the surface energy balance. Each term in Eq. 5 can be expressed as an energy flux density (Wm^{-2}). The transport of sensible and latent heat away from the surface, as a result of the net input of radiation from the Sun, is responsible for most of the heat and all of the water content of the Earth's atmosphere (Oke, 1987).

It is common practice to treat the fluxes represented in Eq. 5 as the equivalent of molecular diffusion processes, and to formalize the dependence of these fluxes on a relevant concentration gradient with an equation of the form:

Eq. 6 $$J_i = \frac{D_i}{\Delta x}(c_{i1} - c_{i2}) \text{ --or-- } J_i = \frac{K_i}{\Delta x}(c_{i1} - c_{i2})$$

where J_i is a flux of species i, c_{i1} and c_{i2} are concentrations of species i at two points, Δx is the distance between these two points and perpendicular to the concentration gradient, D_i (m^2s^{-1}) is a constant, called the diffusivity of species i (a function of the properties of the material and the medium through which it is diffusing), and K_i (m^2s^{-1}) is the eddy diffusivity for species i. The two different forms represent the diffusion process in laminar flow regimes (on the left) and the idealized process for turbulent regimes (on the right). Values of D_i can be determined directly through experiment and are relatively constant for constant temperature, but values for K_i depend on the scale of the turbulent processes and are variable in space and time. As such, they represent an abstraction or simplification of a complex process. They may be difficult to measure, and their use may or may not be appropriate, depending on the particular situation. I introduce the concept here because extensive use is made of this abstraction in micrometeorological studies and in the literature relating multispectral remote sensing information to surface flux and resistance estimation. A **resistance** to the transport of species i (r_i, s m^{-1}) is defined in terms of Eq. 6 as $r_i = \Delta x/D_i$ (for laminar flows) or $r_i = \int dx/K_i$ (for turbulent flows, where K_i is varying with distance from the surface) giving the following general form for fluxes:

Eq. 7 $$J_i = \frac{(c_{i1} - c_{i2})}{r_i}$$

Fluxes of sensible heat (H, W m^{-2}) can be expressed in terms of Eq. 7, where the relevant concentration gradient for heat is related to the temperature difference, resulting in:

Eq. 8 $$H = \frac{\rho C_p (T_s - T_a)}{r_h}$$

where ρ is the density of air (kg m^{-3}), C_p is the heat capacity of air (J $kg^{-1}K^{-1}$), T_s and T_a are the temperatures of the surface and the overlying air (K), and r_h is the resistance to sensible heat transfer across the distance for which the temperature gradient is measured (s m^{-1}). This resistance, r_h, is also commonly referred to as the aerodynamic resistance, since it is largely a function of wind speed and the shape of individual "roughness elements" on the surface (leaves, for example). Under most circumstances, the use of r_h to represent the turbulent transfer processes is a gross abstraction, particularly when considering time

and space scales beneath the characteristic scale of the turbulent processes. With long time periods and large spatial scales the abstraction becomes more plausible, but also much more difficult to measure directly (Crago, 1996). This state of affairs represents one of the most challenging problems facing energy budget research at large spatial scales (Carlson et al., 1995b).

Fluxes of momentum are also typically handled in the form of Eq. 7, and another separate budget, the conservation of momentum, can be invoked to treat this quantity. It is mentioned in the current context because it has been widely used in micrometeorological studies as a tool with which to derive the aerodynamic resistance, r_h, at field scales. The diffusion form for momentum flux is:

Eq. 9 $$\tau = \frac{\nu\rho}{\Delta x}(u_1 - u_2) \quad \text{--or--} \quad \tau = \frac{\rho(u_1 - u_2)}{r_m}$$

where τ is the momentum flux density, or shear stress (kg m^{-1} s^{-2}, or Pa), ν is the momentum diffusivity, or kinematic viscosity ($m^2 s^{-1}$), u_1 and u_2 are wind speeds (m s^{-1}) that define the gradient in momentum between two points, and r_m is the resistance to momentum flux between these points (s m^{-1}).

The flux of water across the surface due to evaporation (E) is the principle link between the energy and mass balances at the surface, since a large amount of energy is associated with the conversion of liquid water to water vapor. This energy, equal to the product of E and the latent heat of vaporization (λ), is transported in the gas phase as chemical potential energy and is released again as kinetic energy if and when the vapor condenses. The mass form of the water vapor flux equation is:

Eq. 10 $$E = \frac{(\rho_{v1} - \rho_{v2})}{r_v}$$

where ρ_{v1} and ρ_{v2} are the vapor densities (or absolute humidities, kg m^{-3}) at two points defining the mass concentration gradient for water vapor, and r_v is the resistance to water vapor transport (s m^{-1}). Eq. 10 can take several other forms by specifying the concentration gradient in terms of a partial pressure. An approximate form that is very useful in mass and energy balance calculations is:

Eq. 11 $$E \approx \frac{\rho_a M_w}{P M_a} \frac{(e_1 - e_2)}{r_v}$$

where ρ_a is the density of dry air (kg m^{-3}), M_w and M_a are the molecular weights of water and dry air (kg mol^{-1}), e_1 and e_2 are the partial pressures (Pa) of water vapor at two points, defining the concentration gradient in vapor pressure, and P is the total atmospheric pressure (Pa). The approximation involved in this form is the assumption that P-e can be simplified to P, since e is very much smaller than P (typical values of e and P give errors in E of ~0.5%). Eqs. 10 and 11 are converted to their energy equivalent forms by multiplying on both sides by λ.

I will briefly present the theory of similarity that is used to relate the flux resistances for sensible heat, latent heat, and momentum in micrometeorological studies, and which is now commonly applied to fluxes at much larger scales over all sorts of surfaces. The method described here is called the "aerodynamic method" for flux estimation, or the "micrometeorological method" or "Monin-Obukhov similarity theory", or simply "similarity theory". The central assumption is that the eddy diffusivities (and by definition the resistances) for momentum, sensible heat, and latent heat flux are all equivalent, or at least related in a simple and consistent way. A second assumption is that the mean wind profile over the surface (variation of wind speed with height) is approximately logarithmic, with wind speed increasing rapidly near the surface and more slowly with height. The simplest form for this relationship is suggested to be:

Eq. 12 $$u_z = \frac{u^*}{k} \ln(z/z_0)$$

where u_z is the wind speed at height z above the ground, u^* is called the friction velocity, k is a dimensionless constant (0.41, the von Karman constant), and z_0 is the height at which the logarithmic wind speed profile extrapolates to 0 (called the roughness length), which is typically assumed to be constant under varying wind speeds for a given canopy. From Eqs. 9 and 12 an expression can be derived for the resistance to momentum transfer between the ground (reference plane) and some height z which is a function only of z, the wind speed at z (u_z), and the roughness length (z_0):

Eq. 13 $$r_m = \frac{[\ln(z/z_0)]^2}{k^2 u_z}$$

The usefulness of this formula derives from the relative ease with which z_0 can be measured at the field scale: a series of wind speed measurements are made over a range of heights above the canopy, and the average results are plotted to determine the y-intercept of the logarithmic line, which is the roughness length, z_0. Afterwards, only one measurement of wind speed at some height z is required to estimate the resistance to momentum transfer from the surface up to that height, and using one more wind speed measurement and Eq. 9 the actual momentum flux can be calculated. The object of these measurements, however, is not usually to estimate momentum resistances and fluxes, but rather to employ the basic premise of the similarity theory to infer from wind speed measurements the resistances to sensible and latent heat fluxes.

There are two important limitations to the application of this theory for flux and resistance estimates. The first, already mentioned, is that it assumes the eddy diffusivities for momentum, heat, and water vapor are equal. In fact, the diffusivity for momentum, from which the others are obtained, is known to be higher in most kinds of canopies than those for heat or water vapor, due to the influence of canopy elements on local air pressure fields (the "bluff body effect", (Jones, 1992, p. 302)). Various approaches have been used to correct for this problem, and some of them will be reviewed in a later section. The second limitation is that the logarithmic profile case is only valid under conditions of neutral stability, in which the buoyant forces and the gravitational forces acting on air at the surface are balanced. The exact specification of these conditions requires estimates of the rate of change of both wind speed and some other variable (for instance temperature) with height in the planetary boundary layer (Iribane and Godson, 1981). The atmosphere is in general not in a neutral stability state, and so further empirical corrections are required, with the further complication that the correction factors themselves do not behave similarly for momentum and the other fluxes (Oke, 1987, p. 382)

3. Review of relevant literature

In this section I present a review of literature concerned with the use of surface temperature estimates in studies of the land-surface energy budget. I give particular attention to studies addressing the

influence of vegetation on the surface fluxes of sensible and latent heat, and to studies incorporating radiometric data obtained by satellite remote sensing instruments. The use of remotely sensed radiometric data to estimate land-surface fluxes of energy and mass (and the resistances to such fluxes) has evolved along multiple interacting trajectories. Here I will follow an approximately chronological progression of ideas.

3.1 Early investigations, pre-1970

The idea that radiometric data can be related to storage and transport of energy and mass at the Earth's surface has its origins in agricultural research. Early investigators, using first mercury thermometers then later bimetallic thermocouples to measure the temperatures of single leaves or individual plants, examined the relationship between transpiration and leaf-air temperature differences (e.g. Clum, 1926a; Clum, 1926b; Curtis, 1936), and identified incident radiation, humidity, and wind speed as important variables influencing this relationship. Later, these qualitative relationships were refined through careful measurements of the leaf radiation budget (Gates, 1964) and extended from single leaves to entire canopies by considering the variation of radiation with depth and the aerodynamic and thermal effects of neighboring plants (Idso and Baker, 1967).

This research proceeded mostly in the field of agriculture as methods were sought to specify irrigation schedules that were based on plant characteristics as opposed to the more common methods that relied either on calendar scheduling or on crude meteorological indices. It is interesting to relate these developments to the expansion of dryland agriculture and the increasing demands being placed on limited irrigation water resources. It has been observed that by the early 1900's most of the world's choice (rain-fed) agricultural land was already in production (Glantz, 1994), and while advances in agricultural technology after 1900 have resulted in enormous increases in the productivity of that land, population growth has mandated the development of increasingly marginal agricultural lands which require intensive irrigation to maintain productivity (Beaumont, 1989). As the limited nature of the arable land resource became more evident, and as the energy cost of development and transportation increased, more attention was focused on ways to schedule the use of irrigation water to derive maximum productive gain from each

unit of water expended (Jackson, 1982). We are just beginning to examine some of the less obvious costs associated with water resource allocation, such as the dependence of aquatic biota on flow variation (Poff and Ward, 1989), and of course human population growth continues. The practical implications of such investigations are therefore of increasing importance.

The development of instruments that could record emitted thermal radiation made it possible to extend temperature measurements from individual leaves to entire plant canopies. It was suggested by Monteith and Szeicz (1962, cited in Jackson, 1982), reasoning from a theoretical model of the canopy energy budget, that the recorded surface temperature could be related to the stomatal resistance of field crops. Wiegand and Namken (1966) showed the sensitivity of leaf-air temperature differences in a cotton crop to variation in leaf turgor and incident radiation using a wide spectrum thermal infrared radiometer. Their paper also illustrates the cumbersome nature of the early instruments. At that time aircraft-mounted scanning thermal radiometers were just being declassified and made available for civilian use, and truly portable field radiometers measuring in the thermal infrared were not available until the early 1970's (Hsiao, 1973). In an early study of the application of aerial thermal remote sensing over forests, Weaver et al. (1969) reported that midday thermal imagery was superior to nighttime imagery for distinguishing different forest species on the basis of their canopy temperatures, and showed that midday canopy temperatures were higher for a water-stressed conifer stand than for an adjacent well-watered broadleaf stand. The relative contributions of canopy and underlying soil and litter were not assessed.

The development of early thermal radiometric instruments and their use in the explicit study of surface energy and mass fluxes coincided with an exploration of reflected visible and infrared spectra of vegetation that could best be described as implicitly concerned with questions of fluxes and resistances. By the early 1900's, the importance of photosynthetic pigments in shaping the reflectance spectrum of green vegetation in visible wavelengths was accepted (Knipling, 1969). As early as 1930, water absorption bands in the mid-IR region had been noted and their usefulness in vegetation water status addressed, (Forsythe and Christison, 1930; cited in Hoffer and Johannsen, 1969), and by the mid 1960's,

photographic infrared techniques were being widely used to record reflectances of vegetation in the near-IR region (Knipling, 1969).

Hoffer and Johannsen(1969) suggested the potential use of one or more spectral regions for the detection of plant water stress, and Jordan (1969) showed how the ratio of near-IR:red transmittance under a forest canopy could be used to estimate biomass. Early speculation that "mesophyll collapse" during wilting would result in notable reductions of near- and mid-IR reflectances were shown to be inaccurate, and it was established that infrared reflectance only decreases under severe desiccation (Knipling, 1969). Multispectral photographic techniques using blue, green, red, and near-infrared filtering and compact housings with motion correction for aircraft deployment were in use by 1967. These were supplemented with optical false-color compositing systems that allowed detailed and accurate land-use classifications, as well as the interpretation of varying degrees of crop development and water stress (Yost and Wenderoth, 1967). These early reflectance studies lacked both a theory and adequate instrumentation for absolute calibrations, and so were mostly concerned with empirical relationships. The impetus for the research was much the same as discussed for thermal studies above, with the additional promise of spectral indices that would allow areal biomass surveys.

3.2 Current literature, 1970-present

3.2.1 Reflectance data (visible and near-IR) for surface resistance and flux estimation

Tucker (1979) provides a review of the use of red and photographic (near) infrared combinations for vegetation monitoring, and documents the introduction, in the early 1970's, of both mechanistic and stochastic canopy radiation models that accounted for the influence of solar geometry on reflectances observed from an arbitrary angle (e.g., the model of Suits, 1972). These models were able to corroborate some of the results of observations and were typically applicable for cases of uniform land-use and flat terrain. Many variations on the simple ratio of infrared to red reflectance had already been proposed by the time of Tucker's review, and others have been proposed since. These variants are mostly concerned with reducing the influence of atmospheric path radiance (light multiply scattered in the atmosphere and making its way to the sensor without interacting with the surface), atmospheric transmissivity effects due to

aerosols and water vapor, and the confounding influence of soil background in applications where the vegetation is the target of interest. I will deal here only with the normalized difference vegetation index (NDVI) defined as (NIR-red)/(NIR+red), where NIR is the near-IR reflectance and red is the red reflectance, as recorded at the sensor, and reflectances are measured as proportions of incident radiation.

Jackson et al. (1983) related red and near-IR indices to water stress in a wheat crop. Hunt et al. (1987) proposed an index using near-IR and mid-IR data from the Landsat Thematic Mapper instrument that they found to be related to the relative water content of leaves, a measure of water stress. An important advance in the use of spectral reflectance indices for quantitative studies of surface fluxes was the demonstration by Asrar et al., (1984) of a nearly linear relationship between the NDVI and the fraction of photosynthetically active radiation (FPAR) absorbed by a wheat canopy. These results were supported by a detailed radiative transfer model (Choudhury, 1987), and it was demonstrated that the NDVI-FPAR relationship would be sensitive to background reflection properties (Asrar et al., 1992). The connection between NDVI and FPAR is relevant to surface resistance because radiation absorbed by leaves is one of the most important variables influencing the stomatal resistance (Jarvis, 1976), the term that dominates the control of evaporation in most vegetation types (Hall et al., 1991). This connection between reflectance and stomatal resistance was subsequently given explicit attention by a number of investigators who tested it against both radiative transfer models (e.g., Myneni and Ganapol, 1992), and field measurements (Verma et al., 1993). A recent development is the idea that vegetation indices can be used to parameterize the ground heat flux component (G) of the surface energy balance (Kustas et al., 1993). This component can be measured in the field with heat flux plates buried at various depths in the soil, but few other methods are available or have been proposed which are operable over large regions. The method proposed by Kustas et al. (1993) relies on the observation that, for a long part of the daylight period, the ratio R_n/G is approximately constant, and on the assumption that this ratio is controlled by the fraction of R_n reaching the soil. Vegetation indices that are sensitive to fractional canopy cover or leaf area index should therefore also be good predictors of this ratio.

3.2.2 Thermal infrared emission data: explicit energy balance formulations for surface resistance and flux

Building on a well-developed theory of radiative and convective energy transfer at terrestrial surfaces (e.g., Gates, 1964; Monteith and Szeicz, 1962), a large body of literature has grown around the notion that remotely sensed thermal imagery can be used to estimate various components of the energy balance over large regions. The bulk of these studies rely on the application of semi-empirical energy balance formulations developed in micrometeorological research for the estimation of local flux components. So, as an early example, Heilman et al. (1976) proposed a method for regional ET estimation that relied on surface temperature measurement from thermal scanner imaging instrumentation flown on an aircraft. Typical of these explicit energy budget studies, the proposed method required that a rather extensive set of ground-based meteorological measurements be collected at the time of thermal imaging, including net radiation (R_n), soil heat flux (G), air temperature and wind speed profiles over the surface, and atmospheric water vapor content (for correcting thermal emission data). Crop emissivities were estimated using infrared thermometers. The resistance to sensible heat transfer was determined from a form of the similarity theory similar to that shown in Eqs. 12 and 13. Sensible heat flux was calculated from Eq. 8, and latent heat flux (evapotranspiration) was calculated as the residual of the other three energy balance terms ($\lambda E = R_n - H - G$).

The problems cited by Heilman et al. (1976) are characteristic of many such investigations, and they deserve mention here. First, the calibration of scanner radiance to actual surface temperature was difficult. Although radiation received by the scanner was compared against two internal calibration standards set at temperatures that bounded the likely range of canopy temperature, the influence of water vapor and other atmospheric components, such as CO_2 and aerosols, compounded the calibration problem. Calibration to thermocouples installed on or in the leaves was required for accuracy within 1.5 °C. Sampling strategy was a problem, since the variation in single leaf temperatures was large, due to variable radiation environments in the canopy. Second, the estimation of resistance to sensible heat transport away from the canopy and into the boundary layer (r_h) relied on empirical relationships represented by Eqs. 12

and 13. In addition to any fundamental inadequacies of the similarity theory, measurement error for wind speed and air temperature profiles impacted the determination of r_h. Third, measurements of R_n and G depended on constant radiation inputs in order to avoid error due to transient temperature changes, produced, for example, by the passage of clouds. Similarly, the timing of overflight had to correspond to a period of steady radiation conditions. Fourth, a field with complete cover had to be selected to avoid the influence of a mixture of soil and vegetation temperatures being observed by the scanner. All of the formulas become more complicated in the case of mixed surface types, and opportunity for error increases. Finally, all of the errors associated with estimating R_n, G, and H are carried to the latent heat estimate, since it is calculated as a residual. Heilman et al. (1976) reported errors in ET estimates ranging from +62.5% to -43.6%, with a mean absolute error of ~26%, as estimated from lysimeter measurements. It should be noted that this method provides an instantaneous estimate of ET flux, but not a cumulative amount.

Soer (1980) describes a similar method for estimation of ET, and includes a comparison between surface temperatures measured remotely and predicted by a soil-plant-atmosphere water and energy balance model. He shows that the model generates crop surface temperatures that agree well with those detected from an aircraft-mounted infrared scanner, and that the estimates of ET from the surface temperature observations agreed with surface water balance measurements to within 30%. He also included an error propagation analysis that showed ET errors to be in the range of 12-15% due only to measurement error. Additional unestimated error, related to the formulation of the flux equations, was acknowledged. This study also introduced a more detailed scheme for the treatment of near-surface atmospheric stability corrections to the sensible heat resistance term, and illustrated that large errors are likely in conjunction with the use of the similarity equations if such corrections are not made. Another novel aspect of this study is that a comparison between simulated values of instantaneous and daily total ET was used to scale the ET rates measured with remotely sensed surface temperature (an instantaneous measurement) to their daily equivalent. Some such treatment of this problem is required if infrequent and

instantaneous measurements of surface temperature are to provide quantitative surface water balance information over long periods.

Many other studies have used essentially the same formulation for the energy budget, with the same or similar ground-based meteorological requirements. Moran et al. (1989) used Landsat Thematic Mapper data to compute instantaneous ET fluxes over an agricultural area and compared the results with Bowen-ratio measurements and with aircraft-flown radiometer estimates. They noted that the influence of incomplete canopy cover was significant, that the estimation of surface roughness parameters over large regions was problematic, and that the difficulties in extrapolating ground-based measurements of air temperature and wind speed over complex terrain had not been adequately addressed. Vidal et al. (1994) used data from the Advanced Very High Resolution Radiometer (AVHRR) to estimate instantaneous ET fluxes over an 18-month period for a forested region in France. They generated an index of temporal variation in forest ET, and then related this index to statistics on fire ignitions over the same period. They found a strong relationship between their index and the number of fires, and produced a regional alarm statistic that was based on the remote sensing data and estimated the probability of a fire event on each observation day. Humes et al. (1994) computed energy balance components across a semi-arid watershed in Arizona by scaling estimates from aircraft-based remote thermal data against measured fluxes at one reference site. Estimates were made at a number of other sites, and they found errors to be in the same range as for comparisons between different ground-based measurement systems. They noted that spatial variability in incoming solar radiation was a significant source of error, and that the level of uncertainty attached to large-scale estimates of the resistance to sensible heat flux can create substantial errors in heat flux estimates that propagate to ET estimates.

Kustas et al. (1996) addressed the problem of partial canopy cover in an energy balance model by testing both single-source (canopy and soil treated as one unit) and dual-source (canopy and soil treated as separate units, acting in parallel) models of land surface temperature. The principle assumption in their treatment of canopy and soil temperatures for dual-source modeling was that the canopy and soil behaved in parallel; that is, energy budgets could be calculated for each partition separately, without having to

account for interactions between the two energy budgets. The single-source model is the more usual approach. They found that the dual-source model gave more accurate estimates of ET over both a semiarid and a humid test site. They also found that the single-source model could be parameterized to mimic the dual-source model, but a general scheme for such parameterization was not found.

Another approach to regional resistance and flux estimation has been to combine remotely sensed thermal data with a numerical model of planetary boundary layer (PBL) dynamics (PBL being the lower 0.5 - 2 km of the atmosphere which is turbulently mixed due to daytime surface heating and the drag exerted by the surface on the atmosphere). Taconet et al. (1986) provide a detailed discussion of one such model. A single midday measurement of surface temperature is the only remote sensing input, and other parameters include vegetation height, LAI, resistance to transpiration, and mean rooting-zone soil water content. They demonstrate that, of these parameters, the midday surface temperature is most sensitive to the mean stomatal resistance of the canopy. Their model can be inverted, with a specified surface temperature, to produce estimates of the canopy resistance to transpiration, and they give a likely error range for such estimates of 20 sm^{-1}, or around 10%. Carlson et al. (1991) use a PBL model to investigate the sensitivity of surface temperature in a vegetation canopy to the diurnal variation of water stress in a progressively drying system. Their simulations show that surface temperatures are not sensitive to soil water content until a critical water content is reached, after which the canopy begins to show a midday depression in ET, with a corresponding rise in T_s. As the soil becomes progressively dryer this daily depression becomes deeper and more prolonged. Their results compare favorably with observations of surface temperature and soil water content for a corn crop. The appearance of cloud cover reduced or eliminated the midday ET depression. They suggest that midday thermal IR data may be the most appropriate for the early detection of water stress. Diak and Whipple (1993) present a method by which radiosonde reports are used to assess the diurnal change in the height of the PBL, which together with satellite surface temperature data can be used to predict surface fluxes. They reiterate the needs expressed by Taconet et al. (1986) for both an improved distribution of surface meteorological data and a better way

to assess the land surface characteristics of relevance to sensible and latent heat fluxes at large spatial scales.

Choudhury et al. (1986) used both similarity theory and the Penman-Monteith equation to derive a system of four separate equations that were then solved iteratively to predict ET, canopy resistance, soil water potential and leaf water potential. Using an initial guess for the value of soil water potential, all four equations were solved simultaneously, then predicted canopy temperature was compared with a satellite measurement, and the estimate of soil water potential was adjusted and this process repeated until the two temperatures converged. Their predictions of ET for half-hourly simulations over six days with varying weather and soil water conditions agreed well with observations over both well-watered and water-stressed wheat.

3.2.3 Thermal infrared emission data: empirical relationships with surface resistances and fluxes

Using aircraft-mounted sensors, Bartholic et al. (1972) demonstrated that the thermal structure of the ground in an agricultural area could be determined remotely using longwave infrared detectors in the 8-14 μm range. They recorded differences of up to 6 °C between well-watered and water stressed cotton crop canopies, and noted the strong influence of soil temperature on their results. Stone and Kanemasu (1975), using hand-held radiometers, observed rapid fluctuations in the canopy temperature of a sorghum crop (3 °C in 3 min.) that were the result of clouds passing over the study site. They concluded that care was required in the interpretation of thermal crop studies under conditions of variable radiation. They observed that canopy heating occurred more rapidly than cooling. Idso et al. (1975) also examined the thermal properties of soil using hand-held radiometers, and discovered that the soil water status could be predicted from the difference between night and afternoon soil temperature, but that the relationship was specific to the soil type.

Jackson et al. (1977) and Idso et al. (1977) used surface temperature data for the assessment of plant stress, defining the "stress degree day" (SDD) concept. They showed that crop yield could be predicted by a daily accumulated index of the temperature difference between the air and the crop canopy at about 1 - 1.5 hours after solar noon, where higher values were associated with lower yield. The

midafternoon temperature difference was also shown to be a good predictor, in conjunction with estimates of Rn and G, of daily evapotranspiration (ET) (Jackson et al., 1977), and other studies soon corroborated their results (e.g. Walker and Hatfield, 1979).

A review by Byrne et al. (1979) examined the SDD in the context of an explicit surface energy balance, and considered the influence of fractional vegetation cover on the integrated thermal response at a sensor, offering one of the first thermal mixing models for two components. They also considered the influence of atmospheric water content and the spatial variation of emissivity in fractional canopies on temperature retrievals, and indicated that errors of 1-3 °C were likely to result from these uncertainties. They suggested that the incorporation of other meteorological variables with remotely sensed surface temperature could lead to more accurate predictions of ET. Their review is also the first explicit reference I find to the utility of combining thermal and reflective remote sensing techniques to expand the range of conditions which can be assessed. Idso et al. (1981) soon thereafter incorporated a measure of vapor pressure deficit (VPD) into the original SDD formula, and found that it helped to normalize the parameters obtained from calibrations at two sites with different climates. Jackson et al. (1981) presented a theoretical basis for the inclusion of VPD, using the energy balance equations and similarity theory.

Seguin and Itier (1983) performed a detailed analysis of the ET estimation formula given by Jackson et al. (1977), examining the theoretical basis for the application of an instantaneous measurement to a daily integrated flux estimate, with attention to the influences of wind speed, boundary layer stability, and surface roughness on the results of the method. This method is know as the "simple ET" method. They concluded that the extension of one-time-of-day measurements to the entire day was only valid for clear days. They also note the problems with equating the diffusivities for momentum and heat, especially for large roughness lengths (> 1 cm), which result in overestimation of fluxes. They commented that although corrections for crops were available, the cases of forest and other very rough canopies had not been resolved. They found the likely errors to be in the range of 20% for estimates of ET, acceptably close to the uncertainties obtained by ground-based methods. They further suggested that for the purpose of calibration, any satellite study would require at least one ground reference temperature.

Carlson and Buffum (1989) contrasted the simple semi-empirical ET formula from Jackson et al. (1977) with other more complex and mechanistic methods, observing that the initialization requirements for the complex models are often prohibitive. They demonstrated the application of a sophisticated PBL model in the estimation of the parameters for the simpler models proposed by Jackson et al. (1977) and Seguin and Itier (1983). Their purpose was to generate a range of meteorological conditions with the PBL model from which the parameters for the simpler model could be determined. They found that their model generated coefficients very close to those estimated by the original authors, but that, in a separate field comparison, small errors in either measurement or model initialization could result in unacceptable errors in daily ET predictions. They noted that, given the sources of error inherent in the simple method, it may be acceptable to use regional-scale surface winds and air temperatures as inputs, and to stratify stability and roughness into two or three categories.

Thunnissen and Nieuwenhuis (1990) proposed a similar simple model, but instead of referencing ET to the difference between remotely sensed crop temperature and ground-based air temperature they used the idea of a reference crop in the thermal IR image, which was taken to be transpiring at its potential rate. Temperature differences were then determined between the rest of the image and the reference field. Potential ET was calculated from the usual ground-based meteorological measurements and then used to scale the relative thermal differences to absolute ET amounts. This method has the advantage of not being sensitive to bias between the ground-based measurements of air temperature and the remote measurements of canopy temperature. They found that large errors were possible in fields where a lot of soil showed beneath the canopy, a problem shared by most other methods. They note that in the absence of ground-based measurements of potential ET rates for the reference field it is still possible to produce a map of ET differences from some arbitrary reference.

In two papers, Seguin et al. (1991 and 1994) further developed the SDD concept, applying it over very large regions using satellite data ranging in resolution from 1 km to 5 km. They redefined the temperature difference to use the daily maximum temperature, regardless of exact coincidence with satellite overpass, and they used the split-window technique (e.g., Ottlé and Vidal-Madjar, 1992) to

estimate the surface temperature from satellite data. They also proposed a normalized SDD, with a summation of temperature differences divided by R_n, as a regional water stress indicator, and argued that it is more physically associated with ET than the integrated NDVI, and therefore deserves more attention as an indicator of productivity, at least in crops. In the later paper they demonstrated the application of the method with AVHRR data, illustrating that estimates can be made once every 5 or ten days with reasonable accuracy, by using the 5- or 10-day maximum air and surface temperatures, from meteorological stations and satellites, respectively. They noted that their results were best when averaged at a scale of about 10 km, and suggested that for intermediate resolutions, down to about 100 m, the accuracy suffers from the influence of heterogeneous land-use.

Carlson et al. (1995a) have shown a simple relationship between a scaled version of the NDVI and the parameters for the simple ET method. Using a PBL model to determine how these parameters should vary with respect to wind speed, surface roughness, and fractional vegetation coverage, they predicted that these parameters are more sensitive to NDVI than to either wind speed or surface roughness. This was a surprising result, since the parameters had been considered as surrogates for the resistance to sensible heat, which micrometeorological studies show to be very sensitive to wind speed. It is unclear at this point whether the result is pertinent only to their PBL model, or if it has more general applicability.

3.2.4 Combined thermal infrared and reflectance data for surface resistance and flux estimation

An early example of the integration of shortwave and longwave remote sensing data in the estimation of surface flux is the study by Reginato et al. (1985), in which they discuss the use of remotely sensed shortwave radiances to estimate the shortwave components of the surface radiation budget. One objective of their study was to show how remote sensing measurements could be used to assess the components of the surface radiation budget that were sensitive to surface conditions and would therefore have complex spatial distributions, while using ground-based meteorological measurements to assess the radiation budget components that were not dependent on surface conditions (incoming shortwave and longwave radiation) and would thus vary gradually with distance. Their objective was to maximize the

distance across which ground-based meteorological measurements could be extrapolated, making the best use of a ground based network of observations.

Goward et al.(1985) were the first to examine how thermal emissions are related to reflective radiances over diverse vegetated land areas. They noticed a strong negative relationship between greenness and thermal counts, relating data from two different sensor systems, and concluded that latent heat exchanges over the vegetation were controlling surface radiant temperature. They speculated that the relationship might be diagnostic of aerodynamic parameters and surface resistances to moisture flux, and that if so it would be influenced by available soil water. The relationship was explored further by Hope et al. (1986) who used a canopy reflectance and energy balance model to simulate the relationships between spectral vegetation index and the canopy radiant temperature over a range of leaf area. These predictions compared well with satellite observations, but they cited the need to consider soil background influences and to compare the model results with other sources of remote sensing observations and ET measurements. Hope (1988) further developed this modeling approach by simulating minimum canopy resistance and the related minimum canopy temperature for a variety of meteorological conditions, and then assessing an unknown canopy resistance on the basis of a near-linear relationship between the ratios of minimum temperature to observed temperature and minimum resistance to actual resistance. Scatter in the low resistance-low temperature range was large, indicating that the accuracy of the method was poor for cases with the highest potential fluxes.

Another approach making use of thermal and reflective remote sensing data to predict surface fluxes is provided by Pierce and Congalton (1988). They used Landsat Thematic Mapper thermal data to determine sensible heat flux component of a surface energy balance, and incorporated visible and near-IR information from the same sensor to perform a classification of vegetation over a mountainous watershed in California. They introduced a suite of ground-based meteorological measurements in their method. They concluded that their method gave reasonable trajectories of sensible and latent flux for coniferous forest, but that more information was required to parameterize the response of vegetation to changes in wind speed, which they found to be significant.

Nemani and Running (1989) tested the sensitivity of the relationship between surface temperature and canopy density to variation in surface resistance. Surface temperature was obtained from AVHRR thermal channels by a split window method, canopy density was estimated from the AVHRR NDVI, and ground-based meteorological measurements were incorporated in a water and energy budget model to generate predicted canopy resistance. The experiment was carried out over a 20x25 pixel (~ 20x25 km) area of conifer forest in Montana. The slope of the T_s-NDVI curve was assessed from eight different midday images through the growing season, and they found a strong correlation between that slope and the independently simulated canopy resistance. The seasonal pattern in both relationships showed the influence of mid-summer drought, with high resistances correlated with low slopes for the T_s-NDVI curve. They concluded that latent heat exchange was the major cause of spatial variation in the surface radiant temperatures, and that the T_s-NDVI slope was a qualitatively useful tool in regional ET research. They did not assess the influence of fractional vegetation cover on their results.

Carlson et al. (1990) used the T_s-NDVI slope and a PBL model to estimate fractional vegetation cover and soil water content in the surface and rooting zones. They assumed that the primary control on the T_s-NDVI relationship for a given point in time was variation in fractional vegetation cover, and the resulting mixture of temperature signals from sunlit vegetation and sunlit bare soil. They further assumed that flux estimation from bare soil and vegetation mixtures could be treated in parallel, without having to account for the interaction of adjacent vegetation clumps and bare soil patches, (Shuttleworth and Wallace, 1985). By using a plot of temperature vs. temperature variance they defined the endpoints of the soil and vegetation mixture, and then adjusted the soil water content in their PBL model for individual cases to reproduce observed slopes in the relationship between NDVI (which they interpreted as fractional cover) and T_s. They used data with a high spatial resolution collected from an aircraft platform for their study, and were cautious about the extension to low resolution data (AVHRR), since it would be more difficult to obtain the complete range of fractional cover in any one cell. They suggested that AVHRR data could be a useful way to test the variation in these relationships over a wide variety of surface types and climates.

Kustas et al. (1990) expanded on the development of the simple daily ET formula (Seguin and Itier, 1983) and the reference field concept (Thunnissen and Nieuwenhuis, 1990) by incorporating visible and near-IR reflectances to calculate surface albedos in the radiation budget, and compared the results with more complex simulations using similarity theory. They found that extrapolations from a reference field were limited by the spatial variability of surface meteorological parameters, that the extension of instantaneous flux estimates to the entire day was valid under clear sky conditions, and that the problems of partial canopy cover were very difficult to address with a one-layer canopy model. Hall et al. (1991) reported much less positive results from experiments over a Kansas prairie during an intensive field measurement campaign, reporting large overestimates of sensible heat flux from measurements of T_s. They note that typical values for r_h are very small, and so small errors in temperature are translated into large errors in fluxes. On the other hand they point out that the stomatal resistances are much larger, making estimates of evapotranspiration more reliable than estimates of sensible heat flux. They noted the good agreement of the simple ET method for daily flux estimation with observations, and illustrated that the ratio between evaporation and net radiation is relatively constant for their site on a sunny day. Carlson et al. (1995a) later criticized the theoretical treatment of similarity in these experiments, and noted that with proper treatment of the difference between momentum and sensible heat flux resistances the agreement between observation and prediction was improved. Sandholt and Andersen (1993) presented a method for interpolating daily ET values in between satellite observations by the definition of a drying curve (Priestley and Taylor, 1972), which is based on the relationship between potential ET, actual ET, and accumulated precipitation. They reported good results when comparing these interpolated estimates to field measurements.

Smith and Choudhury (1991), using Landsat TM data, investigated the T_s-NDVI relationship over crop, bare soil, and forest, and found an inverse-linear relationship for crops but not for forests. These results were corroborated by a two-layer surface heat balance model. Their model also predicted that the variation in slope of the T_s-NDVI line was in opposite directions for decreases in transpiration and for decreases in soil evaporation: their assessment of the usefulness of a simple relationship between the T_s-

NDVI slope and water stress over diverse landscapes was generally pessimistic. Nemani et al. (1993) went on to show, using AVHRR data, that grass, crop, alpine shrub, and forest biomes all had an inverse-linear relationship of T_s to NDVI, and that this relationship could be extracted from a cloud or aerosol contaminated scene using an automated procedure. They illustrated the likely range of variation of the T_s-NDVI slope under extreme conditions of both evaporation and transpiration, and their interpretations of observed relationships agreed with Smith and Choudhury's model results (1991) on this point. They concluded that fractional vegetation cover is the dominant feature controlling the spatial variation of surface temperature, as opposed to an earlier hypothesis that the variation was controlled by differences in transpiration (Nemani and Running, 1989). A strong linear relationship was demonstrated between the T_s-NDVI slope and negative values of the Crop Moisture Index (generated weekly by the National Weather Service) for two bi-weekly AVHRR composites over 20 climatic zones.

Moran et al. (1994) describe a method using both a spectral vegetation index and the surface-air temperature difference that integrates the observed relationship between T_s and NDVI (Goward et al., 1985; Nemani et al., 1993; Nemani and Running, 1989) with an explicit treatment of the surface energy balance. Their objective was to provide a means by which vegetation water stress could be assessed over partial canopies without resorting to intensive micrometeorological measurements. The ground-based meteorological data required by their approach are near-surface air temperature, vapor pressure, wind speed, and net radiation. They used this information in combination with remotely sensed surface temperature and spectral vegetation index to assess the limits to surface-air temperature differences under conditions of maximum and minimum evaporation and transpiration, producing a graphical device they call the vegetation index/temperature trapezoid. Their method assumes a knowledge of LAI, maximum leaf level stomatal conductance, maximum and minimum values for the spectral vegetation index corresponding to complete and absent canopies, and some canopy-dependent parameters for the assessment of aerodynamic resistance. They further assume that one value of R_n is sufficient for both vegetated and bare soil cases, that simple relationships can be derived for aerodynamic resistance as a function of wind speed, and that soil heat flux can be estimated as a function of R_n and fractional cover. Given these

conditions, any pair of temperature difference and vegetation index values defines a point within the trapezoid, and further simple relationships can be applied to derive the possible ranges for the combinations of canopy and soil temperatures that make up the mixed temperature signal. A second measurement, of bare soil temperature for example, uniquely defines the canopy temperature in the mixture. They demonstrate the error associated with a linear mixing model of canopy and soil temperatures, and suggest ways to eliminate this error with a dual-source model that considers radiant and convective exchange of heat between the soil and the vegetation.

In a recent paper (Moran et al., 1996) this logic is applied over a region (~150 km x 100 km) using Landsat Thematic Mapper data to derive the required vegetation index and surface temperature estimates. Simple interpolations of near-surface air temperature, wind speed, and humidity observations provided the necessary meteorological data . Solar input was assumed constant over the region. Using eddy flux correlation equipment installed on one site within the region, they estimated the error for predicted ET to be about 15%.

3.3 Summary of literature review with respect to the current research problem

The first conclusion I draw from the foregoing review is that, while many authors have highlighted the importance of accurate spatial distributions of surface meteorological parameters in studies of the relationship between remotely sensed surface temperature and various aspects of the surface energy balance, no adequate effort has yet been made to generate and use such information for study areas larger than a single agricultural field or for periods longer than a single growing season (Moran et al. 1996, Carlson et al. 1995b, Humes et al. 1994, Seguin et al. 1994 Diak and Whipple 1993, Nemani et al. 1993, Kustas et al. 1990, Moran et al. 1989). Having, in Chapter 2, developed a collection of methods which begins to satisfy the needs outlined by these authors, I set out in the current study to implement these methods, and to augment them where necessary to better match the spatial and temporal characteristics of the remote sensing data. The surface meteorological parameters suggested to be of general importance in explaining the observed variation in surface temperature are, in approximate order of importance, incident shortwave radiation, air temperature, vapor pressure deficit, wind speed, and some measure of antecedent

precipitation. Of these, all but wind speed are parameters that are explicitly treated by the methods in Chapter 2. There are various possible solutions to the generation of distributed surface wind fields over large regions of complex terrain, but in view of their complexity I have elected to ignore the influence of variable wind speed on surface energy fluxes, assuming a constant speed of 1 m s^{-1} throughout. Because the satellite observations are essentially instantaneous, modifications and additions to the methods described in Chapter 2 are required in order to generate instantaneous estimates of incident radiation, air temperature, and humidity using only the observed values of daily maximum and minimum temperature and total precipitation. These modifications and demonstrations of their effectiveness are presented in the description of methods, below.

My second conclusion from the literature survey is that, while the specification of a resistance to sensible heat flux (r_h) is an important first step in the quantitative explanation of observed surface temperature, the most common methods of determining r_h depend on micrometeorological observations that are impossible to obtain over large areas, especially if those areas include rugged or remote terrain (Kustas et al. 1996, Carlson et al. 1995b, Humes et al. 1994, Moran et al. 1994, Thunnissen and Nieuwenhuis 1990, Moran et al. 1989, Taconet et al. 1986, Soer 1980). There is the additional complication for studies using relatively coarsely resolved observations, such as from the AVHRR instrument, that the heterogeneity of landcover within a single field of view can be great. This brings into question the effectiveness of the theoretical treatment outlined above (Eqs. 5-11) in explaining the observed variation of aggregated surface temperature. If predictions from this theory can be substantiated by observations at coarse spatial resolution and across large domains, there remains the problem of estimating an appropriate value for r_h over a region characterized by sub-resolution heterogeneity.

In combining interpolated and extrapolated surface meteorological fields with a relatively long record of longwave emissions and shortwave reflectances, I see a unique opportunity to test the applicability of the surface energy balance theory embodied by Eqs. 5-11 to large, topographically and floristically diverse landscapes. This theory is also at the foundation of the algorithms describing surface energy dynamics in the general terrestrial ecosystem process model described in Chapter 3, and such a test

is a step toward the estimation of the accuracy of that system of algorithms as a whole. Such validation efforts for regional-scale flux estimation methods have been indicated by many investigators as a research priority (Carlson et al., 1995b; Kustas et al. 1994a, Kustas et al. 1994b, Brutsaert and Sugita 1992, Choudhury, 1991). The use of remotely sensed radiometric data shows promise for the accomplishment of that goal, but these data by their nature introduce many complications, requiring that progress along this avenue be painstaking.

4. Operational statement of the research problem

Eqs. 8 and 11, describing the fluxes of sensible and latent heat, taken together with Eq. 5, describing the form of conservation of energy at the surface, can be solved for the surface temperature, T_s, by introducing the approximation $\Delta e = VPD + s\,(T_s - T_a)$ (i.e., the Penman approximation, where VPD is the vapor pressure deficit (Pa) and s is the slope of the water vapor pressure saturation curve with temperature (Pa $°C^{-1}$)), to give:

Eq. 15 $$T_s = \frac{R_n}{\frac{\rho c_p}{r_h} + \frac{M_w \rho \lambda s}{M_a P r_v}} - \frac{VPD}{\frac{M_a P c_p}{M_w \lambda}\frac{r_v}{r_h} + s} + T_a$$

This equation is very useful because it separates the influence of the three important meteorological parameters (radiation, humidity, and temperature) into three linear terms. This expression shows that the quantity $(T_s - T_a)$ should be more closely related than T_s to meteorological variables having direct interaction effects with the surface. This is so because T_a on the right side of Eq. 15 is a simple additive term, and is not influenced by r_v or r_h, which are the two intrinsic surface variables in the relationship. I therefore rearrange Eq. 15 as follows, where ΔT is the difference between surface and air temperature, and combine physical constants (c_p, M_w, M_a), parameters that are weak functions of temperature (ρ, λ), and parameters that are functions of elevation (P) into new terms, to give:

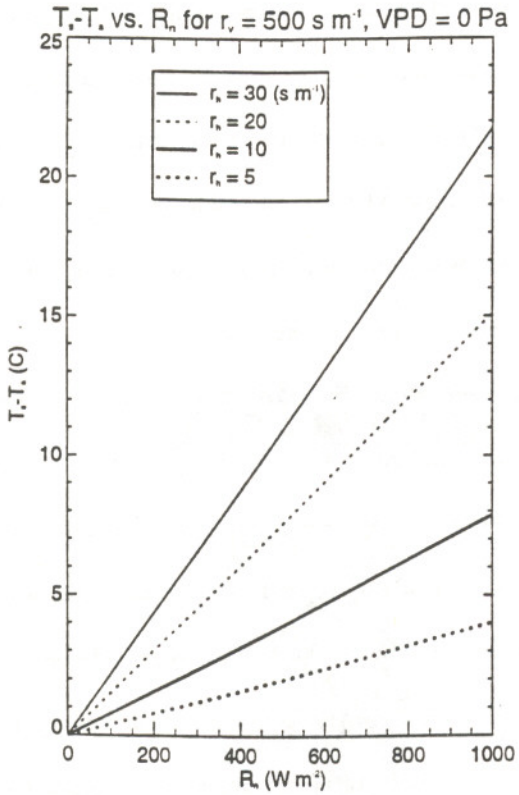
Eq. 16 $$\Delta T = \frac{R_n}{\frac{c_1}{r_h} + \frac{c_2 s}{r_v}} - \frac{VPD}{c_3 \frac{r_v}{r_h} + s}$$

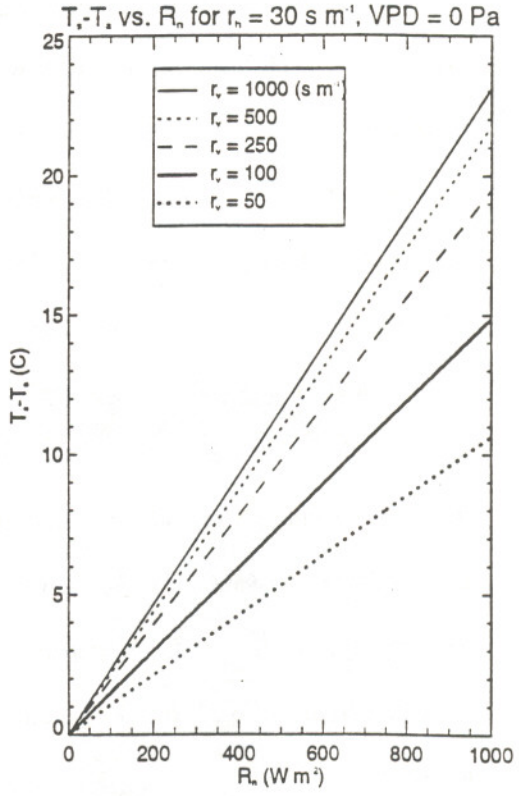
The behavior of Eq. 16 over a range of R_n, VPD, r_h, and r_v is rather complex. In general, an increase in R_n leads to an increase in ΔT, with the magnitude of the increase controlled largely by r_h when r_v is higher than about 500 s m^{-1}. At lower values of r_v the magnitude of the increase in ΔT with increase in R_n is reduced. An increase in VPD is predicted to decrease ΔT for a constant R_n, with a greater influence of VPD on ΔT for lower ratios of r_v/r_h. These relationships are illustrated in Figure 1, which shows the influence of varying R_n on ΔT under a range of conditions for r_h, r_v, and VPD.

In general, r_v will always be higher than or equal to r_h. In the case of evaporation from a water surface, both sensible heat and water vapor must pass through the same boundary layer before mixing with the overlying atmosphere, and $r_h = r_v$. For soil or a leaf, sensible heat is generated by molecular kinetic energy of the surface at the interface between the air and the surface, and moves through the surface boundary layer to mix with the atmosphere. Evaporating water must move across an additional distance, to the surface from the liquid water evaporation front, and must then cross the same boundary layer as does the sensible heat. In this case r_v will be larger than r_h, since the two pathways are in series and their resistances are therefore additive. As the difference between r_h and r_v increases, the denominator in the R_n term on the right side of Eq. 16 becomes progressively dominated by r_h.

For the case of a bare soil, typical values for r_h are in the range 30-60 s m^{-1}, and r_v can range from this lower limit, for a saturated soil, to 1000-10000 s m^{-1} for a field-dry soil, to practically an infinite resistance for exceptionally dry soils (desert sand, for example) (Daamen and Simmons 1996, van de Griend and Owe 1994, Cahnzy and Bruckler 1993). For a typical value of r_h for a soil, the influence of r_v on ΔT (at constant VPD) is negligible for r_v greater than about 500 s m^{-1}.

Typical values of r_h for single leaves range from 6 s m^{-1} for needle-leaf vegetation to about 20 s m^{-1} for grasses, and up to 60 s m^{-1} for very broad-leafed vegetation, banana, for example. The minimum value for r_v, except in the case of intercepted water evaporating from the leaf surface, is determined by the stomatal density and maximum average stomatal aperture, and typical values range from 150 s m^{-1} for crops to 400 s m^{-1} for many tree species (Körner 1995). These would be the values for r_v under ideal field conditions, i.e., plenty of water, available nitrogen, warm temperatures, and saturating radiation loads. For





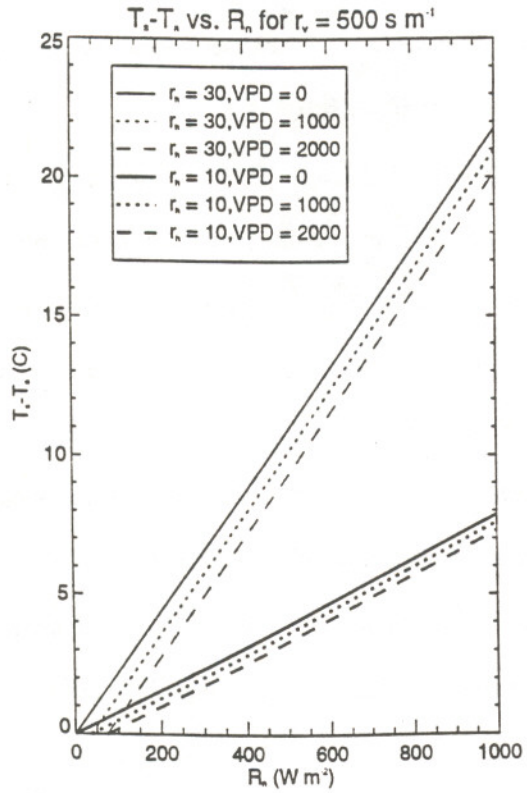


Figure 1
Variation in the difference between surface and air temperature (T_s-T_a) with respect to variation in net radiation (R_n), under a variety of conditions for the sensible heat flux resistance (r_h), the water vapor flux resistance (r_v), and the vapor pressure deficit (VPD), as predicted by Eq. 16.

less favorable conditions, r_v increases due to stomatal closure. With the typically lower values of r_h for vegetation than for soils, the lower limit at which r_v begins to have a significant influence on the R_n term in Eq. 16 is also reduced, and for values of r_v above 400 s m^{-1} there is little influence of r_v on ΔT through the influence of radiation.

Assuming constant values of r_v, r_h, and VPD, the derivative of ΔT with respect to R_n ($d\Delta T/dR_n$) is:

Eq. 17 $$\frac{d\Delta T}{dR_n} \approx \frac{1}{\frac{c_1}{r_h} + \frac{c_2 s}{r_v}}$$

where the expression is approximate because s varies with temperature. Given the observed constraints on the values for r_h and r_v for soil and vegetation surfaces, the denominator on the right side of Eq. 17 can be simplified by assuming that $c_1/r_h >> c_2 s/r_v$, so that $(c_1/r_h) + (c_2 s/r_v) \approx c_1/r_h$. Rearranging to solve for r_h then gives:

Eq. 18 $$r_h \approx c_1 \frac{d\Delta T}{dR_n}$$

In preliminary investigations of the relationship between ΔT and other meteorological variables produced by the methods in Chapter 2, I found that there was a strong positive relationship between ΔT and R_n which had higher slopes for regions with lower NDVI. This preliminary support led me to present the following hypothesis:

> **The variation in surface temperature as derived from observations by the AVHRR instrument, normalized as ΔT by including an estimate of the instantaneous near-surface air temperature, and considered with respect to the variation in R_n as predicted from sun-slope geometry and time of satellite data acquisition, can provide an estimate of r_h, assuming that r_h is not varying greatly over time.**

My method in investigating this hypothesis was to perform a multiple regression analysis of the variation in ΔT with respect to R_n and a number of other important variables, so that I could isolate the observed variation in ΔT that was due to the variation in R_n (i.e., the multiple regression coefficient for R_n, or $d\Delta T/dR_n$).

In testing this hypothesis, I selected as my study area a region in the northwestern United States which spanned seven degrees of latitude and ten degrees of longitude (about 800 km by 800 km). The area was chosen for its combination of diverse terrain, including several major mountain chains as well as a variety of large and small alluvial plains, and diverse climate and vegetation, including hot and dry semi-desert, warm and dry shrubland, cooler and moderately dry grassland and forest, cool and wet forests, and cold and wet alpine environments.

Radiometric observations from the AVHRR instrument, composited over bi-weekly intervals, were converted to surface temperature values by first estimating the spatial variation in spectral emissivity of the land surface and then performing a correction for atmospheric influences. The emissivity estimation and atmospheric correction used multiple spectral bands, an independent estimate of atmospheric water vapor content, and an assessment of the long-wave radiation incident on the surface based on estimates of near-surface temperature and humidity. The spatial resolution of the radiometric observations was 1 km, with an approximate temporal sampling frequency of 14 days. I developed a simple algorithm for the detection of snow and cloud cover in the composited satellite data, and removed grid cells so identified from further analyses. Screening was performed to remove from further analysis all observations made from excessive off-nadir angles. The yearday of compositing was determined for each composited cell. From ancillary sun-sensor geometry data accompanying each composited image, I estimated the local solar time at which each composited cell was observed.

Using the methods described in Chapter 2, I generated estimates of daily maximum and minimum temperatures, precipitation, humidity, radiation, and daylength over a grid that corresponded with the satellite observation grid. From an independent dataset of observed hourly temperatures, I derived an empirical expression which related the time of day of the maximum temperature to a set of independent meteorological variables available in the gridded dataset. Also from the hourly dataset, I derived an expression relating the temperature at a particular time, within the four-hour span before or after the time of the maximum temperature, to the maximum temperature, the minimum temperature, and a collection of other predictors available from the gridded dataset. From these two expressions and a knowledge of the

composite days and observation times for each cell, I predicted the air temperature at the time of satellite observation for each composited cell. I likewise predicted the incident solar radiation at the time of observation using gridded terrain data, established geometric expressions for solar illumination on slopes, and potential atmospheric transmittances, assuming cloud-free observations. Vapor pressure deficit (VPD) was assessed at the time of satellite observation, using the estimated air temperature at the observation time and other variables from the gridded dataset, as described in Chapter 2.

As a preliminary test of the correspondence between surface temperature observations, estimated meteorological fields, and surface energy balance theory, I performed a regression analysis on the surface temperature observations, with estimated fractional cover and estimated meteorological parameters coincident with the observations as independent variables. The objective of this analysis was to compare the variation in surface temperature due to variation in individual meteorological parameters, as predicted by theory, with the coefficients for those same interactions from a multiple regression analysis. Assuming some reasonable qualitative agreement could be found, the next step in the analysis was to use the regression coefficient for R_n to estimate the spatial distribution of r_h as given in Eq. 18. This parameter has been shown to have characteristic values for different vegetation types and for soil, but it is difficult to assess r_h over large spatial areas without an accurate knowledge of the distribution of vegetation types and the fractional vegetation coverage. If the estimated values for r_h correspond well with independent estimates of the vegetation cover type and the fractional vegetation cover, that result would lend weight to the validity of my hypothesis.

Independent estimates of vegetation type and fractional cover were derived from classification work performed over parts of my study area using other satellite sensor data, and from spectral information from other channels on the AVHRR instrument. In the final analysis, I compared the values of r_h predicted from Eq. 18 over the range of classified vegetation cover types to test two hypotheses: that the predicted values of r_h were related to fractional vegetation cover, and that the relationship between r_h and fractional cover was distinct for different landcover types. If these two hypotheses are supported by the data, I will consider it evidence in favor of my overall hypothesis, that the surface energy balance theory presented

here is applicable over the spatial resolution and scale represented by AVHRR data, and that this theory is consistent with observations of surface meteorological variables and surface radiometric characteristics. In addition, if I come away with confidence in the two specific hypotheses just stated, then the observed relationships between r_h and NDVI could be used in future surface energy balance studies to parameterize r_h from a knowledge of covertype and fractional coverage. The investigation of these hypotheses is far from a complete validation of the surface energy balance theory incorporated in the process model described in Chapter 3, but it is an important first step in that direction.

5. Methods

5.1 Study area and topographic data layers

The study area was chosen to lie within the U.S. borders, since topographic and meteorological data were available for the region, and to encompass a diverse range of topography and landcover. The selected region is the entire area between 110 and 120 degrees west longitude, and between 42 and 49 degrees north latitude. The region is approximately 850 km east-west by 830 km north-south. All the spatial data was projected in Lamberts Azimuthal Equal Area (LAZEA) projection, with the center of the projection located in the center of the region. Figure 2 illustrates the topography of the region with some political boundaries for reference. For a general description of the terrain, vegetation, and climate of this region, see the discussion in section 2.7 of Chapter 2, which is for a region of nearly the same geographical extent.

Topographic data was obtained from the United States Geological Survey (USGS) EROS Data Center (EDC), Sioux Falls, SD, in a geographic projection, with spacing along both longitude and latitude of three arc-seconds, or approximately 90 m. A subset of this data was extracted, and the subset was reprojected to the LAZEA projection on a grid with 100m grid-cell spacing, using a nearest-neighbor resampling scheme. A new grid with a grid-cell spacing of 1 km was then generated by taking the average of the 100 cells inside each 1 km^2 region. Through this method, I gain confidence in the sub-grid accuracy of the new 1 km grid that would be unfounded if I had used, for example, the digital elevation data

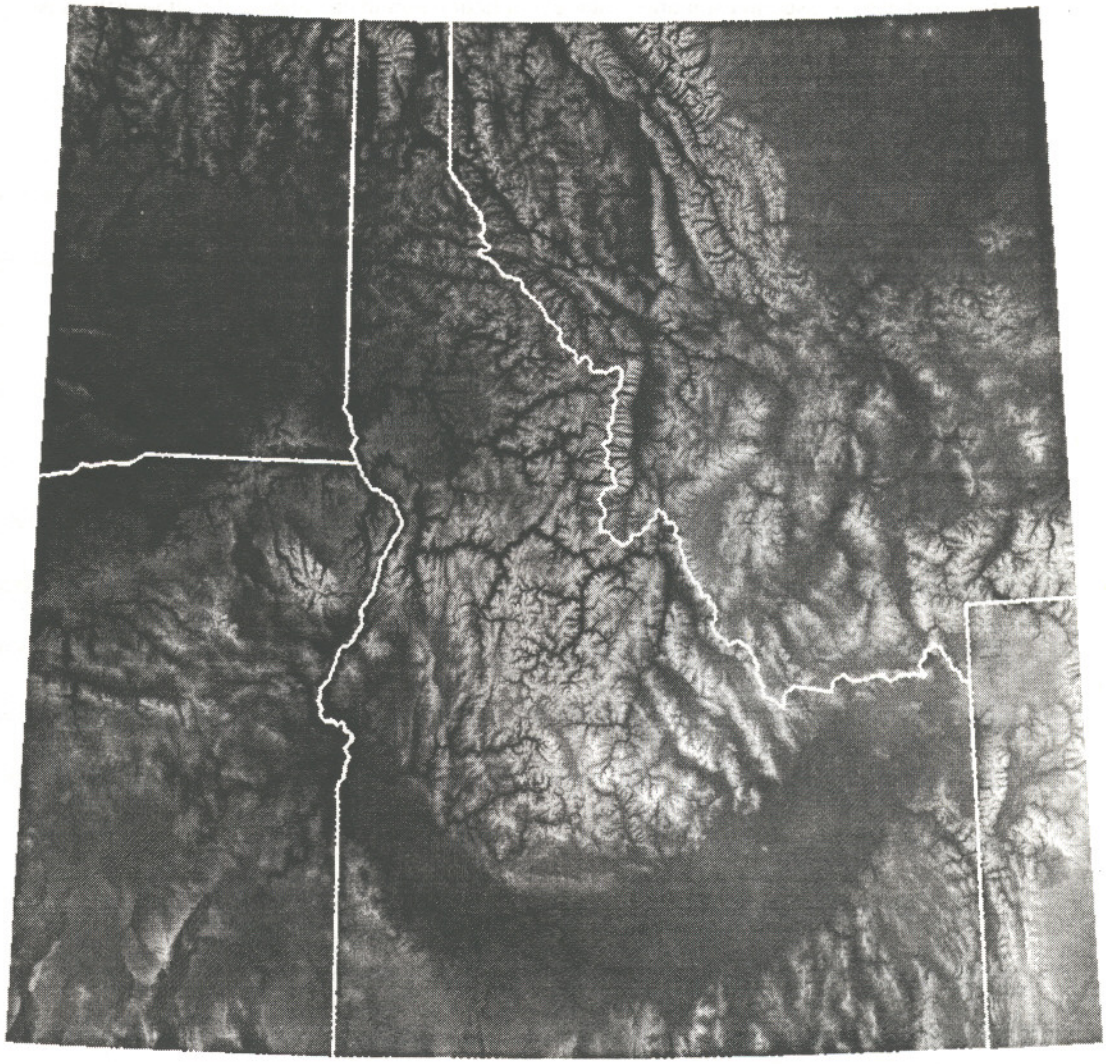

Elevation over study region (m)
500 1000 1500 2000 2500 3000 3500

Figure 2

provided by USGS-EDC as a 1 km grid in conjunction with their AVHRR satellite data. The 1 km elevation grid was employed in conjunction with standard methods to generate coincident grids of terrain slope and aspect, used in estimating incident radiation and in calculating the sun-earth-sensor geometry.

5.2 Estimation of daily surface meteorological parameters

5.2.1 Generation of Daymet input database

The raw sources of observed daily maximum temperature (Tmax), daily minimum temperature (Tmin), and precipitation (Prcp) are the same as described in Chapter 2, i.e., the cooperative network administered by the National Weather Service, and the SNOTEL network of high elevation sites administered by the Natural Resources Conservation Service. These databases were obtained in electronic format. I wrote software to extract all of the station records for a specified geographic range meeting certain quality criteria. A single record here refers to one month of data for a particular station for a particular meteorological parameter. In order to be included at this stage in the database generation, each record in the geographic range corresponding to my study area had to have no more than seven missing daily observations in the month. Missing observations were marked as such in the newly generated data files. Records were extracted for the five-year period from 1990 to 1994, chosen to correspond with the period for which remote sensing data were available (see below).

Because station locations are periodically changed, inducing changes in both horizontal and vertical relationships between stations that bear on the Daymet algorithms, a meta-data record was created for each monthly data record, recording its location as determined from the station history database supplied by the network administrators. Subsequent to extracting all data records, the horizontal station locations for each record were amended through the use of the observed station elevations and the 100 m grid-cell spacing elevation data (described above), for the reasons and according to the methods described in Chapter 2, section 2.7.

In preliminary investigations with this dataset, I found that there were a number of obviously erroneous observations of temperature that were not flagged as such in the original databases. These were found mostly in the SNOTEL data, and were evidenced by rapid changes in temperature or repeated

temperature values for multiple days when no such activity was observed at neighboring stations. Because these errors were substantially above the level expected for typical instrument error, they occasionally "hijacked" the automated parameterization process, resulting in poor parameter choices that reflected the influence of one or several outliers. I identified and eliminated such errors by performing a preliminary cross-validation assessment of daily temperature errors, using a fixed set of parameters taken from Chapter 2, Table 2. Any observation which differed from the cross-validation prediction by more than 10 °C was reassigned as missing and dropped from further analysis. This threshold was set conservatively: I wanted to be careful not to eliminate data which was simply not being predicted well by the interpolation and extrapolation algorithms, but at the same time eliminate as much of the obviously bad data as possible. Through manual observation of the data points so identified, I determined that this threshold eliminated few if any reasonable data points, and passed over a number of obviously bad ones. If this screening resulted in any record exceeding the previously established threshold of seven missing observations per month, the entire record was discarded. The number of observations removed in this way averaged around 75 per variable per year, out of approximately 180,000 observations per variable per year. No such screening method is possible for the precipitation data since, due to the spatial variability of the process being undersampled by the station network, the accuracy of daily precipitation amounts is too low to allow discrimination between bad data and poor predictions. As a result, some bad data are sure to remain undetected in the precipitation observations.

5.2.2 Parameterization of Daymet for the study region and period

In order to implement the Daymet methods, a number of parameters must first be established. A complete discussion of this parameterization process is given in Chapter 2, and here I will focus only on aspects unique to this study. The two most important parameters are the Gaussian shape parameter (GSP) and the average number of stations (ANS) to include in predictions at a given point, which define the behavior of the interpolation algorithm. Also of importance is the spatial smoothing width applied to the topographic data for use in the vertical extrapolation algorithm. In Chapter 2 I showed how the mean absolute prediction error statistic (MAE), generated through a cross-validation procedure, could be related

to GSP and ANS in order to select an optimal combination of these two parameters, with respect to minimizing prediction error. I found that for any value of ANS within a broad range, a single value of GSP could be found to minimize the mean absolute prediction error. I also found that the MAE at this minimum is not much changed over a broad range of ANS, i.e., there is typically a trough of minimized error as opposed to a well. In this study I took advantage of that characteristic of the interpolation algorithm by specifying one value of ANS for all predictions for a given variable and using an automated search procedure to find the value of GSP which minimized the MAE.

In the example application in Chapter 2, I found a single optimal combination of ANS and GSP for each predicted variable and used that combination for predictions on all days and across the entire spatial domain. I noted there that it seemed likely that better results could be obtained by varying the parameterization either spatially or temporally, or both. In the current study I implemented a temporally variable parameterization for GSP, finding the value that minimized MAE at the fixed ANS for each month independently. Cross-validation predictions were made for each observation that was not flagged as missing in the original dataset, for each station in the region, and for all days in a single month. All these prediction errors were averaged to obtain the relevant MAE statistic for selecting the optimal value of GSP for a given month. There was substantial temporal variation in the resulting optimal values of GSP for all three primary variables, as well as temporal variation in the minimized values of MAE (Figure 3). The value of ANS was fixed at 20 for Prcp predictions, and at 30 for both Tmax, and Tmin predictions, following the results in Chapter 2, Table 2. Average MAE and bias were lower in trials using the temporally-varying parameterization scheme than in trials using a single annual optimization for most variables, with the exception of a significant increase in the magnitude of the estimation bias for Tmax and Tmin under the variable parameterization scheme. Differences were small but highly significant in all cases (Table 1), with the greatest improvements from the temporally-varying parameterization coming in the winter months for Prcp and in the summer months for both Tmax and Tmin.

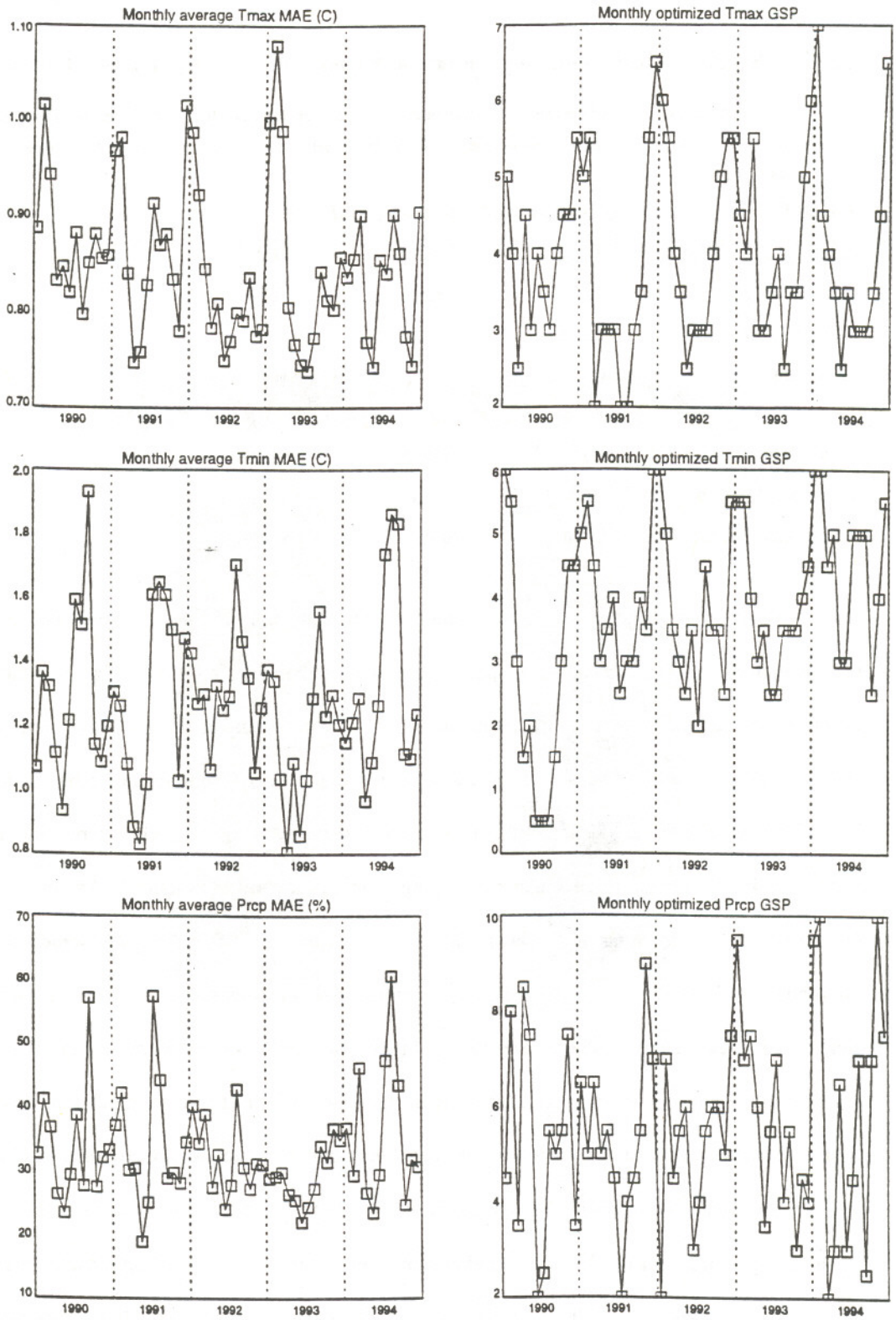


Figure 3
Mean absolute errors (MAE) for monthly averages of predicted Tmax, Tmin, and Prcp, together with the Gaussian shape parameters (GSP) found to minimize MAE, for the five year period 1990-1994.

Table 1. Results of t-tests comparing constant with temporally-varying parameterization of GSP.

Errors for precipitation are from predicted and observed monthly totals, and expressed on the basis of average daily total precipitation. Temperature errors are from predicted and observed daily values. These trials used 24 years of data, from 1970 through 1993, and excluded SNOTEL stations.

Variable	Constant GSP	Variable GSP	P
Prcp: MAE (cm day^{-1})	0.0268	0.0266	<0.01
Prcp: MAE (%)	23.7	23.5	<0.01
Prcp: bias (cm day^{-1})	+0.00079	+0.00055	<0.01
Prcp: bias (%)	+0.7%	+0.5%	<0.01
Tmax: MAE (°C day^{-1})	1.629	1.627	<0.01
Tmax: bias (°C day^{-1})	-0.033	-0.035	<0.01
Tmin: MAE (°C day^{-1})	1.864	1.861	<0.01
Tmin: bias (°C day^{-1})	+0.000	-0.002	<0.01

Based on my findings in Chapter 2 and further investigation with the current dataset, I determined that the inclusion of a spatial smoothing parameter was only beneficial in the case of precipitation estimates. By repeatedly performing the automatic optimization of GSP using a series of increasingly smoothed elevation grids, I was able to assess the impact of elevation smoothing on Prcp prediction error. I found that monthly errors (% of monthly totals) were minimized using a linearly tapered smoothing window with a circular extent and diameter varying by month as shown in Figure 4. As I have observed for previous implementations of the Daymet logic, optimal prediction of Prcp requires some terrain smoothing throughout the year, with the most smoothing in the summer months and the least in the winter months. This result suggests that the sensitivity of precipitation intensity to terrain is greatest in the winter and least in the summer, in correspondence with the increased organization of precipitation events in the winter (frontal systems) and the more random spatial distribution of summer precipitation (convective storms). The inclusion of terrain smoothing reduced the MAE for predictions of monthly total precipitation from 33.5% to 33.0%, a small but significant improvement, while significantly increasing the magnitude of the bias, from –0.0041 cm day^{-1} to –0.0061 cm day^{-1} ($P<0.05$ and $P<0.01$, respectively).

Based on results summarized in Chapter 2, Table 2, and further testing using the current dataset, I set the temporal smoothing parameters, used in regressions of temperature and precipitation against station

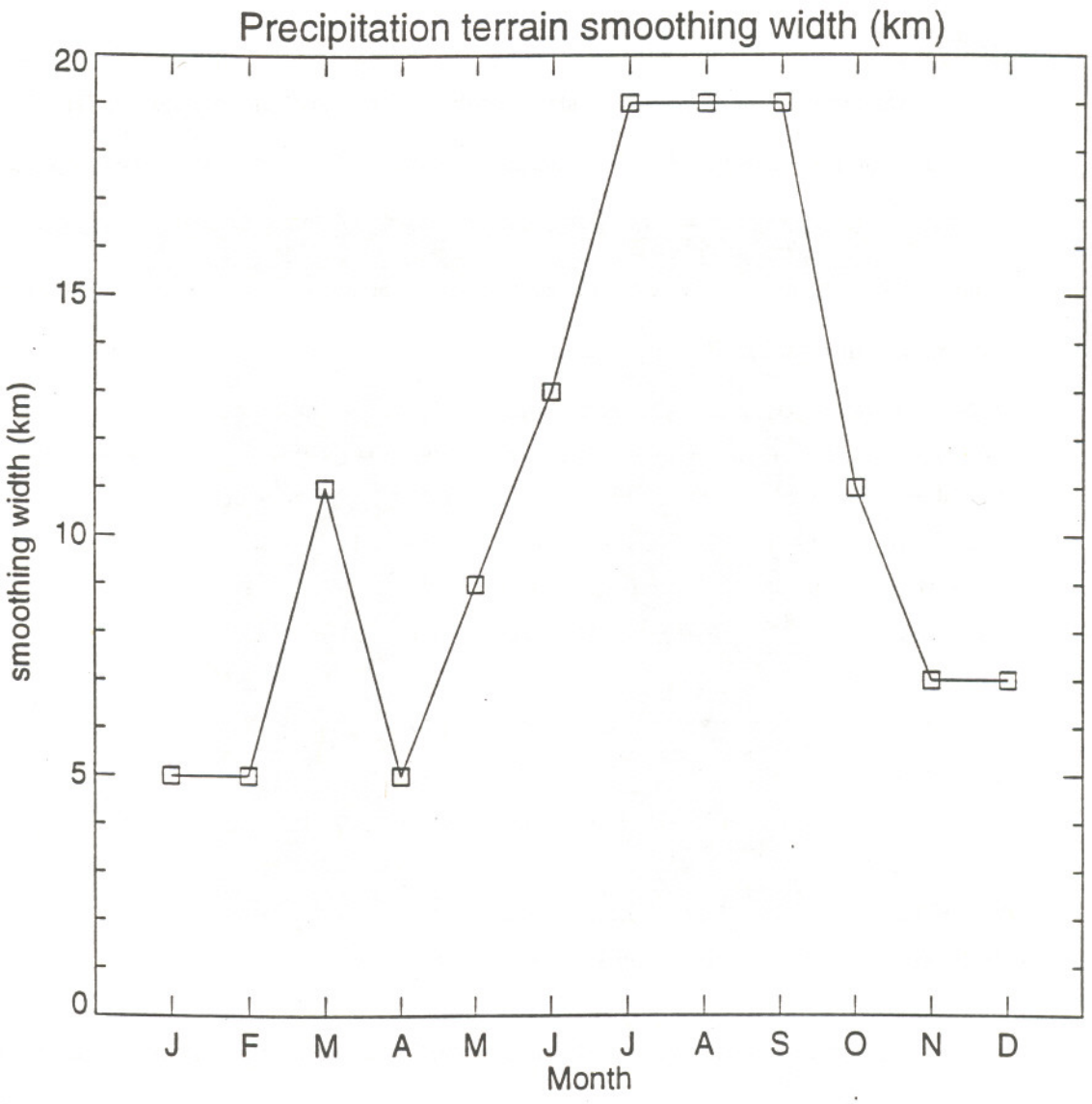

Precipitation terrain smoothing width (km)
smoothing width (km)
20
15
10
5
0
J
F
M
A
M
J
J
A
S
O
N
D
Month

Figure 4

elevation, to 1 day for both Tmax and Tmin, and 5 days for Prcp. The critical proportion of stations receiving precipitation, required in estimating precipitation occurrence, was set at 0.5. The f_{max} parameter was kept at 0.95, and the number of iterations used in estimating station density was kept at 3 (Chapter 2, Table 2).

The results of the final cross-validation analysis of error and bias are shown in Table 2. Errors for Tmax and Tmin are given both for daily predictions and for the 12-month average of daily predictions accumulated to monthly average predictions, for each year from 1990 to 1994, and for the average of all years. Only the 12-month average of summarized monthly prediction errors are presented for Prcp, for the same reasons discussed in Chapter 2, Section 3.2.

Table 2. Final cross-validation analysis estimates of error and bias.

MAE and bias have units (°C day^{-1}) for Tmax and Tmin, and (%) for Prcp. MAE_d = daily MAE, MAE_m = monthly MAE.

Variable	1990	1991	1992	1993	1994	5-year mean
Tmax MAE_d	1.63	1.58	1.52	1.62	1.55	1.58
Tmax MAE_m	0.87	0.87	0.82	0.85	0.83	0.85
Tmax bias	-0.038	-0.031	-0.032	-0.034	-0.034	-0.034
Tmin MAE_d	1.91	1.84	1.86	1.77	1.89	1.85
Tmin MAE_m	1.29	1.27	1.31	1.18	1.32	1.27
Tmin bias	+0.014	+0.003	+0.004	+0.007	-0.003	+0.005
Prcp MAE_m	33.9	33.9	32.3	29.2	35.9	33.0
Prcp bias	-7.0	-8.1	-6.4	-4.7	-10.7	-7.4

Because the bias statistics for Prcp are computed on the basis of individual months, and because bias (underestimation) in dry months is greater than in wet months (see Chapter 2, Figures 6 and 7, and accompanying discussion), the magnitude of Prcp biases in Table 2 are misleadingly high. As shown in Figure 5, where the monthly Prcp prediction biases are plotted against observed monthly total precipitation amounts for all months over the period 1990 to 1994, the biases in the driest months are very large (and negative), but for increasingly wet months the bias rapidly approaches zero. When Prcp bias is computed

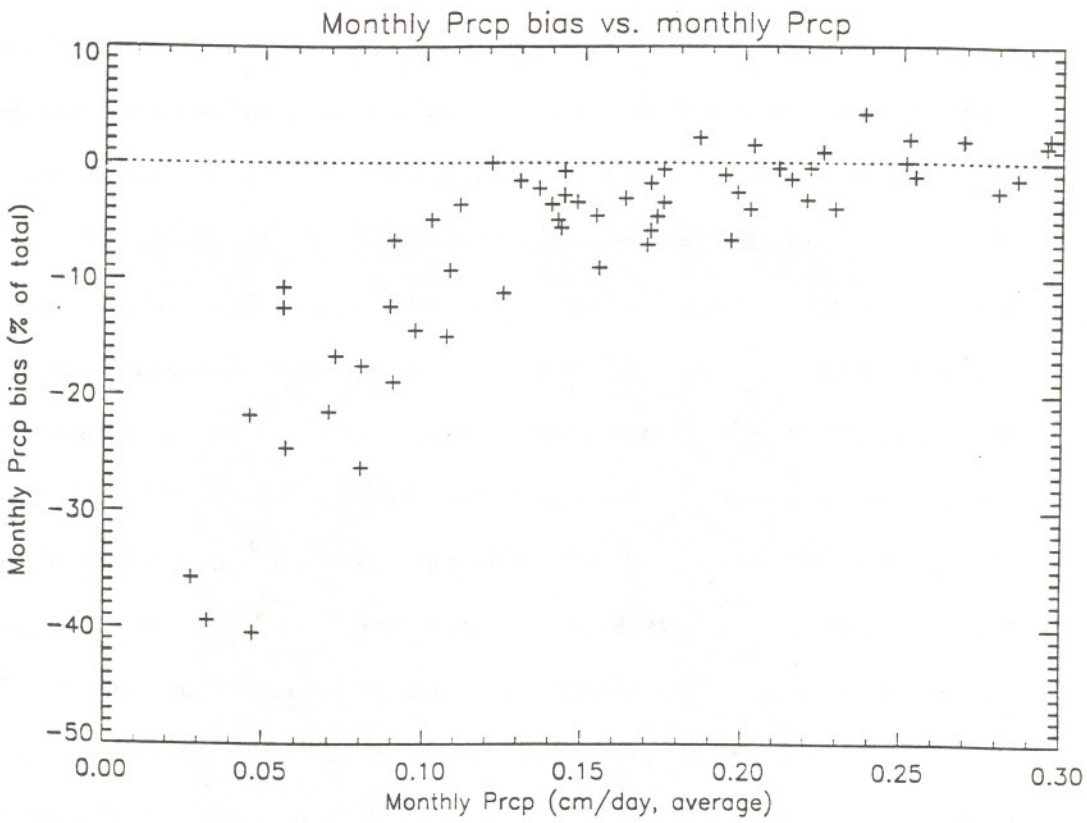

Monthly Prcp bias vs. monthly Prcp
Monthly Prcp bias (% of total)
10
0
-10
-20
-30
-40
-50
0.00
0.05
0.10
0.15
0.20
0.25
0.30
Monthly Prcp (cm/day, average)

Figure 5

by weighting each month according to its contribution to the annual total precipitation, the resulting 5-year mean bias is –3.5%, with annual values ranging from –5.0% to –2.9%.

5.2.3 Daymet simulations and summary of the resulting daily surface meteorological fields

The methods described in Chapter 2 for the final predictions of Tmax, Tmin, and Prcp were implemented here with minor modifications. One significant modification to the logic described in Chapter 2 is the implementation of a sign-switching mechanism in the regression estimates of the influence of elevation on temperature and precipitation. Recall from Chapter 2 that these regressions are based on all possible unique pairs of stations in the station list at a given point. Differences in either temperature or precipitation are regressed against differences in elevation, in order to diagnose the magnitude and sign of the local relationship. In Eqs. 3 and 4, Chapter 2, the intercept from these regressions is included as part of the final prediction. Since the regressions are based on differences between stations, in theory the intercepts should be very close to 0.0, and should not have a significant impact on the predictions. In preliminary testing with the current database, I found that the regression intercepts were significantly different from zero, and that inclusion of them in the prediction equations significantly reduced prediction error and bias. I determined that this was the result of always taking the difference between stations in a unique pair as the first minus the second, given that my database design placed SNOTEL stations (with generally high elevations) at the bottom of station lists and that the algorithm for finding unique station pairs always began at the top of station lists. By implementing a sign-switching convention that alternated the order of differencing between pairs, I found that the magnitude of the regression intercepts was reduced by two orders of magnitude (e.g., from a Tmax trial, from –5.2e-4 to –2.2e-6 °C m^{-1}), that the variance in the intercepts was also greatly reduced, that right skewness observed in the original intercepts was eliminated, and that the inclusion of the intercepts in the prediction equations no longer had a significant effect on errors or biases. I therefore retained the sign-switching convention and dropped the regression intercepts from the prediction equations.

The only other significant modification to the Daymet logic from the description in Chapter 2 is for the estimation of the daytime average vapor pressure deficit (VPD). In the example in Chapter 2, I

used the assumption from Running et al. (1987) that dewpoint temperature was equal to night minimum temperature in calculating VPD. In the current implementation I could not find sufficient evidence to support the assumption that the proposed relationship between minimum temperature and dewpoint temperature is valid for the case of minimum temperature varying over short distances in response to elevation. I therefore generated a second set of minimum temperature predictions for estimating dewpoint temperature, by eliminating the correction for elevation.

The final modification to the Daymet methods was a correction to the values for the monthly variation in the solar constant. I found that the average of the values from the original MTCLIM code (Running et el. 1987, Hungerford et al., 1989) was too high by 43 W m^{-2}, or 3.1% (1403 W m^{-2} instead of 1360 W m^{-2} as given by Gates (1980), Jones (1992), and others). I subtracted 43 W m^{-2} from all the original monthly values to obtain the new solar constant array.

Results from the Daymet simulations are summarized in Table 3. For temperatures as well as for precipitation, 1990 and 1991 were typical of the 5-year mean conditions, while 1992 and 1994 were warmer and drier than average, and 1993 was cooler and wetter than average. Solar radiation was higher than average in 1990, 1992, and 1994, and lower than average in 1993. Monthly spatial averages of temperature, precipitation, and solar radiation (Srad) for the five years are shown in Figure 6.

Table 3. Annual and 5-year mean statistics for Daymet output

Values for Tmax and Tmin are for 365-day average temperature (°C), values for Prcp are for 365-day total precipitation (cm), and values for Srad are 365-day averages of daylight radiant flux density (W m^{-2}). Mean and standard deviation (stdev) are taken over all 1 km^2 pixels (n=606526).

Variable	1990	1991	1992	1993	1994	5-year mean
Tmax (mean)	12.90	12.73	13.71	11.13	13.46	12.78
Tmin (mean)	-0.98	-0.86	-0.21	-1.88	-0.58	-0.90
Prcp (mean)	54.7	54.5	47.5	59.8	48.3	53.0
Srad (mean)	289.4	283.1	290.4	273.2	291.0	285.4
Tmax (stdev)	3.53	3.19	3.35	3.00	3.23	3.25
Tmin (stdev)	2.92	2.74	2.76	2.71	2.80	2.77
Prcp (stdev)	39.6	28.4	26.4	24.6	29.0	28.7

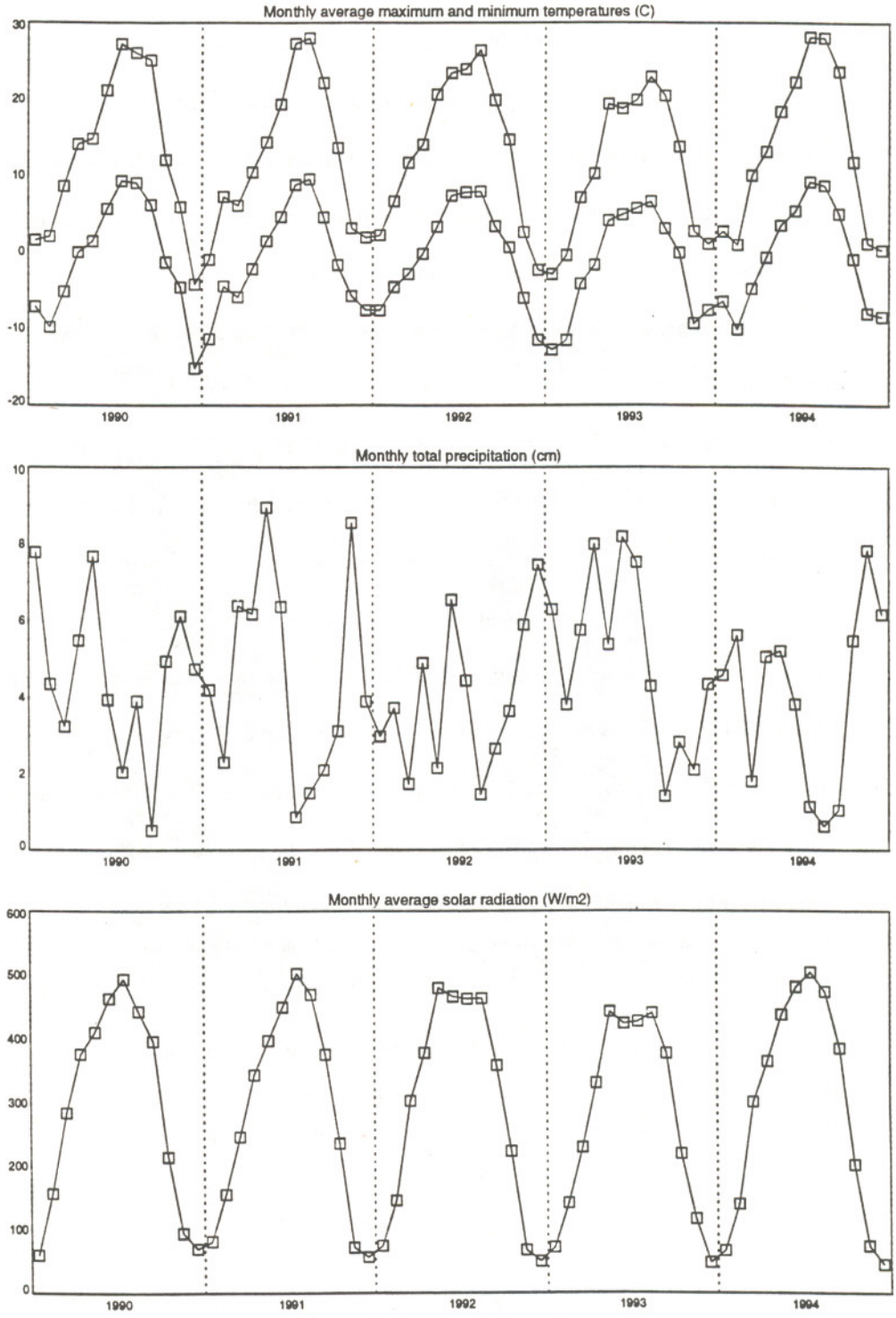


Figure 6
Results of Daymet estimation of meteorological parameters. Temperatures and radiation are shown as averages of the daily results over the study region for each month, and precipitation is shown as the monthly accumulated depth (water equivalent) averaged over the study region.

5.3 Estimation of subdaily meteorological parameters

In order to assess the influence of meteorological conditions on observed variation in landsurface temperature as detected from a polar-orbiting spacecraft, I require estimates of the most important meteorological variables at the time of data acquisition, i.e. at the time of day of satellite overpass. In particular, estimates of the near-surface air temperature, the incident solar radiation, and the vapor pressure deficit at the time of overpass must be obtained. The methods described in Chapter 2 are designed to provide daily summaries of these variables. Qualitative relationships could be drawn between these daily average estimates and the observed surface temperatures, but my purpose here is to investigate the quantitative relationship between meteorological parameters, surface energy balance theory, and the remotely observed surface temperatures. In this section I describe the methods by which I arrived at the subdaily estimates of meteorological parameters that allowed me to make a quantitative as opposed to qualitative analysis.

The application of these methods depends on a knowledge of the date and time of day of satellite data acquisition for each grid cell. In this section I will only discuss the general case of subdaily meteorological variable estimation, and will postpone a discussion of the specific application until after I have described the database of radiometric observations and associated temporal compositing and viewing geometry variables.

5.3.1 Subdaily air temperature estimates

I reduced the problem of subdaily air temperature estimation to two separate sub-problems. Given that I already had an estimate of the daily maximum temperature, and that the time of satellite overpass was in the approximate neighborhood of the time of the maximum temperature, my tasks were first to estimate the time of day of the maximum temperature, and second to estimate the shape of the diurnal temperature curve around the time of maximum temperature. Then, from the Daymet database already described and a knowledge of the local time of satellite data acquisition, I could derive the relevant air temperature. My

criteria in tackling these two problems was that the eventual solution should rely only on the information already available from the Daymet database.

In order to assess the factors influencing the time of day of maximum temperature, I gathered hourly meteorological data from several sources. From the Western Regional Climate Center, I obtained hourly temperature data for nine airport stations within the study region, over the period 1983-1994. Preliminary tests with those data indicated a possible inverse relationship between the long-term growing-season average time of day of maximum temperature (T_{x0}) and the long–term average maximum temperature. I intended to use these hourly data to derive an empirical relationship that I could then apply to predict T_{x0} over the entire study region, but the range of topographic settings for the airport stations was limited, and the highest elevation was 1357 m (Pocatello, ID). I sought additional data from higher elevations and in more topographically complex settings to test the generality of the patterns observed in the airport data. The science staff at Glacier National Park (GNP) provided hourly observations of temperature from four stations within the Park, having various topographic settings and ranging in elevation from 965 m to 2276 m.

Observations at these stations are made either on the hour (airport stations) or as hourly averages reported at the end of the hour (GNP), in both cases with respect to local standard time. Because the stations are located at various longitudes within their time zones, the local solar time for reporting is not directly comparable between stations. Since it is the diurnal pattern of temperature with respect to local solar time which is of interest in coordinating air temperature estimates with satellite observations of surface temperature, and since the difference between stations in T_{x0} is of the same order of magnitude as the potential error from this longitudinal bias, a correction was required. An offset (ΔT) was assigned to each station on the basis of its longitude (L_s), the central longitude of its time zone (L_z), and the number of degrees of longitude per hour of solar time (15.0):

Eq. 19 $\Delta T = (L_s - L_z) / 15.0$

Because the GNP stations recorded the average temperature over the previous hour, the most likely time of the true maximum temperature is one half hour prior to the reporting time at which T_{x0} is observed, and so

for these stations ΔT from Eq. 19 was amended by subtracting a half hour. Because the airport stations report the instantaneous temperature on the hour, the most likely time for the true T_{x0} is on the reporting hour, and so this additional correction is not necessary for these stations. Table 4 lists all stations, their elevations, geographic coordinates, and values for ΔT.

Table 4. Station data for 13 stations used in assessment of subdaily air temperature.

Station name	Elev. (m)	Lat. (°)	Long. (°)	ΔT (hrs)
Airport stations				
Boise, ID	865	43.57	-116.22	-0.75
Lewiston, ID	438	46.38	-117.02	+0.20
Pocatello, ID	1357	42.92	-112.60	-0.51
Great Falls, MT	1116	47.48	-111.37	-0.42
Helena, MT	1167	46.60	-112.00	-0.47
Kalispell, MT	909	48.30	-114.27	-0.62
Missoula, MT	972	46.93	-114.10	-0.61
Pendleton, OR	452	45.68	-118.85	+0.08
Spokane, WA	718	47.63	-117.53	+0.16
GNP stations (all in MT)				
Mt. Brown	2276	48.63	-113.83	-1.09
Mt. Otokomi	2227	48.71	-113.53	-1.07
Sun Point	1453	48.68	-113.58	-1.07
West Glacier	965	48.51	-114.00	-1.10

The hourly temperature data were screened to remove days with missing data between the hours of 9 AM and 9 PM. Since the analysis of observed surface temperatures is concerned only with the growing season, defined here to be the period from April 1 to October 31, only those days were extracted from the original hourly temperature data for further analysis. For each day, the time of the maximum temperature was determined as the final reporting period having the day's maximum temperature. Many days have the same temperature recorded for multiple reporting periods in the middle of the day, and this method assured a consistent interpretation of the shape of the diurnal temperature trend.

I found that the long-term average growing-season maximum temperature ($Tmax_0$) was a good predictor of the long-term growing-season average of the daily times of maximum temperature obtained by this method (T_{x0}). This relationship is illustrated in Figure 7, where different symbols are used for airport stations (diamonds) and GNP stations (stars). The regression line has the formula:

Eq. 20 $T_{x0} = 12.51 + 0.1247\ Tmax_0$

with an R^2 of 0.79 and a mean absolute prediction error of 0.21 hrs. Figure 7 demonstrates that the relationship observed for the airport stations extrapolates reasonably well to the higher elevation stations in complex terrain.

Using Eq. 20 as a first order approximation of the spatial variability in the time of Tmax, I explored methods to allow predictions at finer temporal resolution. For these further investigations I used only the airport stations, since their records were more complete and of longer duration. My approach was to calculate the deviation in the time of Tmax away from T_{x0} for each 14-day period throughout the growing season (T_{x1}) as a new dependent variable. I found that a multiple regression based on the 14-day average deviations of Tmax, Tmin, and Srad from their long-term growing season averages ($\Delta Tmax_1$, $\Delta Tmin_1$, and $\Delta Srad_1$), and the average warming or cooling per day over the 14-day period ($dTmax_1$, °C day^{-1}), gave reasonable predictions of T_{x1}. The observed and predicted 14-day anomalies from T_{x0} are shown in Figure 8a. The regression equation is:

Eq. 21 $T_{x1} = 0.005 + 0.036\ \Delta Tmax_1 - 0.030\ \Delta Tmin_1 + 0.032\ \Delta Srad_1 + 0.177\ dTmax_1$

with an R^2 of 0.35 and a mean absolute prediction error of 0.33 hrs.

As the final step in predicting the time of Tmax, I calculated the deviations from T_{x1} for each day (T_{x2}) as a new dependent variable. I found that a multiple regression similar in form to Eq. 21 helped to explain a small proportion of this daily variance. The independent variables selected for this prediction were the daily deviations of Tmax and Tmin away from their 14-day average values ($\Delta Tmax_2$ and $\Delta Tmin_2$), the difference between this day's and the next day's Tmax ($dTmax_2 = Tmax_{i+1} - Tmax_i$), and the difference between this day's and the previous day's diurnal temperature range ($dDTR_2 = DTR_i - DTR_{i-1}$). The observed and predicted daily anomalies from T_{x1} are shown in Figure 8b. The regression equation is:

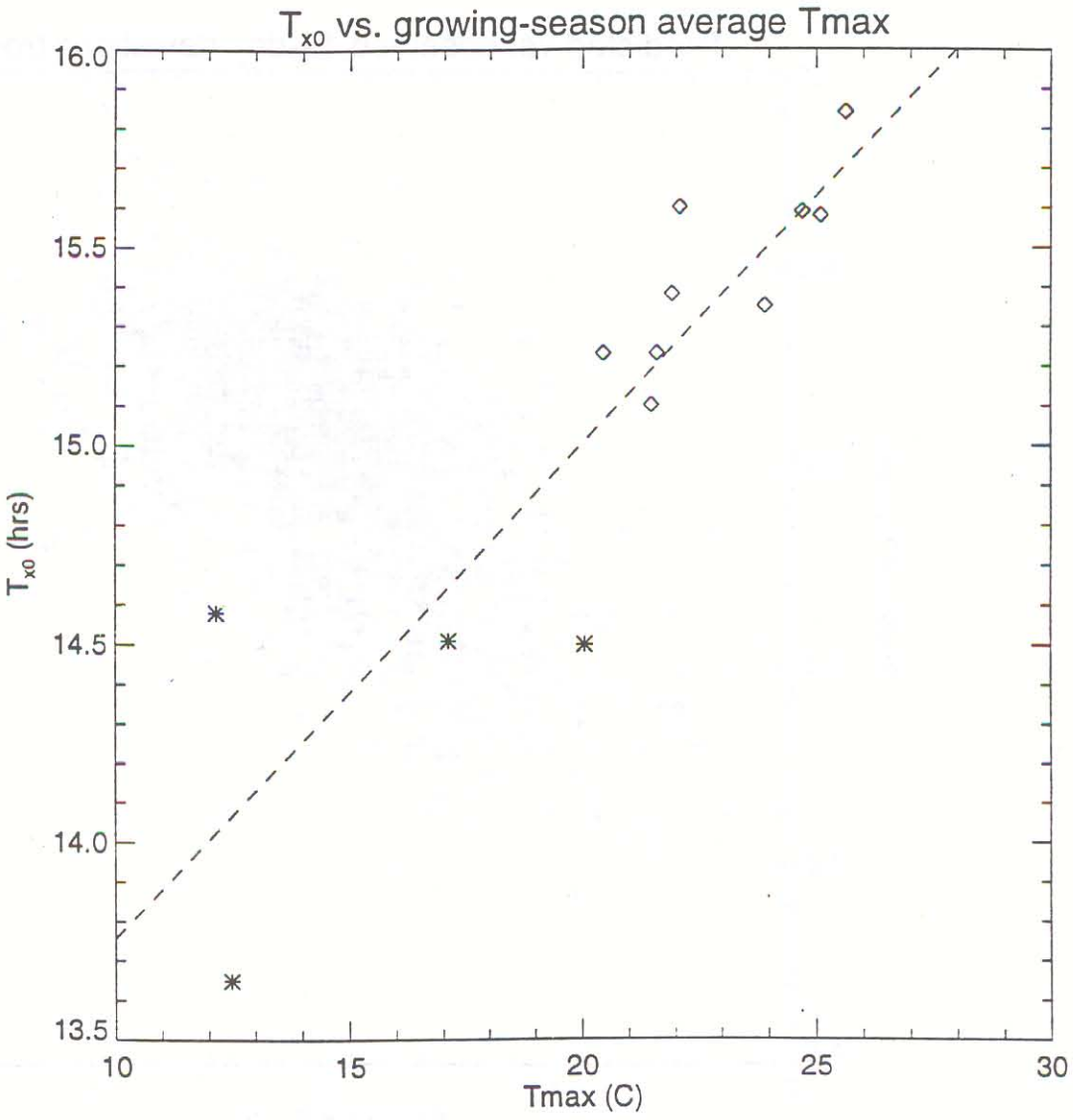

T_{x0} vs. growing-season average Tmax
16.0
15.5
15.0
14.5
14.0
13.5
T_{x0} (hrs)
10
15
20
25
30
Tmax (C)

Figure 7

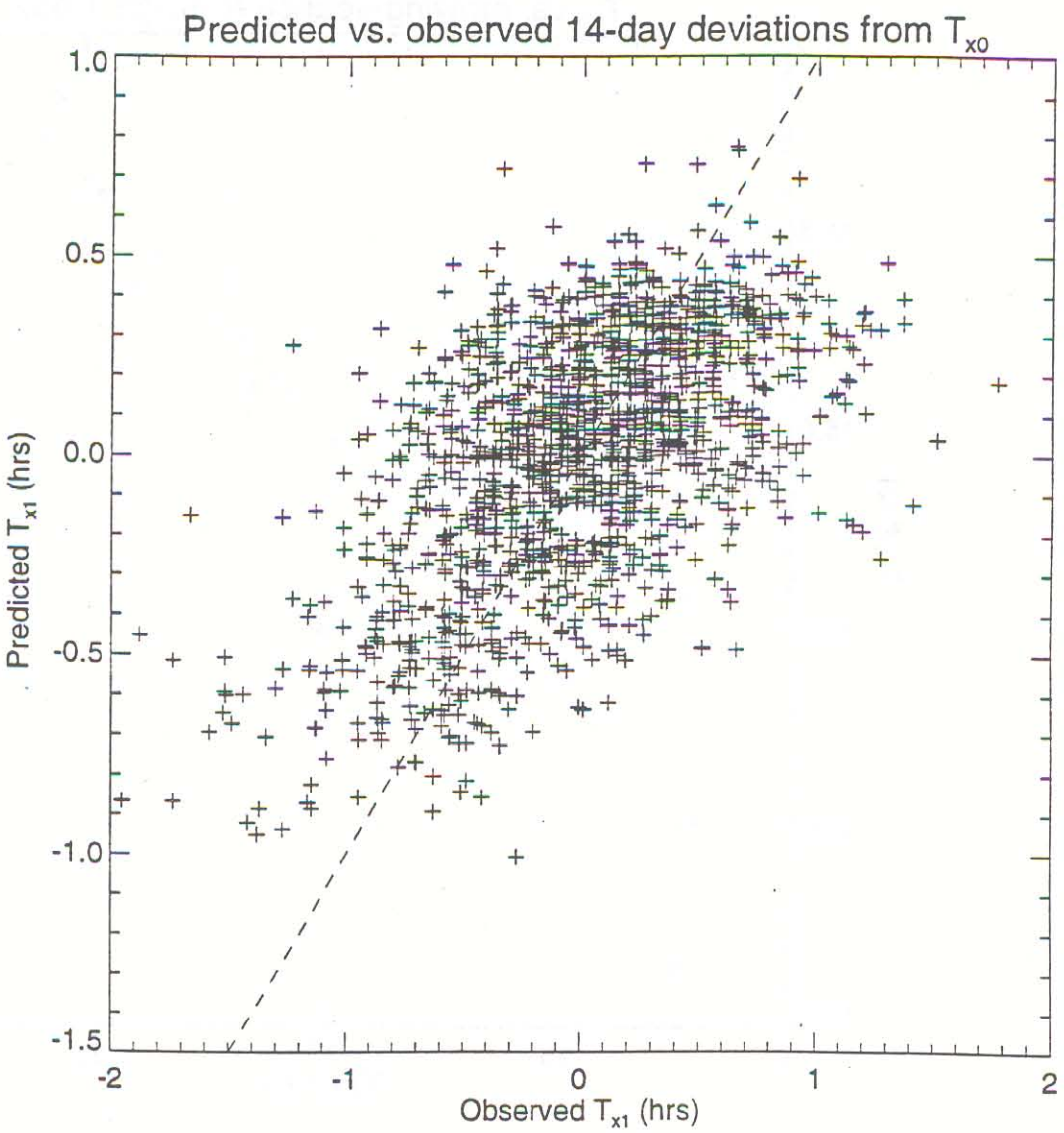

Predicted vs. observed 14-day deviations from T_{x0}
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2
-1
0
1
2
Predicted T_{x1} (hrs)
Observed T_{x1} (hrs)

Figure 8a

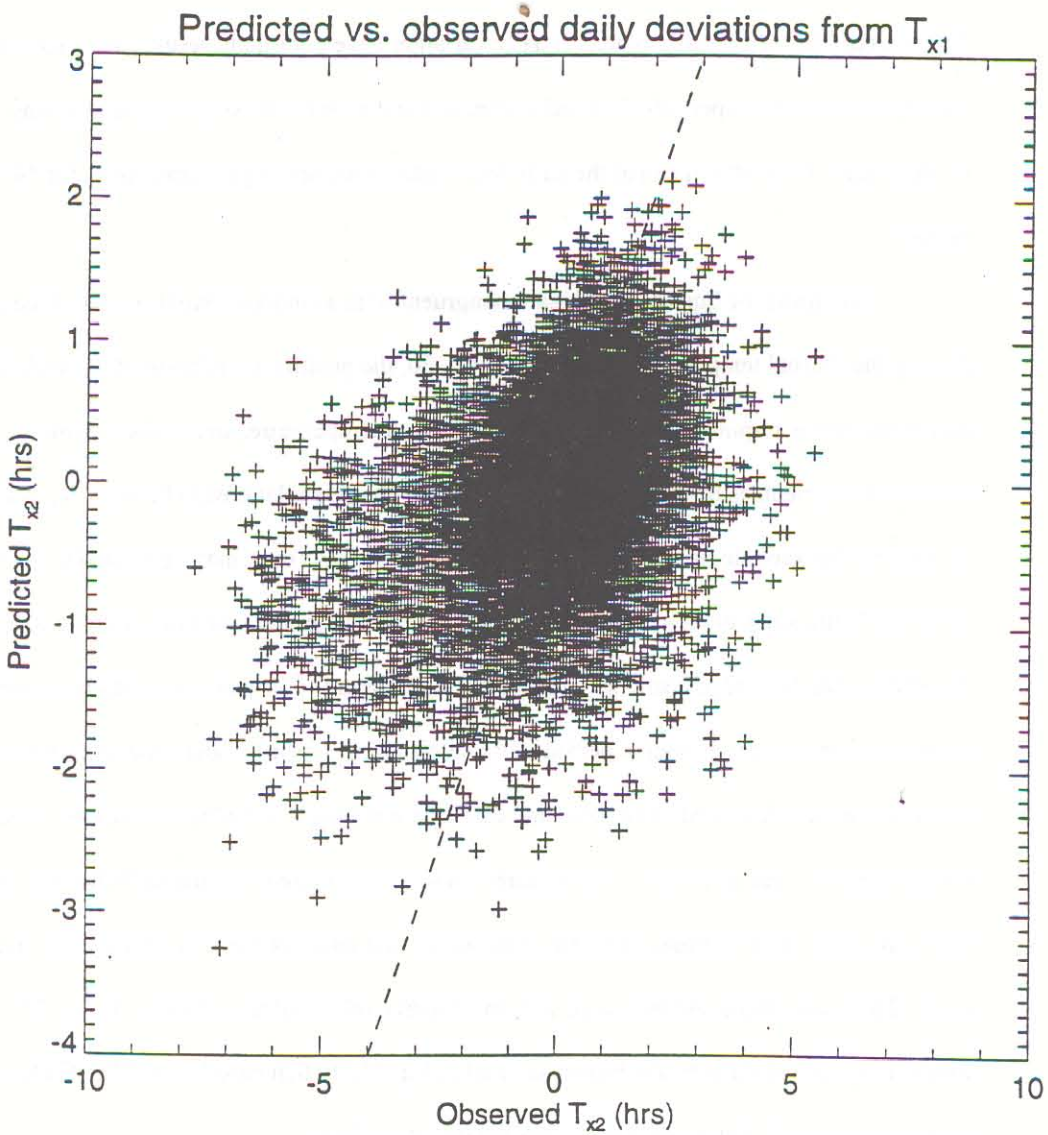

Predicted vs. observed daily deviations from T_{x1}
Predicted T_{x2} (hrs)
Observed T_{x2} (hrs)
3
2
1
0
-1
-2
-3
-4
-10
-5
0
5
10

Figure 8b

The purpose for the step-wise regression employed here is to ensure that at least the medium- and long-term averages of T_{x2} are reasonably well predicted. I was unable to find any regression relationship that provided good predictions at the daily level, using either step-wise or one-step regressions. I found that a single-step regression to predict daily variation gave poor daily results, and that upon averaging to 14-day and long-term periods the results were worse than with the step-wise regressions. The step-wise method gains what little it can at the daily level without sacrificing accuracy over the 14-day and long-term periods.

The forms of Eqs. 2, 3, and 4 are congruent with a simple thought model of the physical forces driving the diurnal temperature variation. In Eq. 20, the positive correlation of T_{x0} with $Tmax_0$ suggests that more time is required to reach a higher maximum temperature, and so the maximum temperature in hotter environments is reached later in the day. The positive relationship between T_{x1} and $\Delta Tmax_1$ in Eq. 21 follows this same logic. The negative relationship with $\Delta Tmin_1$ may be related to the influence of cloud cover on Tmin, since cloudy conditions would be expected to increase Tmin and to decrease the energy absorbed at the surface, causing an earlier peak temperature. The positive relationship with $\Delta Srad_1$ tends to support this inference. During a period of warming a shift toward a later time of maximum temperature should occur, as illustrated by the extreme case of a warming trend which causes the maximum temperature to occur after dark. By the same logic a cooling trend should shift the time of Tmax to earlier hours, as again in the extreme example of the maximum temperature at sunrise before an arriving cold air mass. These arguments are in accord with the positive relationship to $dTmax_1$ in Eq. 21. All the same arguments apply to the relationships expressed in Eq. 22. In that case I found that DTR was a better predictor than Srad, but the two quantities are directly related.

The second part of the subdaily air temperature prediction problem is to estimate the shape of the temperature curve in the vicinity of the maximum temperature. My approach was to generate separate regressions to predict the temperature for the four hours preceding and the four hours following the time of Tmax. Each regression predicts the relevant temperature as a fraction of the difference between Tmin and Tmax, where the current day's value of Tmin is used for the hours before Tmax, and the next day's Tmin is

used for the hours after Tmax. Preliminary examinations of temperature curves for the nine airport stations showed that there were few significant differences between them in their shapes, when normalized according to the difference between Tmax and Tmin. Figure 9 shows the long-term average values for the normalized temperature curve (NTC) before and after the time of Tmax, with the average of all stations plotted as a solid line and symbols marking the values for individual stations. I was not able to explain the variation between stations in the long-term averages of NTC, and so I did not include a spatial component of variation (such as Eq. 20) in the regressions for NTC.

I found that one third to one half of the variance in 14-day average values for NTC ($NTC14_h$, where h indicates the number of hours away from T_{x2}) could be explained using a multiple regression with $Tmax_{14}$, DTR_{14}, $Prcp_{14}$, and $Srad_{14}$ (the 14-day means of these parameters) as predictors. Table 5 lists the resulting regression coefficients for each hour.

Table 5. Regression coefficients for the prediction of $NTC14_h$.

Coefficients marked * are not significantly different from 0.0 (P>0.1).

MAE is for the normalized fraction of DTR, unitless.

Other units: Tmax and DTR (°C), Prcp (cm), Srad (W m^{-2}).

Hour	R^2	MAE	intcpt.	Tmax	DTR	Prcp	Srad
-4	.46	.043	.561	-.0001*	.0049	-.1431	.0055
-3	.43	.035	.682	.0008*	.0047	-.1227	.0028
-2	.40	.027	.789	.0007	.0046	-.0895	.0010
-1	.34	.018	.882	.0007	.0033	-.0587	-.0002*
+1	.50	.019	.810	.0007	.0047	-.0886	.0002*
+2	.42	.033	.686	.0011	.0027	-.1147	.0029
+3	.32	.041	.582	.0012	-.0004*	-.0948	.0044
+4	.19	.044	.488	.0004*	-.0001*	-.0538	.0038

I was not able to account for a significant amount of the daily variance in NTC observations using the variables available from Daymet predictions.

In some cases, the temperature before or after the 8-hour period covered by the regressions in Table 5 may be required. In those cases I use a simple linear interpolation from the temperature at the last predicted hour to that day's or the previous day's minimum temperature, assuming that the minimum

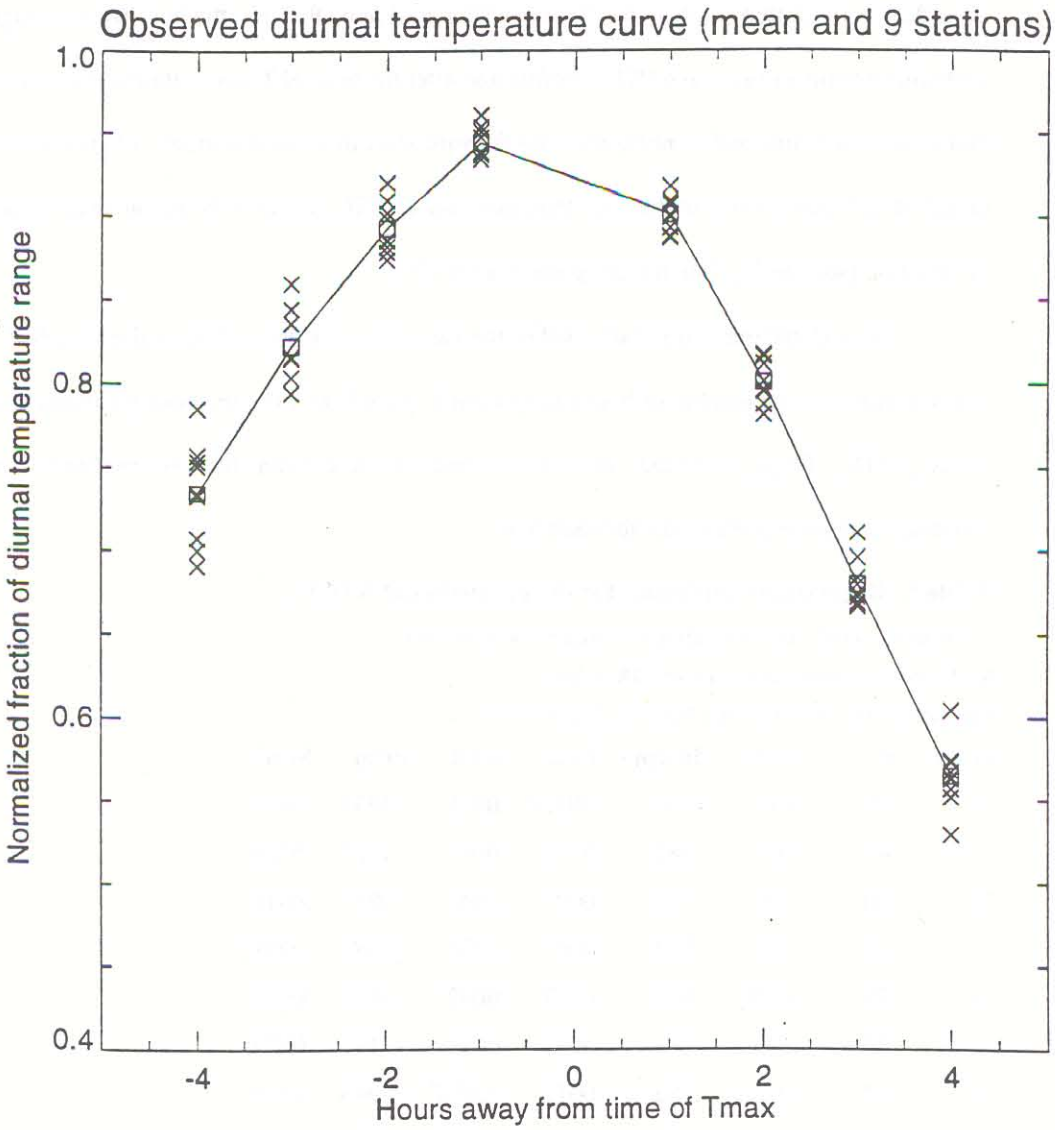

Observed diurnal temperature curve (mean and 9 stations)
Normalized fraction of diurnal temperature range
1.0
0.8
0.6
0.4
-4
-2
0
2
4
Hours away from time of Tmax

Figure 9

temperature occurs at sunrise and using the Daymet predictions of daylength to estimate the time of sunrise. A schematic of the complete diurnal temperature prediction algorithm is given in Figure 10. In order to illustrate the effectiveness of the entire subdaily temperature prediction algorithm, I generated predictions of the instantaneous hourly temperature between the hours of 12 and 6 pm (local solar time) for all of the days with good observations from the airport station network, over the 11-year period from 1983-1994. Observed temperatures were interpolated between reporting times to obtain values on the hour in local solar time for the generation of error statistics. The mean absolute prediction errors for each station at each hour are listed in Table 6.

Table 6. Mean absolute prediction errors (°C) for subdaily temperature.

	Hour of prediction (pm)						
Station	**12**	**1**	**2**	**3**	**4**	**5**	**6**
Boise, ID	2.64	2.38	2.23	2.11	1.77	1.58	1.76
Lewiston, ID	2.17	2.13	2.12	2.05	1.63	1.61	1.91
Pocatello, ID	2.29	2.04	1.88	1.75	1.51	1.59	1.66
Great Falls, MT	1.95	1.78	1.72	1.68	1.63	1.80	1.96
Helena, MT	2.13	2.08	2.07	1.97	1.80	1.98	2.08
Kalispell, MT	2.52	2.45	2.40	2.22	1.85	1.98	2.01
Missoula, MT	1.78	1.76	1.77	1.68	1.50	1.71	1.97
Pendleton, OR	2.55	2.42	2.36	2.27	1.80	1.79	1.91
Spokane, WA	2.19	1.99	1.86	1.79	1.46	1.51	1.72

Mean prediction errors summarized by yearday are plotted for the growing season (April through October, yeardays 90 to 300) in Figure 11, illustrating the seasonal pattern of prediction accuracy. As an illustration of the distribution of prediction errors, all observations and predictions are plotted in Figure 12. An example of the temporal pattern of prediction skill is shown in Figure 13, where observed (solid lines) and predicted (dashed lines) hourly temperatures between noon and 6 pm are plotted together for two 14-day periods for the Missoula station. Note that during a warming trend or when the maximum temperature is high and stable over a number of days the predictions are quite good, but that during a period of cooling

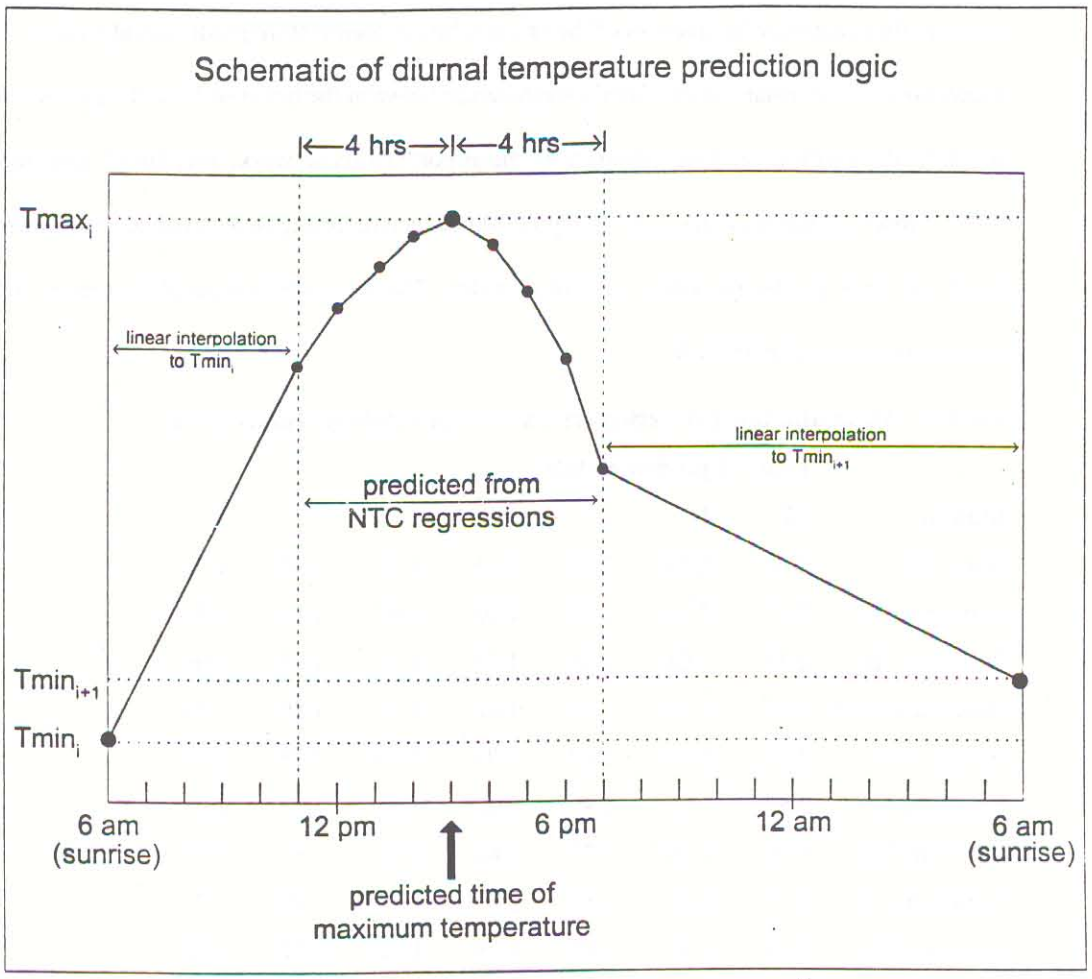

Schematic of diurnal temperature prediction logic
4 hrs
4 hrs
Tmax$_i$
linear interpolation to Tmin$_i$
linear interpolation to Tmin$_{i+1}$
predicted from NTC regressions
Tmin$_{i+1}$
Tmin$_i$
6 am (sunrise)
12 pm
6 pm
12 am
6 am (sunrise)
predicted time of maximum temperature

Figure 10

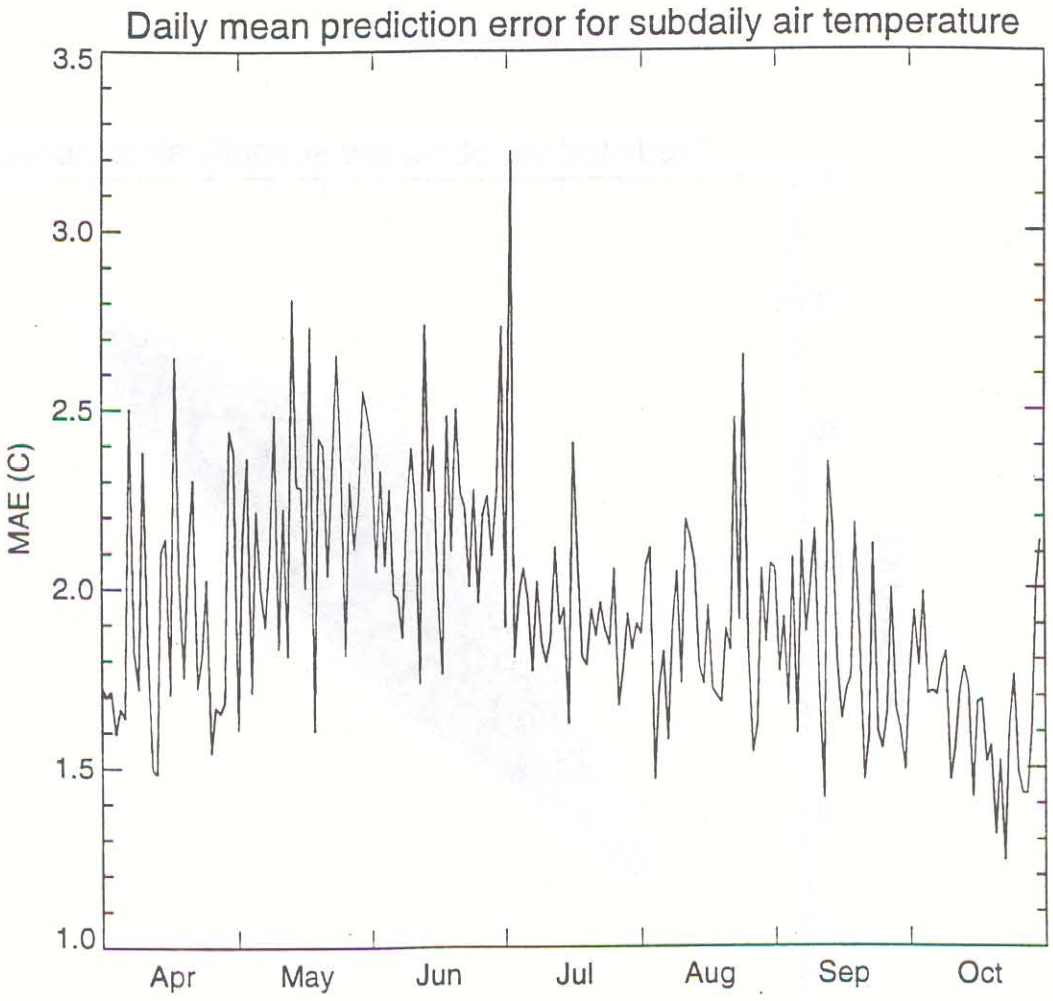

Daily mean prediction error for subdaily air temperature
MAE (C)
3.5
3.0
2.5
2.0
1.5
1.0
Apr
May
Jun
Jul
Aug
Sep
Oct

Figure 11

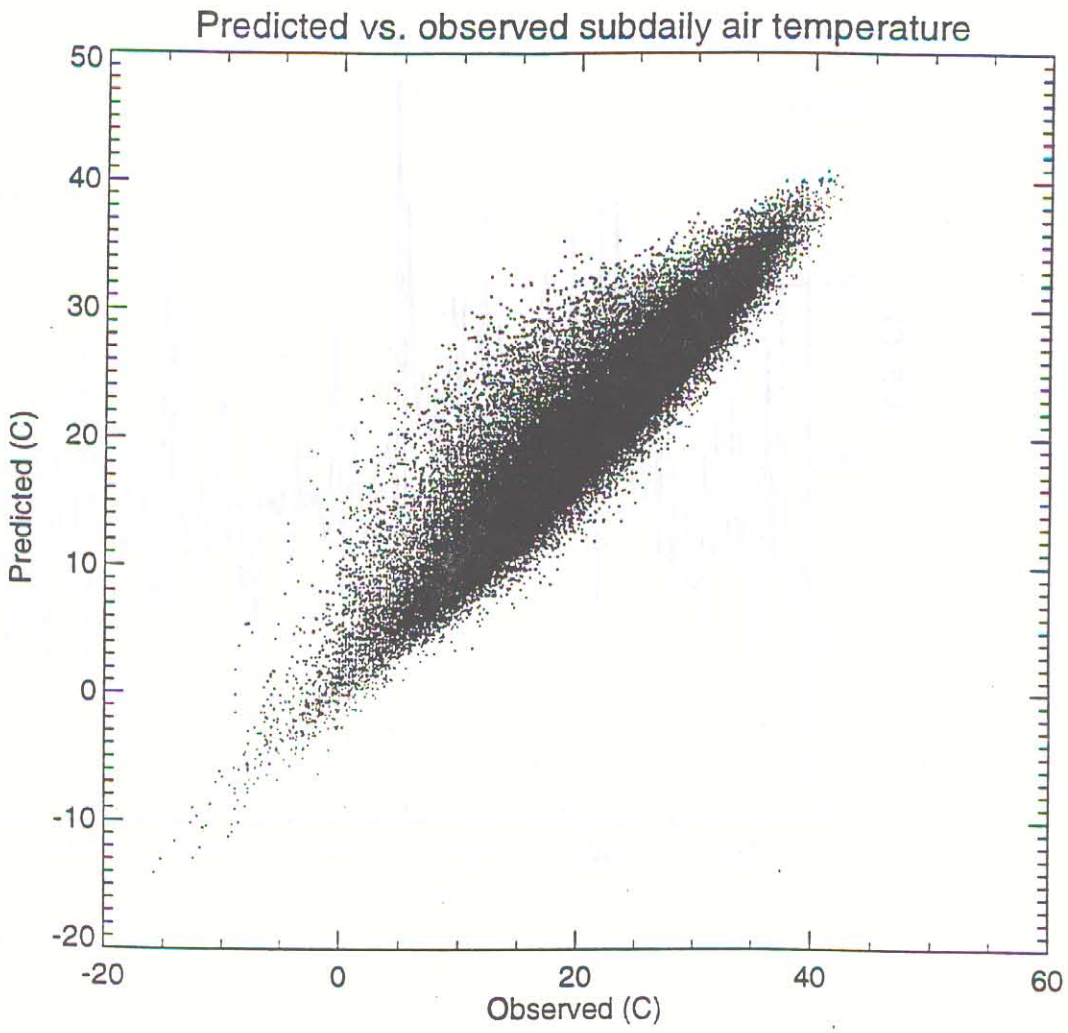

Predicted vs. observed subdaily air temperature
Predicted (C)
Observed (C)
50
40
30
20
10
0
-10
-20
-20
0
20
40
60

Figure 12

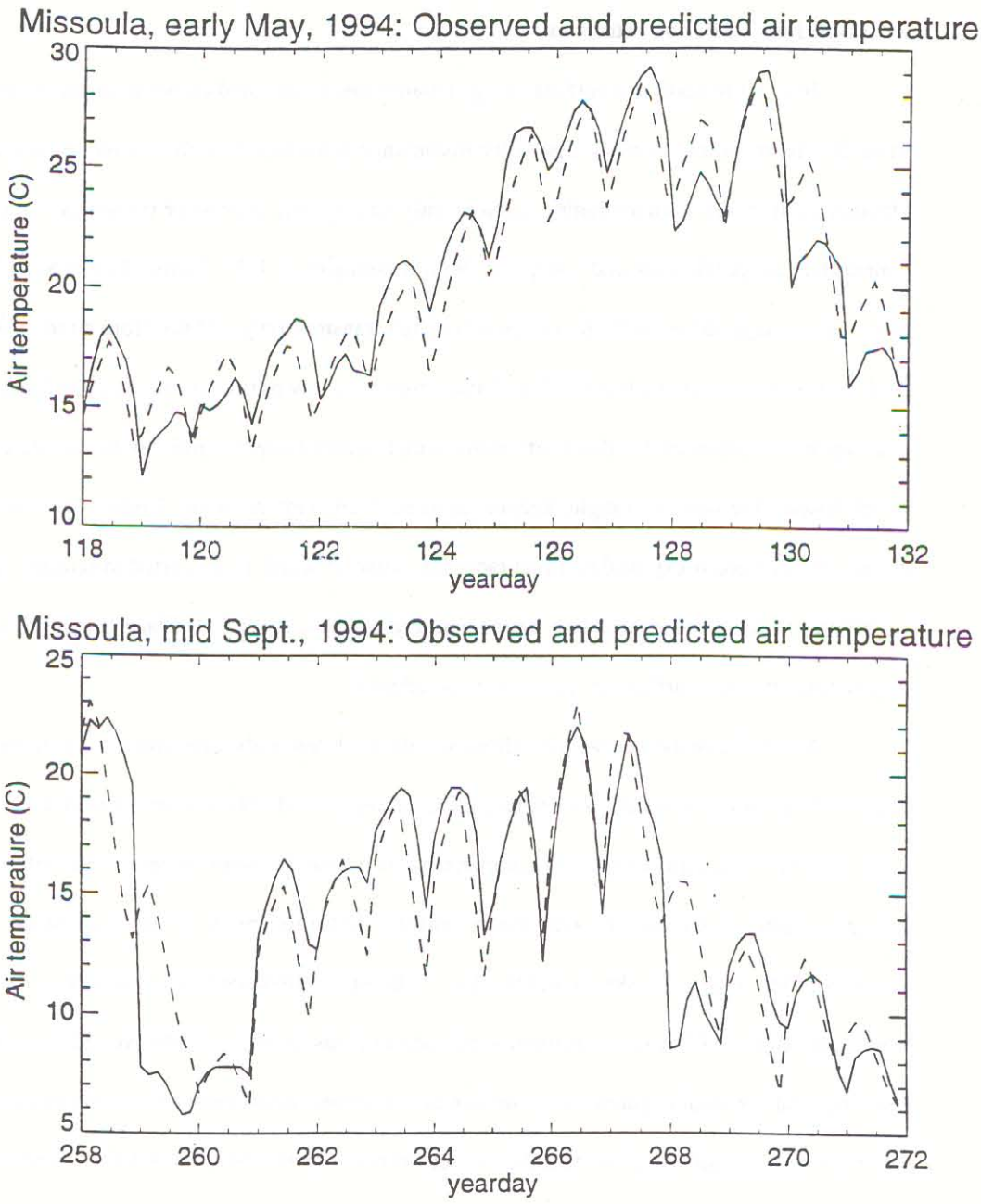


Figure 13
Observed and predicted hourly air temperatures at Missoula, MT. Observed (solid line), predicted (dashed line). Hourly values from noon till 6 PM are plotted for each day.

passage of fronts with strong spatial temperature gradients. The mean prediction error over all hourly predictions was 1.95 °C, with a bias of +0.09 °C.

5.3.2 Subdaily radiation estimates

In order to apply the surface energy balance theory outlined earlier to instantaneous observations of surface temperature, a knowledge of the instantaneous net radiation flux density is required. The Daymet methods use a logic relating average daily atmospheric shortwave transmissivity to the diurnal temperature range (Bristow and Campbell 1984, Running et al. 1987, Hungerford et al. 1989). It has been shown that a high value for DTR is related to a high transmissivity and therefore an increased radiation load, and conversely that a low DTR is characteristic of a low transmissivity and a reduced radiation load. The physical explanation for this relationship is that clouds damp the mid-day temperature increase and radiate toward the surface at night, keeping minimum temperatures high. Under clear conditions during the day, the surface is exposed to direct radiation, some of which is converted to sensible heat which contributes to an increase in surface air temperature, while at night the surface radiates to a cold atmosphere and near surface air temperature is reduced.

As I will discuss below, the radiometric data for this study were composited from a large number of daily observations in an effort to reduce the occurrence of cloud cover in the final dataset. Although the cloud removal by compositing is not complete (also discussed below), in general the satellite view is through a relatively unobscured atmosphere, and at least for the instant of data acquisition, and in the direction of the sensor, the sky should be clear. I therefore eliminated the correction to clear-sky transmissivity from DTR in the estimation of instantaneous incident shortwave radiation. By doing so I am assuming that the sampled area on the surface has an unobstructed view of both the sensor and the sun at the time of data acquisition, i.e., that the surface observed by the sensor is not shadowed by clouds. This assumption is not likely to be met during relatively cloudy periods when views through broken cloud fields are incorporated in the final composited data. I am also assuming that haze and aerosols are not contributing significantly to spatial variation in otherwise clear-sky transmissivities. Both of these

assumptions, if not met, will contribute to error in my results. I do include spatially homogeneous corrections of clear-sky transmissivity due to elevation and optical air mass.

After computing a transmissivity, the rest of the original MTCLIM radiation algorithm is employed unchanged (except for solar constant correction noted above). The original logic iterates through a series of calculations for each day of prediction, accumulating the solar radiation inputs according to varying sun-slope geometric relationships over the course of the day. Here I use those same equations, but perform only one calculation of instantaneous incident radiation, for the local solar time of satellite data acquisition.

5.3.3 Subdaily humidity estimates

Vapor pressure deficit (VPD) is an important variable in the surface energy balance equation, contributing to the control over partitioning of absorbed radiation between sensible and latent heat fluxes. VPD is exponentially related to air temperature for a constant absolute humidity, or mass of water vapor per volume of air, as a result of the exponential relationship between air temperature and the saturation water vapor pressure. Diurnal variation in air temperature therefore has a strong influence on VPD, while the absolute amount of water vapor in the air may remain relatively constant over the course of a day. In estimating instantaneous (subdaily) VPD, I assumed that the near-surface atmospheric vapor pressure was constant over the day, and that the minimum temperature (without correction for elevation) was an adequate surrogate for the dewpoint temperature (Running et al. 1987). I then used the instantaneous air temperature derived by the previously discussed methods and the daily average dewpoint temperature to assess the instantaneous VPD, according to the empirical relationship between temperature and saturation water vapor pressure given by Abbott and Tabony (1985):

$$VPD = es(Tair) - es(Tmin)$$

$$es(T) = 610.7 \exp(17.38\ T / (239.0 + T))$$

where Tair and Tmin are in °C, VPD is in Pa, and es(T) is the function giving the saturation vapor pressure in Pa at a given air temperature, T in °C.

5.4 AVHRR database generation

The radiometric information for this study is from the AVHRR instrument onboard the National Oceanographic and Atmospheric Administration's NOAA-11 polar-orbiting spacecraft. The spacecraft was launched in September 1988, and the AVHRR instrument was operational between November 8, 1988 and September 13, 1994. The instrument records radiance in five spectral bands, or channels. Channel 1 records visible wavelengths (0.58 – 0.68 μm), channel 2 records near-infrared wavelengths (0.73 – 1.10 μm), and channels 3, 4, and 5 record longwave, or thermal, infraded wavelengths (3.55 – 3.93 μm, 10.3 – 11.3 μm, and 11.5 – 12.5 μm, respectively) (Kidwell, 1995). I required radiometric data from AVHRR channels 1 and 2, for the estimation of vegetation density and fractional cover, and channels 4 and 5, for the determination of surface temperature. Channel 3 data were incorporated in a cloud-screening algorithm. My objective was to obtain these data, together with the relevant view timing and geometric variables, for the 600,000 km^2 study region, over as many days as possible in the spring, summer, and fall seasons, for a sequence of years that included some unusually wet and some unusually dry periods.

Obstacles in the processing of raw AVHRR data include the geographic registration of imagery, the calibration of digital sensor output to radiometric information and then to reflectance vlaues (channels 1 and 2) or temperature values (channels 3, 4, and 5), the mosaicing of overlapping orbits, the derivation of view timing and geometry, and the contamination of imagery by cloud cover and cloud shadow (Eidenshink 1992, Kidwell 1995, Moody and Strahler 1994, Stoms et al. 1997). All of these problems are addressed and at least partially overcome in the 14-day composited AVHRR imagery available on CD-ROM for the conterminous U.S. from the U.S. Department of Interior, Geological Survey, National Mapping Division, EROS Data Center (EDC) in Sioux Falls, SD (Eidenshink 1992, Moody and Strahler 1994). The EDC database resamples the AVHRR data to a grid with cell spacing of 1 km. Rather than undertake all of these corrections to raw imagery myself, I used the EDC data.

Compositing refers to the selection of one value for each grid-cell in an image from a collection of possible values. In the case of EDC bi-weekly composited AVHRR data, the collection of possible values for each cell in a single bi-weekly image is from all of the satellite orbits which recorded data from that cell during the compositing period. The EDC data is composited using a maximum NDVI algorithm. The

primary purpose for this compositing algorithm is to reduce the occurrence of cloudy or hazy views, since both clouds and atmospherically scattered radiation tend to reduce the NDVI below its observed value through a clear atmosphere (Eidenshink 1992). The data provided for each compositing period include the calibrated reflectances/temperatures for the five AVHRR channels, a date image indicating the day of data acquisition for each grid-cell, a satellite zenith angle image, a solar zenith angle image, and a relative satellite-sun azimuth angle image. Details of the geographic registration, calibration and compositing process are given by Eidenshink (1992).

For the years from 1990 through 1994, I extracted a spatial subset of the biweekly composite imagery (all channels and ancillary information) from all compositing periods between April and October. The EDC dataset uses an Alber's conformal equal area projection that is appropriate to imagery over the entire conterminous U.S. I reprojected the spatial subset to a grid with 1 km cell spacing and a Lambert's azimuthal equal area projection centered on my study area, retaining only those grid cells within the region bounded by longitudes 110 °W to 120 °W and latitudes 42 °N to 49 °N, using nearest neighbor resampling. I wrote a program to perform the subsetting and nearest-neighbor resampling. The projection transformation was accomplished with the General Cartographic Transformation Package, a series of software subroutines released by the U.S. Geological Survey, which I incorporated in my program.

In preliminary investigations of the seasonal patterns of reflectances and temperatures, I noticed that several compositing periods in 1990 had unusually high temperatures. On consultation with EDC personnel it seemed that there may be a data scaling error for the 1990 data on the CD-ROM (J. Eidenshink, personal communication, 1997), and so I eliminated that year's data from further investigation. The remaining four years included two warm, dry years (1992 and 1994), one cool, wet year (1993), and one year with precipitation and temperature close to the 5-year mean (1991), so I felt that dropping 1990 (another rather typical year) would not impair my analysis.

5.4.1 Cloud and snow detection and screening

Another conclusion from preliminary investigation was that significant cloud contamination remained in the composited data, as has been noted by others (Moody and Strahler 1994, Eklundh 1995).

Since cloud contamination has a strong influence on radiometric temperature as observed by the sensor (Ou et al. 1996), and since the observations of radiometric surface temperature are central to my analysis, it was imperative that as much cloud contamination be removed as possible, including contamination by thin cirrus clouds, which are relatively transparent but which influence the observed temperature (Ou et al. 1996).

In addition to cloud contamination, I noticed a pattern of low surface temperature and high surface reflectance, localized over high terrain and diminishing in spatial extent through the spring. I interpreted this pattern as mountain snow cover. Since the existence of a snow pack, whether underneath an open canopy or on top of short vegetation, will have a strong influence on the observed surface temperature, and since I was not confident in my ability to generate an accurate prediction for snowpack surface temperature, I tried to eliminate snowy grid cells along with cloud contaminated grid cells from further analysis. In fact, I was not able to find in the literature or derive on my own a general method for discriminating between cirrus cloud contamination and snow cover in the AVHRR data. Other authors have commented on the difficulty of discriminating complete or partial cloud cover from snow in AVHRR data (Gutman et al. 1994, Cihlar 1996).

The problem of cloud detection in AVHRR imagery has been addressed at length in the literature, and the importance of identifying and removing cloud contamination from imagery used in land-surface analysis has been stressed by many authors (Saunders and Kriebel 1988, Eck and Kalb 1991, Gutman et al. 1994, Cihlar 1996, Simpson and Gobat 1996). A principal distinction between methods proposed for the removal of cloud contamination is whether they are designed for application to composited or uncomposited imagery. The methods applied to uncomposited imagery (single orbital samples) typically rely, in part, on estimates of spatial heterogeneity. These methods are not applicable to composited imagery, since the composite mosaic introduces spatial patterns unrelated to cloudiness. Detection in composited imagery must rely on values for individual grid cells, and most methods are based on a single temporal sample, although methods have been proposed that use the temporal sequence of observations at a single point.

For methods using a single temporal sample, four basic cloud detection criteria have been suggested. First, the visible reflectance values (channel 1) for cloudy grid cells are usually higher than for vegetation or bare soil, and so a channel 1 reflectance threshold for clouds can be established. Second, the ratio between near infrared and visible reflectance (Q) is generally lower for clouds than for vegetation or bare soil, but is typically higher than for water. In order to retain water grid cells in the filtered database, two thresholds are set, but if water can be excluded then only a single (lower) limit for Q needs to be established. Third, the radiometric temperature from AVHRR channel 4 or channel 5 is typically lower for clouds than for vegetation or bare soil, and so a lower temperature threshold can be established. Fourth, and finally, the difference between radiometric temperatures in channel 4 and 5 (dT_{45}) is high for optically thin clouds or for partially-cloudy gridcells. Such partial cloud cover can be confused with soil and vegetation if only the first three criteria are used. In addition to these criteria, I found that the difference between channel 3 and channel 4 radiometric temperature (dT_{34}) was also diagnostic of partial cloud cover at the edges of cloudy regions detected by the other criteria.

Given these criteria, the basic problem becomes one of selecting appropriate thresholds. One difficulty in parameterizing the single-channel temperature threshold is that cloud temperature can vary greatly depending on air temperature, and surface air temperature is generally not known over the spatial domain of the satellite data. In particular, for areas with strong diurnal variation in surface air temperature it may be difficult to assign an appropriate threshold value (Gutman et al. 1994). Here I take advantage of the instantaneous estimate of air temperature at the time of satellite data acquisition to normalize this threshold. Surface energy balance theory suggests that in very few cases will the surface temperature be lower than the air temperature under sunny conditions, and so observed surface temperatures substantially below air temperature are likely cloud-contaminated. Water surfaces are the primary exception, being usually cooler than air temperature during the day, but I am not concerned with water surfaces in this study and so they may be safely confused with cloudy gridcells and eliminated from the dataset. A similar approach, using interpolated air temperatures with very coarse spatial resolution (1° x 1°) and a general circulation model to estimate the diurnal temperature patterns, was used in a study by Eck and Kalb (1991).

Based on values suggested in the previously cited studies and on a graphical analysis of instantaneous air temperature and all five channels of AVHRR data for each 14-day composite, I arrived at the cloud-detection parameters given in Table 7. A gridcell was classified as cloud-contaminated and dropped from further analysis if it met one or more of the criteria in Table 7.

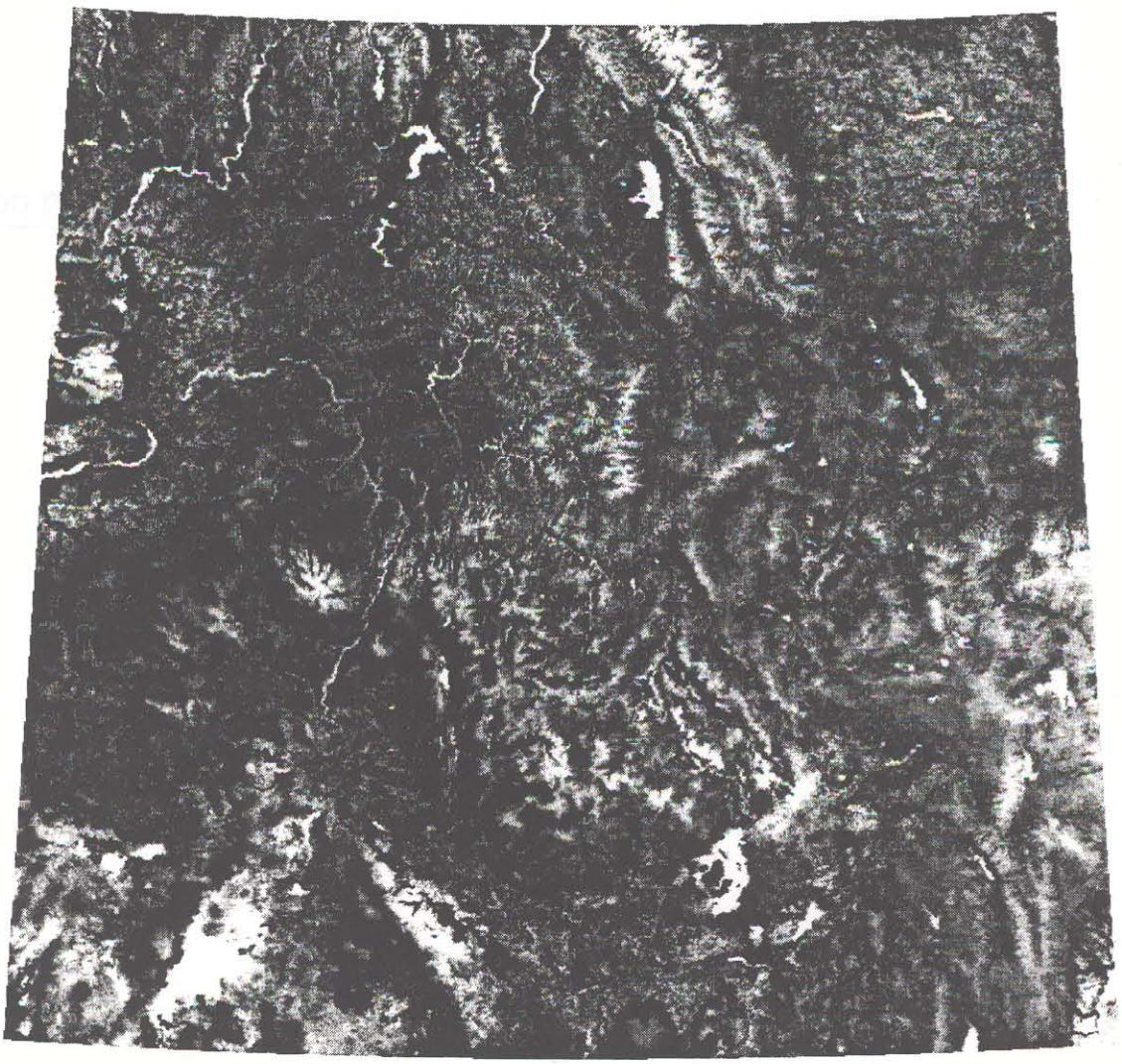
Table 7. Cloud detection criteria.

1) Channel 1 reflectance > 0.20
2) $Q < 1.20$
3) $dT_{45} > 4.5$ (K) or $dT_{45} < -1.5$ (K)
4) $dT_{34} > 15$ (K)
5) $(T_4 - T_{air}) < -5$ (K)

The number of composite periods that were dropped due to cloud and/or snow contamination, for each gridcell, is shown as a gray-scaled image in Figure 14. The total number of composite periods in the analysis is 57, and the average over all cells of the number of periods dropped due to cloud/snow contamination is about 7. As shown in Figure 14, some areas are habitually identified as cloud/snow contaminated, particularly over water bodies, over barren land (the lava fields north of American Falls, ID, are a good example), and over the highest elevation terrain. The histogram of the number of cloud/snow contaminated periods over all gridcells is shown in Figure 15, where it can be seen that the great majority of cells have 20 or fewer contaminated periods. The maximum number of contaminated periods for a single cell was 57, and the minimum was 0.

5.4.2 Satellite zenith angle screening

Subsequent to the cloud detection and screening process, I screened the remaining cells in each composite period to remove from further analysis any cell with the satellite zenith angle (SATZ) greater than 50°. Large angles are problematic because the sensor's field of view increases with increasing SATZ, making accurate registration difficult and introducing bias from surrounding terrain in regions of strong spatial heterogeneity in reflectances or emittances (Moody and Strahler 1994). The number of composite periods that were dropped due to high SATZ, for each gridcell, is shown as a gray-scaled image in Figure 16. The average over all cells of the number of periods dropped due to high SATZ is about 5. Areas that



Number of cloud/snow contaminated periods

0 5 10 15 20 25(+)

Figure 14

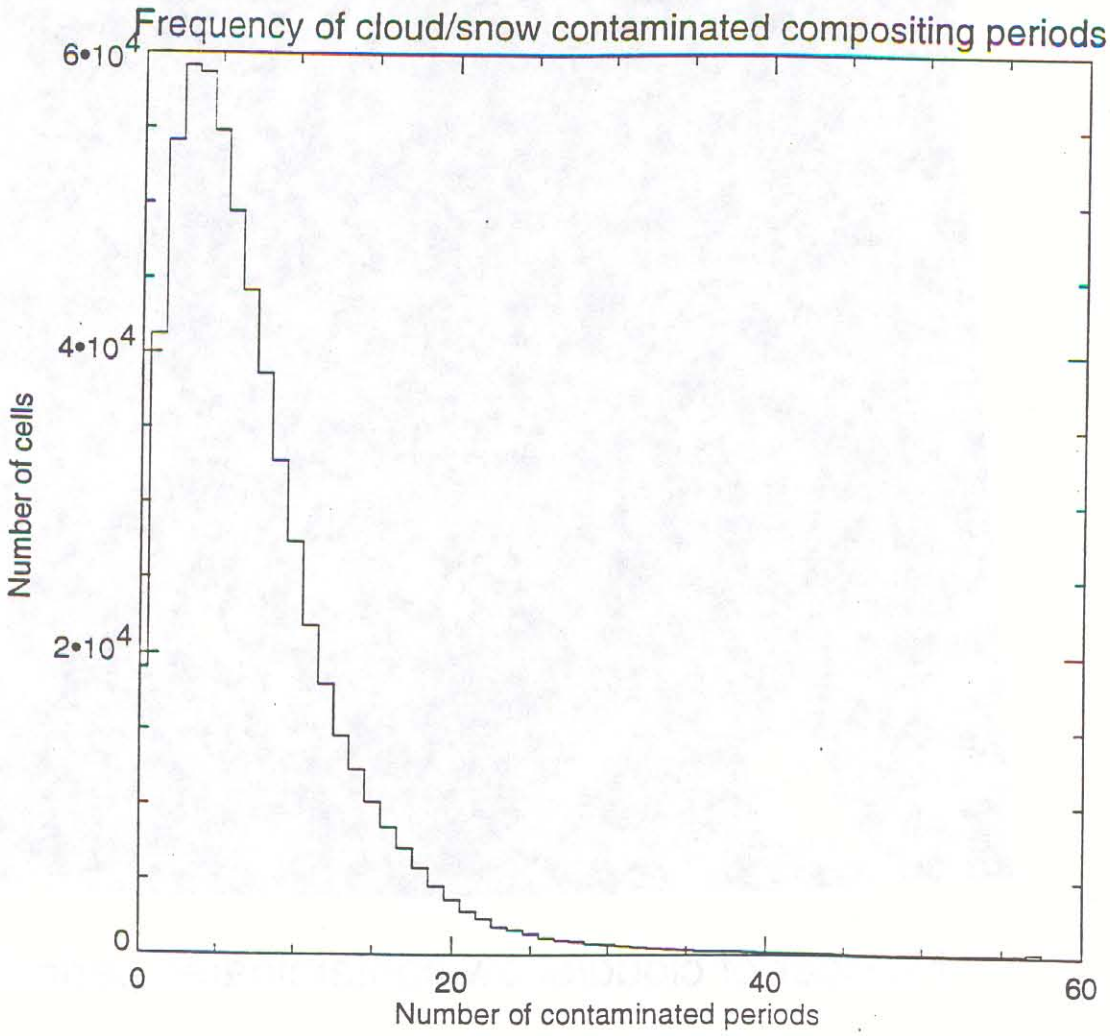

Frequency of cloud/snow contaminated compositing periods
6•10^4
4•10^4
2•10^4
0
Number of cells
0
20
40
60
Number of contaminated periods

Figure 15

tend to have high satellite zenith angles in the composited imagery include water bodies, steeper west-facing slopes, and edge regions between dense and sparse vegetation (e.g., around the irrigated agricultural regions in the Snake River Valley). The histogram of the number of cloud/snow contaminated periods over all gridcells is shown in Figure 17, where it can be seen that the great majority of cells have 15 or fewer contaminated periods. The maximum number of contaminated periods for a single cell was 26, and the minimum was 0. Considering both cloud/snow contamination and high SATZ, the average number of periods dropped per gridcell was about 11, or about 20% of the original dataset.

5.4.3 Estimation of time of day of satellite data acquisition

The local solar time of day of satellite overpass and data acquisition (t_{so}, hours) for each cell in the original composited data was not given explicitly. I estimated it using the solar zenith angle, which was given, and the same equations for earth-sun geometry as used in MTCLIM and Daymet for the estimation of radiation. These equations allow a prediction of the hour angle (h, 0° = noon) from the solar zenith angle (z), the cell's latitude (l), and the declination of the Earth's rotational axis with respect to the principal solar plane (decl), as:

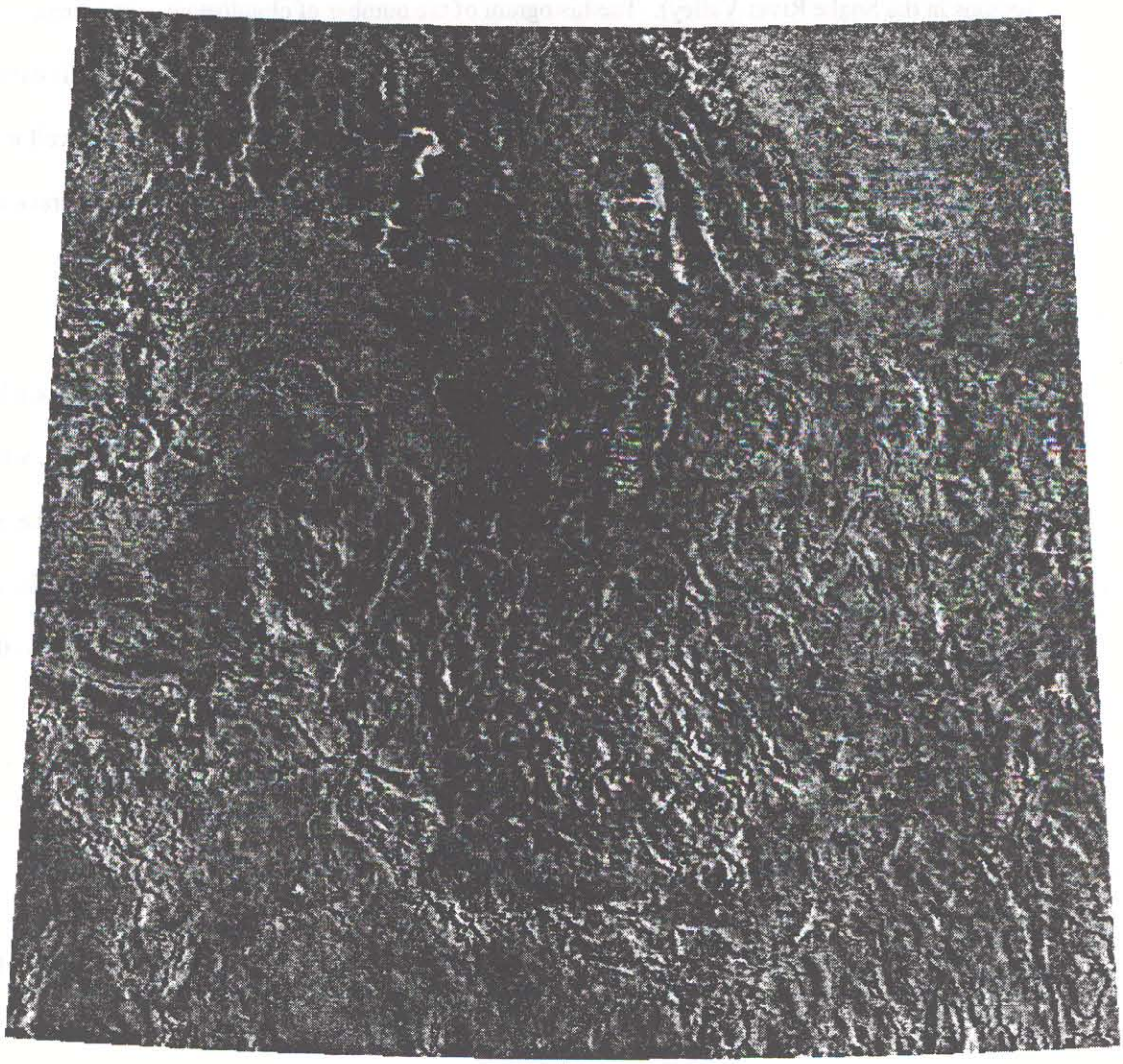
Eq. 23 $h = \text{acos}[(\cos(z) - \sin(l) \sin(decl)) / (\cos(l) \cos(decl))]$

$decl = -23.45 \cos[(yearday + 11.25)\ 0.9863]$

$t_{so} = 12.0 + (1/15)\ h$ (to convert hour angle to hours)

Because the original data store z as whole degrees, the value for z entering Eq. 23 has a potential error of ±0.5°, which causes some error in the estimate of h and therefore t_{so}. As shown in Figure 18, the error in t_{so} is small, on the order of 5 minutes, for typical growing-season ranges of the variables. The largest errors occur near solar noon and in the spring and fall. Note that this method is unable to distinguish between true values of t_{so} occurring the same distance away from solar noon in opposite directions, i.e., z has the same value at 11:00 and at 13:00 hours. Since the times of acquisition are all after noon this does not impede the current application.

The orbital parameters of the NOAA-11 spacecraft have changed slowly since its launch, resulting in the local solar time of equator crossing drifting to later times (Gutman et al. 1994). Orbital drift resulted



Number of periods with excessive SATZ

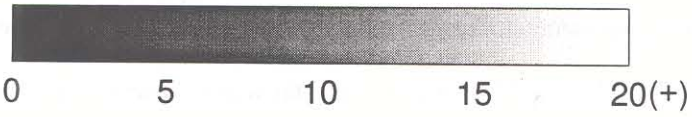


Figure 16

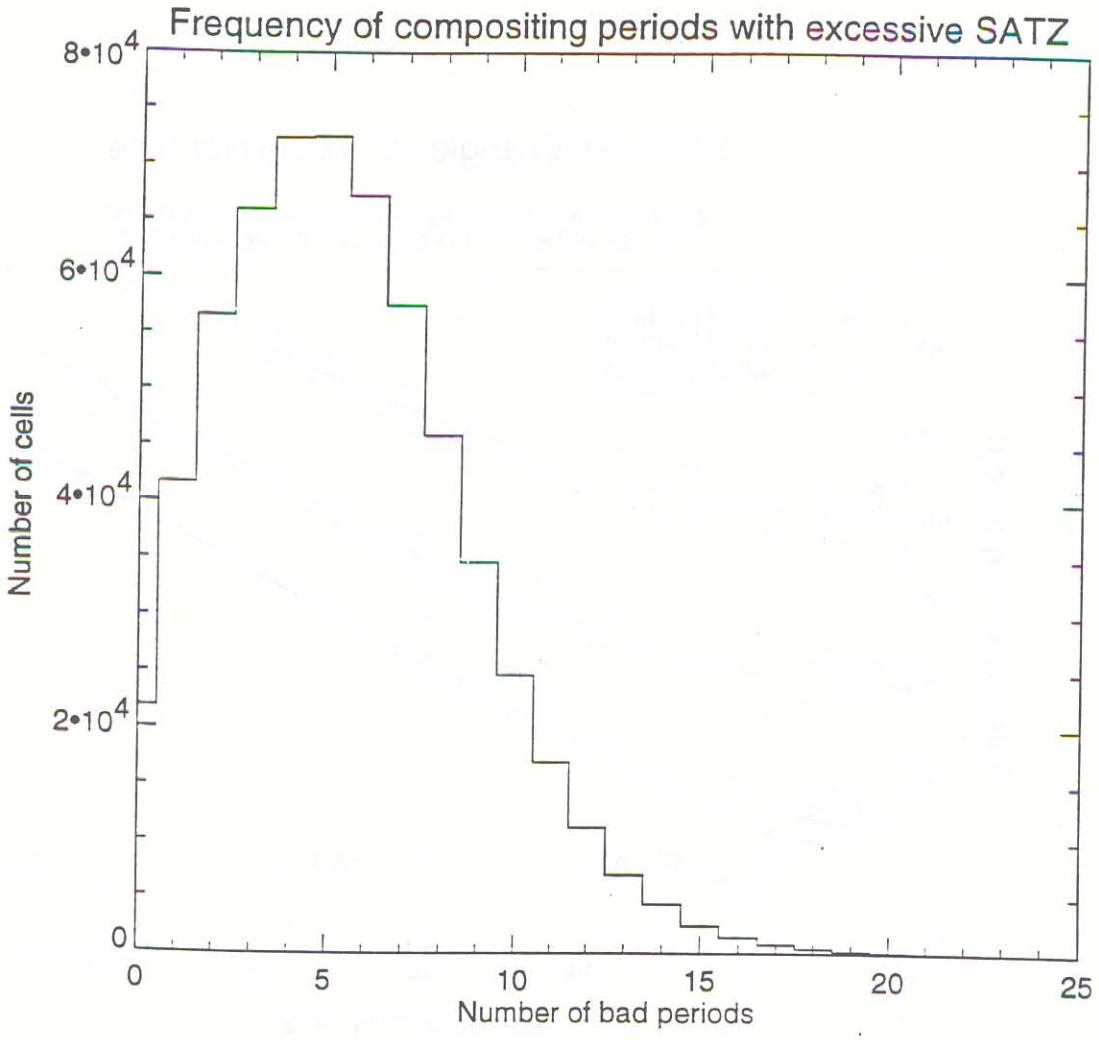

Frequency of compositing periods with excessive SATZ
8•10^4
6•10^4
4•10^4
2•10^4
0
Number of cells
0
5
10
15
20
25
Number of bad periods

Figure 17

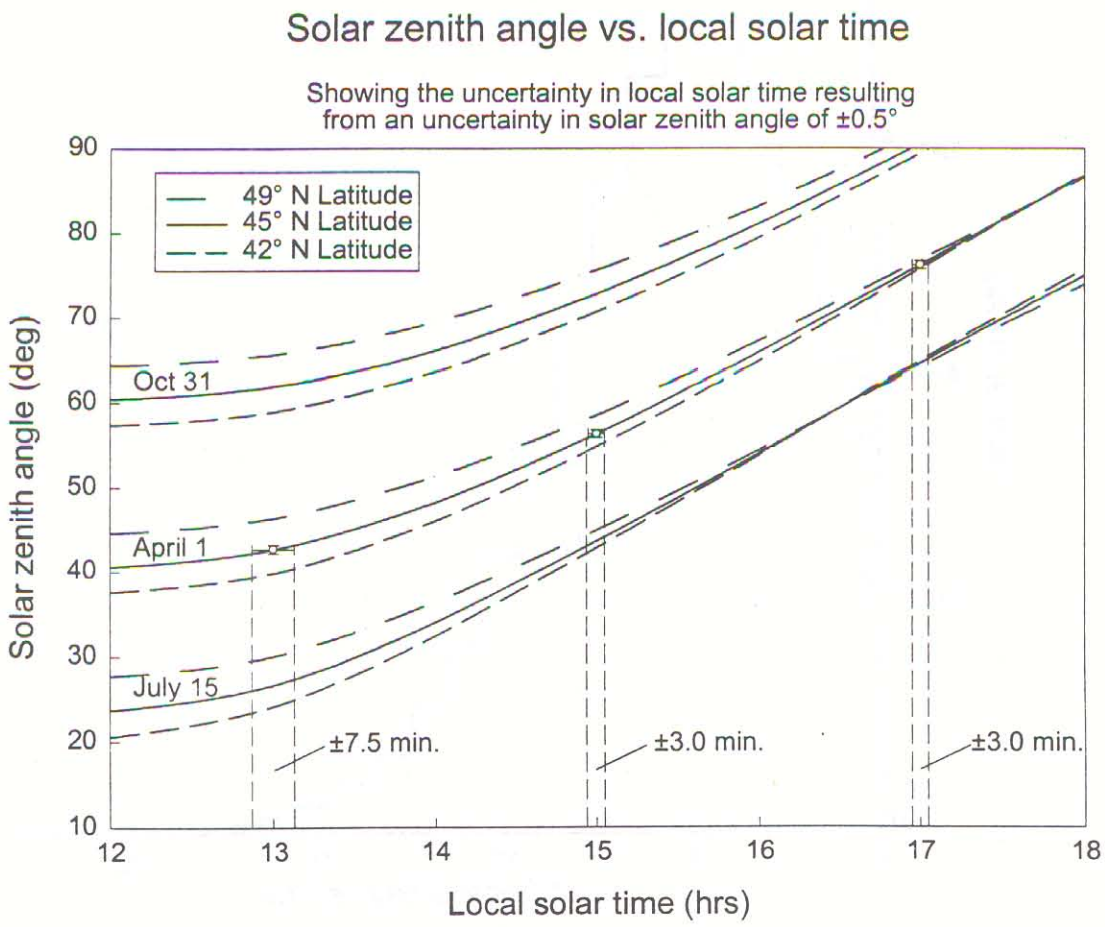

Solar zenith angle vs. local solar time
Showing the uncertainty in local solar time resulting from an uncertainty in solar zenith angle of ±0.5°
49° N Latitude
45° N Latitude
42° N Latitude
Oct 31
April 1
July 15
±7.5 min.
±3.0 min.
±3.0 min.
Solar zenith angle (deg)
Local solar time (hrs)
90
80
70
60
50
40
30
20
10
12
13
14
15
16
17
18

Figure 18

in a significant variation in t_{so} over the period 1990-1994 (Privette et al. 1995). The average value of t_{so} for each growing season compositing period is plotted in Figure 19, showing a shift of more than three hours. The shift in t_{so} is associated with changing solar zenith and satellite zenith angles (not shown), all of which influence the observed surface temperatures.

Having estimated t_{so} for each good gridcell in each composite image, I proceeded to generate composited, instantaneous values of air temperature, incident solar radiation, and vapor pressure deficit, using the methods already described.

5.4.4 Estimation of sun-sensor angle

The plane angle at a particular gridcell between the sun and the sensor at the time of data acquisition (θ) should have an influence on the observed surface temperatures. When this angle is zero all shadowing in the sensor's field of view is eliminated, whether from terrain or canopy or soil roughness elements. This is true regardless of the zenith or azimuth angles. As θ increases the amount of shadow in the field of view increases, until for $\theta = 180°$ the entire view should be shadowed (with the sensor looking directly into the sun). The original dataset did not include θ explicitly, but it did include the solar zenith angle (z), the satellite zenith angle (Z), and the relative azimuth between the sun and sensor (RELA, the angle formed by the horizontal projection of the two zenith angles). From these angles I derived the following expression for θ:

Eq. 24 $\theta = \text{acos}[\cos(z)\cos(Z) + \sin(z)\sin(Z)\cos(RELA)]$

5.5 Estimation of land surface temperature using a split-window algorithm

The AVHRR thermal IR sensors for channels 4 and 5 record radiances in a spectral region where the atmosphere is mostly transparent (10.3 – 12.5 μm, transmissivity of 0.7-0.8). Much of the thermal energy emitted from the surface in these wavelengths can reach the sensor. Atmospheric gases (primarily H_2O, CO_2, and O_3) absorb some of this energy, and also emit energy in these same wavelengths according to their vertical distribution, the temperature profile of the atmosphere, and their emissivity (0.2-0.3, or 1.0-transmissivity). Without knowing the vertical distribution of these gases in the atmosphere and the atmospheric temperature profile, measurement in a single channel can only provide an approximate

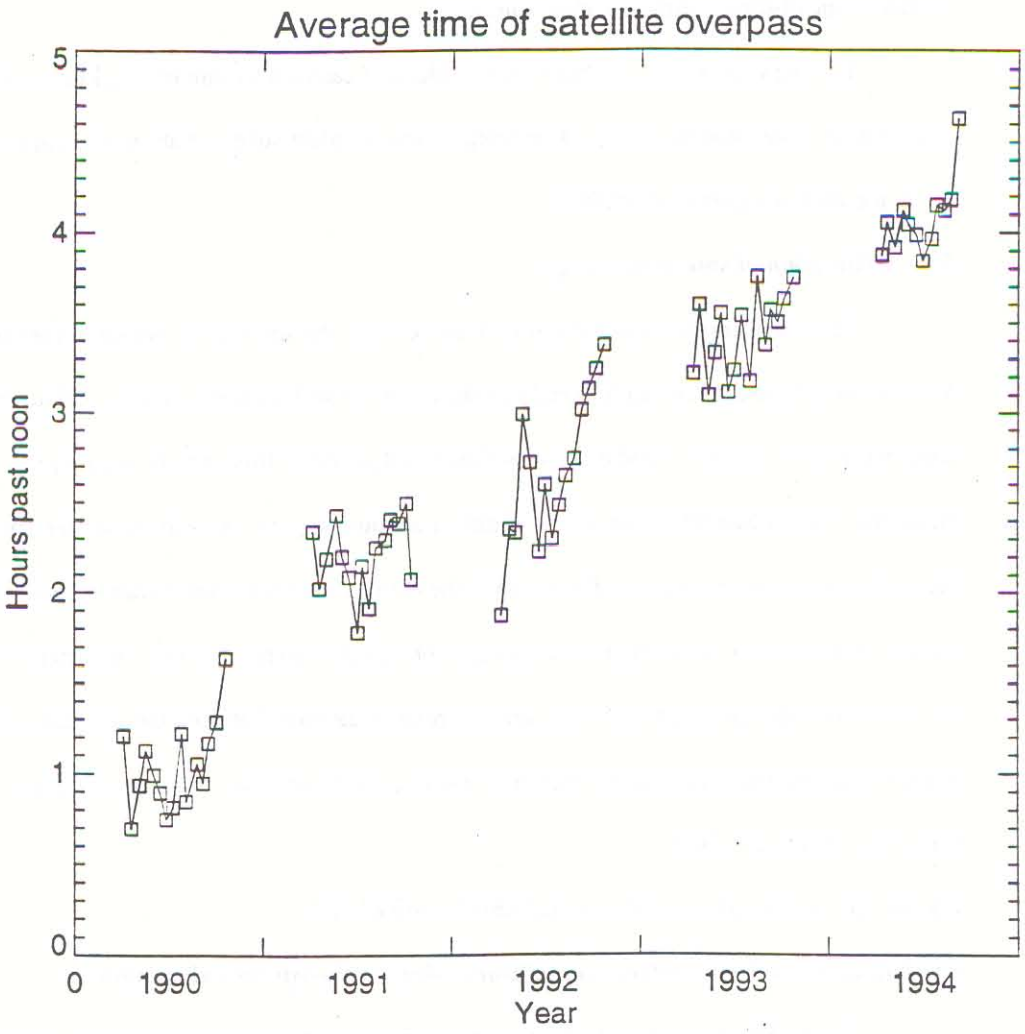

Average time of satellite overpass
Hours past noon
0
1
2
3
4
5
0
1990
1991
1992
1993
1994
Year

Figure 19

estimate of true surface temperature (worst-case errors for water surface temperature of around 7 K, Sobrino et al. 1991). It has been demonstrated, for the case of surface temperature retrievals over the ocean or other water body, that when observations in two closely-spaced spectral bands are available, a correction for this atmospheric interference is possible without having to know the distribution of absorbing gases or the atmospheric temperature profile (Prabakhara et al. 1974). This is known as the split-window method for temperature retrieval, and it relies on differential atmospheric transmissivities in the two observation spectral bands. This method is particularly valuable for temperature retrievals over water surfaces, because the emissivity of water in the 10-12 μm spectral range is uniform and very close to 1.0 (about 0.99). Typical emissivity values for the land surface vary between 0.9 and 0.99 (Salisbury and D'Aria 1992), with significant differences occurring between emissivities in the AVHRR channel 4 (ε_4) and channel 5 (ε_5) spectral bands (Rubio et al. 1997). These factors make the application of a split-window algorithm difficult, since the spectral variation in emissivity can not be directly separated from the effects of differential atmospheric transmissivity (Price 1984, Becker 1987).

In general, the solution to this problem requires some knowledge of the average surface emissivity as well as the spectral variation in emissivity between the two observation spectral bands. Of these, the difference ($\Delta\varepsilon$, defined here as $\varepsilon 4$-$\varepsilon 5$) is the more important in terms of accurate temperature predictions from a split-window algorithm (Becker 1987). Coll et al. (1994b) present a form of the split-window method which is applicable to midlatitude atmospheric conditions and which contains a term for the correction of surface emissivity effect, depending on knowledge of both ε_4 and $\Delta\varepsilon$. In a second paper by Coll et al. (1994a) a method is presented by which an estimate of $\Delta\varepsilon$ can be made from observed differences between brightness temperatures in AVHRR channels 4 and 5, given some knowledge of the atmospheric state. They develop their method with the use of radiosonde measurements of vertical temperature and humidity profiles, and make use of an atmospheric radiative transfer model to derive certain parameters.

I adopted the general approach outlined by Coll et al. (1994a) for the estimation of $\Delta\varepsilon$, but I made certain key simplifications to avoid using radiosonde data or a radiative transfer model. Various authors

have suggested that both ε_4 and $\Delta\varepsilon$ should be closely related to the fractional cover of vegetation over land surfaces, since living vegetation has generally high and spectrally uniform emissivity in this spectral region while vegetation litter, soil, and rock typically have lower average emissivity and larger absolute values for $\Delta\varepsilon$ (Caselles et al. 1997, Rubio et al. 1997, Salisbury and D'Aria 1992, Price 1984). The influence of errors in ε_4 on eventual surface temperature retrievals has been shown to be rather small (Becker 1987), and so I made the simplifying assumption that fractional vegetation cover was linearly related to ε_4, and that NDVI could be used as a linear surrogate for fractional vegetation cover (Asrar et al. 1992). Data collected by van de Griend and Owe (1994) show a linear relationship between emissivity and NDVI, and a theoretical basis for this relationship has been developed by Valor and Caselles (1996). I estimate ε_4, using a linear function, fit by hand to the data of van de Griend and Owe (1994) (see Valor and Caselles 1996, Figure 6), as:

Eq. 25 $\varepsilon_4 = 0.99 - 0.09\ [(0.7 - \mathrm{NDVI})/0.6]$

The estimation of surface temperature is more sensitive to error in $\Delta\varepsilon$, with an accuracy of ± 0.005 required in order to keep temperature errors below 0.4 K (Coll et al. 1994b), and so I took more care in its estimation. Coll et al. (1994a) present the following equation for calculating $\Delta\varepsilon$:

Eq. 26 $\Delta\varepsilon = [T_4^* - T_5^* - (1.0 - \varepsilon_4)(b_5 - b_4)] / b_5$

where T_4^* and T_5^* are, respectively, the brightness temperatures that would be observed immediately above the surface for channels 4 and 5 (i.e. atmospherically corrected values, assuming surface emissivity of unity), and b_4 and b_5 are parameters having units of temperature. In Coll et al. (1994b), the atmospherically corrected brightness temperatures are given as:

Eq. 27 $$T_i^* = T_i + \left(\frac{1-\tau_i}{\tau_i}\right)\left(T_i - T_a^{\uparrow}\right)$$

where i represents the channel (4 or 5), τ_i is the atmospheric transmittance in channel i, and $T_a^{\uparrow}$ is the radiative temperature of the atmosphere in the upward direction.

Values for b_4 and b_5 depend on atmospheric conditions and on the sensor characteristics, and are given by:

Eq. 28 $$b_i = \frac{T_i^*}{n_i} + y\left(\frac{n_i - 1}{n_i}T_i^* - T_a^{\downarrow}\right)(1 - \tau_i)$$

where n_i and y are parameters that depend on the sensor characteristics ($n_4 = 4.673$ and $n_5 = 4.260$ for NOAA-11 (Sobrino and Caselles 1990), and $y \cong 0.51$ (Schmugge et al. 1991)), and $T_a^{\downarrow}$ is the temperature of the atmosphere corresponding to the downward radiance emitted by the atmosphere.

Using a radiative transfer model and considering only the influence of water vapor, Coll et al. (1994b) showed that τ_4 and τ_5 are related to each other through the following two relationships:

Eq. 29 $$1 + 0.5\ W = (1 - \tau_4)/(\tau_4 - \tau_5)$$

$$0.0474\ W = \tau_4 - \tau_5$$

where W is the total column water vapor content of the atmosphere (g cm^{-2}). These two expressions can be solved simultaneously to give τ_4 and τ_5 as functions of W:

Eq. 30 $$\tau_4 = 1.0 - 0.00474\ W - 0.000237\ W^2$$

$$\tau_5 = 1.0 - 0.00948\ W - 0.000237\ W^2$$

I obtained a dataset of daily, gridded, total column atmospheric water vapor content from the National Aeronautic and Space Association's Water Vapor Project (NVAP). This dataset has global coverage on a 1° by 1° grid, and includes daily observations for two of the years in my study, 1991 and 1992. The data are derived from a combination of radiosonde and satellite observations. I extracted a spatial subset of the data and used a cubic convolution algorithm (Research Systems, Inc. 1995) to interpolate the coarse grid over my study area.

The final variables required to estimate $\Delta\varepsilon$ using Eq. 26 are the apparent atmospheric temperatures for upward and downward atmospheric radiances, $T_a^{\uparrow}$ and $T_a^{\downarrow}$. I estimated $T_a^{\downarrow}$ using an empirical expression given by Idso (1981), which relates the apparent atmospheric temperature to the instantaneous air temperature (T_{air}, K) and vapor pressure (e_o, Pa) as:

Eq. 31 $$T_a^{\downarrow} = T_{air}\sqrt[4]{0.7 + 0.00000595 \cdot e_0 \cdot \exp(1500.0/T_{air})}$$

where T_{air} was obtained from the composited instantaneous air temperature database described earlier, and e_0 was estimated from the composited daily minimum temperature. Sobrino et al. (1991) note that, since

water vapor is primarily responsible for the atmospheric radiance in the spectral region in question, and since it is concentrated at lower layers, the downward radiance and apparent temperature will be slightly greater than the upward radiance and apparent temperature. Lacking any justification for a correction between these two temperatures, I assumed that they were the same, which is the same assumption made by Sobrino et al. (1991) in their study. By substituting from Eqs. 27, 28, 30, and 31 into Eq. 26, I arrived at an estimate of $\Delta\varepsilon$ for each grid cell with good data for each compositing period in 1991 and 1992 (the years for which I had data for W).

The possible sources of error in this process are numerous. Perhaps most significant is the lack of spatial resolution in the data for W, and its apparently noisy temporal patterns (not illustrated). Of secondary importance are the errors related to subdaily predictions of air temperature and daily predictions of humidity. I have little confidence, therefore, in the estimated values for $\Delta\varepsilon$ for individual grid cells for single compositing periods. On the other hand, I know that the biases in the subdaily and daily surface meteorological data are quite low, and I can only assume that the same is true for data from the NVAP dataset, in which case the temporal average values of $\Delta\varepsilon$ for individual grid cells should be meaningful. I tested the two-year average values of estimated $\Delta\varepsilon$ against the suggestion from observation and theory that they should have values close to 0 for grid cells with high NDVI, and should have increasingly larger negative values for lower NDVI, approaching approximately –0.02 for bare soils and rock (Rubio et al. 1997, Salisbury and D'Aria 1992). Average values of $\Delta\varepsilon$ are plotted against values of NDVI averaged over the same two-year period in Figure 20, together with the regression line predicting $\Delta\varepsilon = f(\text{NDVI})$:

Eq. 32 $\Delta\varepsilon = -0.02938 + 0.04957\ \text{NDVI}$ ($R^2 = .254$)

Although this relationship explains only 25% of the variance in predicted $\Delta\varepsilon$, it is highly significant ($P<0.01$) and consistent with theory, giving $\Delta\varepsilon = -0.003$ for NDVI = 0.6 (dense canopy), and an intercept at NDVI = 0.0 which is in the range typical of bare soils and rock. The mean absolute prediction error from this relationship is 0.006, which indicates that the likely error in surface temperature introduced by using Eq. 32 in place of the estimated values of $\Delta\varepsilon$ would be about 0.6 K.

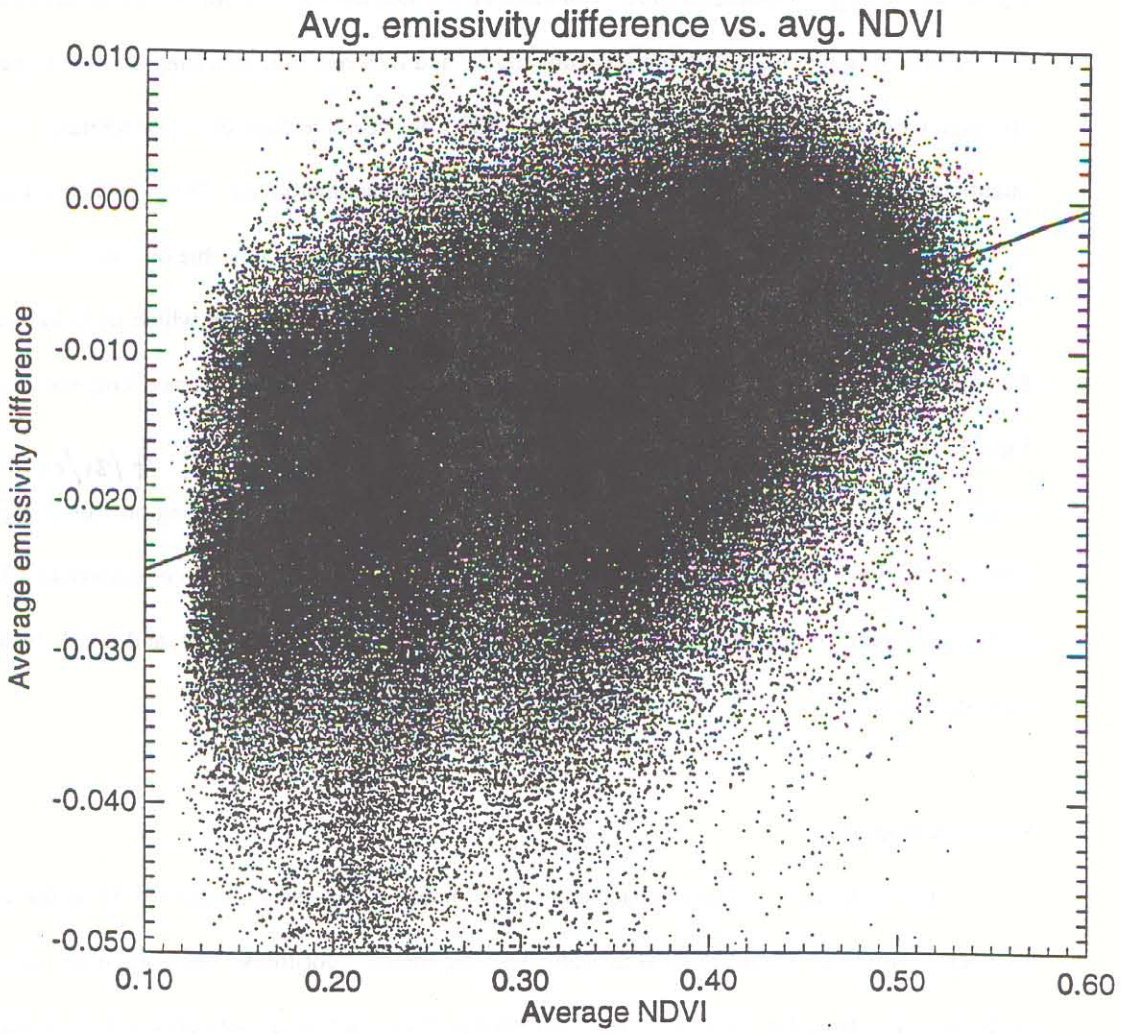

Avg. emissivity difference vs. avg. NDVI
Average emissivity difference
0.010
0.000
-0.010
-0.020
-0.030
-0.040
-0.050
0.10
0.20
0.30
0.40
0.50
0.60
Average NDVI

Figure 20

Given the magnitude of errors in air temperature estimates (1.5 to 2.5 K), and postulating that a substantial amount of the variance exhibited in Figure 20 is due to bias in estimation of $\Delta\varepsilon$ at a single point, I replaced the original estimates of $\Delta\varepsilon$ with values estimated from Eq. 32 for the remainder of my analysis. The purpose of going to the trouble of estimating $\Delta\varepsilon$ and then resorting to a simple linear regression is that, although a general relationship between $\Delta\varepsilon$ and NDVI has been mentioned in the literature, I could find no quantitative examples of how such a relationship would be parameterized. Through this investigation I was able to derive a simple, plausible relationship, relying as much as possible on a theory-driven method.

I used the following generalized form of the split-window equation which includes separate terms for atmospheric and emissivity effects (Coll et al. 1994b) to predict actual surface temperature (T_s):

Eq. 33 $T_s = T_4 + [1.0 + 0.58 (T_4 - T_5)](T_4 - T_5) + A(1-\varepsilon_4) - B\,\Delta\varepsilon + 0.51$

Where A =50.0 is a constant that is valid globally, and B is a constant depending on climatological atmospheric type. Coll et al. (1994b) suggest B = 150 for mid-latitude winter conditions and B = 75 for mid latitude summer conditions. Since my analysis covers the spring, summer, and fall, I chose B = 100 as a constant value.

5.6 Prediction of r_h

For each good grid-cell in each composited image, I generated a value for ΔT as the difference between T_s from Eq. 33 and T_a from the subdaily temperature algorithms. The core of my analysis is the prediction of r_h using Eq. 18, and for this I required a single estimate at each grid-cell of the influence of R_n on ΔT, other factors assumed invariant. I used a multiple regression to obtain this estimate, by including R_n as one independent variable, in addition to several other variables that I knew from preliminary investigations to influence ΔT. The primary purpose of including other variables, instead of performing a simple linear regression on R_n, is to isolate the partial effect of R_n on ΔT in the face of variation due to other influences. A secondary purpose is to be able to make at least some qualitative statements about the influence of these other variable on the observed variation in ΔT.

The variables that I elected to include in the multiple regression, in addition to R_n, are NDVI, VPD, an index of antecedent precipitation (P_{pre}), and the plane angle between the sensor and the sun with respect to the observation point (θ). VPD is included because of its theorized role in determining ΔT, according to Eq. 17. The antecedent precipitation is included as a crude surrogate for the variation in r_c that is expected as the result of changing water availability. NDVI is included because it is assumed to be linearly related to fractional vegetation cover (Asrar et al. 1992), and because the difference in energy balance components between canopy elements and the underlying soil or litter has been shown to result in an approximately linear mixing relationship for sensor fields of view large enough to integrate contributions from each in a single recording (Norman et al. 1995, Caselles et al. 1992, Gillespie 1992). This effect of fractional cover has been shown to occur in conjunction with variation in the illumination and viewing geometry (Caselles et al. 1992, Kimes et al. 1980). Because of the theorized importance of sunlit vs. shaded canopy and soil fractions in determining the radiance observed at the sensor (Caselles et al. 1992, Sobrino and Caselles 1990), I included the sun-sensor angle θ as the final independent variable in the multiple regression for ΔT.

The instantaneous net radiation at the time of data acquisition (R_n) depends on the incident solar radiation, the surface shortwave albedo, and the ground heat flux. I ignored the angular effects of surface reflectance, and took the albedo (α) to be the average reflectances from AVHRR channel 1 (visible band) and channel 2 (near infrared band) converted from percent to a proportion. I ignored the contribution of incident solar radiation to heat storage in the soil (i.e., $G = 0$ in Eq. 5).

The antecedent precipitation index at each grid-cell for each composited day was calculated as the weighted total precipitation over the previous 20 days, with weights decreasing linearly away from the prediction day. Estimated daily precipitation from the Daymet database was used to generate this index.

For each grid-cell, a single multiple regression was performed, using as samples all of the composited daily observations that were not rejected due to cloud contamination or scan-angle violations. A prediction for r_h was then performed for each grid-cell according to Eq. 18, using the regression coefficient for R_n as $d\Delta T/dR_n$ and calculating c_1 as the composited average air density (ρ) multiplied by c_p.

5.7 Evaluation of predicted r_h

I obtained a spatial database of landcover classification and qualitative classes of canopy closure from the Wildlife Spatial Analysis Laboratory at the University of Montana, for the purpose of comparing the predicted values of r_h against independent estimates of two of the physical factors hypothesized to exercise significant control over this parameter, namely, characteristic leaf dimension and fractional canopy cover. From a large database covering all of Montana and parts of Idaho, I obtained a series of spatial subsets, chosen to include a variety of topographic, floristic, and land-use patterns (see Figure 22 for the locations of these spatial subsets and their reference numbers). These landcover data are the result of supervised classifications performed on Landsat Thematic Mapper imagery, guided by observations of landcover, canopy closure, and other parameters from many thousands of field plots. The original database has a nominal spatial resolution of 30 m, and is grouped into polygons of grid-cells having similar spectral characteristics prior to final classification (R. Redmond, personal communication, 1997). The spatial subsets used in this study were reprojected and resampled to a regular 100 m grid, using nearest neighbor resampling. The size of the original polygons is such that this resampling results in very little loss of spatial detail.

From approximately 90 landcover classes in the original dataset, I merged classes to obtain the list shown in Table 8. The original data set contains a categorical canopy closure variable for forest, shrub, and grass classes that ranges from 1 (open canopy) to 3 (closed canopy). These rankings were assigned on the basis of spectral vegetation indices (from single-date Thematic Mapper data) and ground observations of canopy closure. The total frequency across all five spatial subsets is given for each merged landcover type in Table 8, as well as the frequencies of different canopy closure classes where appropriate.

Table 8. Frequencies of landcover and canopy closure classes

All frequencies are given as km^2.

		Canopy closure classes			
		(low)		(high)	no canopy
Covertype	Total	1	2	3	closure class
Conifer forest	10759	2246	6284	2228	--

Broadleaf forest	179	45	121	13	--
Mixed conifer/broadleaf	98	62	36	--	--
Xeric shrubs	1868	1682	133	44	9
Mesic shrubs	1026	243	314	447	22
Grassland/meadow	3339	1117	1112	710	409
Dry agriculture	115				115
Irrigated agriculture	729				729
Standing burnt forest	1290				1290
Barren	783				783

Landcover classes were merged in a way that was consistent with the hypothesis that characteristic leaf shape and vertical canopy structure would be the strongest determinants of landscape-level values for r_h.

After merging landcover classes, I aggregated the 100 meter resolution data to coincide with the 1 km^2 cells used in the rest of the study. This aggregation was performed by assigning to the new grid cell the class with the most frequent occurrence within it from the original dataset. In the case of a tie between class frequencies, one was selected at random. The same procedure was performed for both cover type and canopy closure data.

As discussed in the literature review, the assignment of relevant values of r_h for coarse resolution gridcells characterized by sub-grid scale landcover heterogeneity, such as is true for the ~1 km^2 grid cells from the AVHRR database, is a difficult problem. It is generally supposed that the grid-scale value of r_h relevant to aggregated fluxes of heat and water vapor should be inversely related to the fractional vegetation cover within a grid cell (Qualls and Brutsaert 1996, Norman et al. 1995). If that is the case, then the predicted values of r_h from Eq. 18 should show an inverse relationship with fractional cover, since they rely on the spatially aggregated surface radiometric temperature. Furthermore, at any given level of fractional cover, a grid cell dominated more by needleleaf than by broadleaf vegetation, or one dominated more by a vertically complex canopy structure than a vertically simple structure, should have a lower value for r_h, assuming a linear mixing model and the constancy of the characteristic value for r_h over the non-vegetated fraction over all grid cells. In other words, assuming that the "soil" component of the heterogeneous sensor view is similar between two vegetation components with different characteristic leaf

dimensions or vertical structure, then the relative values for r_h between these two vegetation types should reflect the expected differences due to canopy architecture. If, on the other hand, the expected r_h for the "soil" fractions are greatly different between vegetation types, for example if in one type the non-vegetated view-fraction is dominated by standing dead litter while in another type it is dominated by bare soil, then this relative difference in r_h between the vegetation types at the same fractional cover (as predicted by Eq. 18) would not necessarily be representative of the difference expected due to vegetation structure.

I tested the dependence of predicted r_h on canopy cover in a three step process. First, I used the independent dataset of vegetation type and canopy closure to stratify the r_h predictions over a number of spatial subsets of the study region. In this step, I have only qualitative fractional cover information (the canopy closure classes) to work with. There is not necessarily a correspondence of the canopy closure classes between vegetation types in the same spatial subset, nor between spatial subsets considering the same vegetation type. In this step then, I was interested only in establishing, from independent data, that a qualitative relationship between fractional cover and r_h existed. I compared the ranking of average r_h, as stratified in each spatial subset and in each vegetation type for which canopy closure was included as a variable, against the ranking of canopy closure codes.

In the second step of this analysis I accepted the hypothesis that fractional cover was related to r_h and sought to place this relationship on a more quantitative basis by comparing average values for NDVI with average values for r_h, using the same stratification of the spatially subsetted data as in the first step (i.e., stratification by independent estimates of landcover type and canopy closure). In this step I assumed that NDVI was closely related to fractional cover, and that this relationship was consistent between vegetation types (Asrar et al. 1992). The new hypothesis being explored in this step was that at least two different responses would be observed between r_h and NDVI, one for needleleaf forest vegetation and another for the combination of shrubby and broadleaf vegetation.

In the third and final step of this analysis, having accepted the hypothesis that at least two distinct relationships could be observed between r_h and NDVI in the spatial subset of my study region, I employed another landcover database, created by Loveland et al. (1991) from AVHRR data, which was available

over the entire study region. I reclassed this data to a few very coarse vegetation types (Table 9), and evaluated the relationship between r_h and NDVI separately for each of these classes. This evaluation was performed by taking the average predicted r_h for equally sized subsets of the total population of cells in a given landcover class, ranked by NDVI. These average values of r_h were plotted against average NDVI to obtain curves for each vegetation type over the range of NDVI. A map of the recoded landcover types is shown in Figure 21.

Table 9. Landcover types reclassed from the Loveland classification

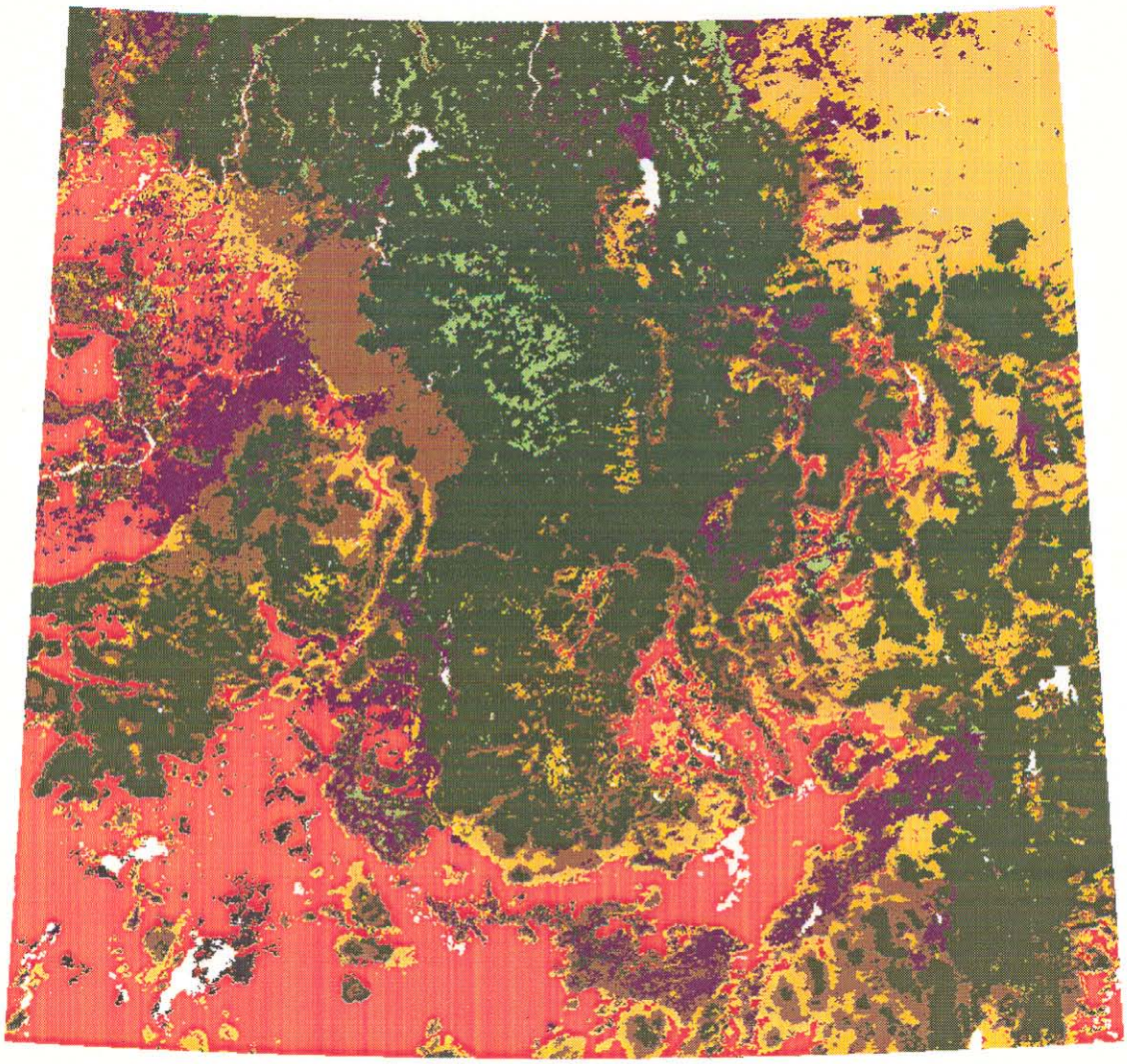
Class	# 1 km^2 cells	Original Loveland values
Agriculture (wet and dry)	50391	1-34
Mixed ag/grassland/woodland	81917	35-54
Grassland	109088	55-65, 157-159
Grassland/shrubland	151072	66-83
Deciduous forest	8565	90-97
Evergreen forest	286161	86-89, 99-132, 136, 137, 141-148, 156
Barren	4560	155

6. Results

6.1 Multiple regression analysis

Results from the multiple regression analysis are summarized by landcover type (Loveland recode) in Table 10. The multiple regressions explain on average 50% of the variance in observed temporal variation of ΔT, with somewhat more variance explained in sparsely vegetated types and somewhat lesss in forested types. Mean absolute prediction errors for instantaneous ΔT predictions are 2.0-2.5 °C for non-forest covertypes, and about 1.5°C for forested covertypes.

The influence of R_n is seen to be positive for all covertypes, and to increase for more arid covertypes, as anticipated by Eq. 15. The orbital drift mentioned earlier, and illustrated in Figure 19, was actually fortuitous in this study, since it allowed a wider range of R_n to enter into the regressions than would have been the case if the orbit were stable. As the orbit drifted to later overpass times, the sensor



Agriculture
Mixed agric./grassland/woodland
Grassland
Grassland/shrubland
Deciduous forest
Evergreen forest
Barren

Landcover classes, recoded from Loveland

Figure 21

recorded the response of the surface to progressively reduced R_n. The other primary source of variation in R_n is due to the seasonal cycle. The influence of NDVI is always negative, as expected from the arguments presented above concerning fractional vegetation cover and the difference in r_h between vegetation and soil. The influence of VPD is mixed. A negative influence is expected from Eq. 15, and is observed for forested covertypes, but a positive relationship is observed for all other covertypes. It may be that the influence of air temperature on some parameter such as **s** (a positive, non-linear effect), which is not included in the regression, is overwhelming the expected behavior of ΔT with respect to VPD, since VPD is very closely and positively related to air temperature. The influence of antecedent precipitation on ΔT is negative in all cases, as expected from the arguments concerning the influence of soil water content on r_v for soil and for vegetation, and the influence of varying r_v on ΔT, illustrated in Figure 1. There appears to be a strong influence of covertype on this relationship, with a progressive increase in the magnitude of the response with increasingly arid vegetation types. This is the variation in sensitivity that would be expected if the typical values for r_h were higher in the more arid vegetation types, since the potential reduction in ΔT for a given reduction in r_v is greater for the case of large r_h than for small r_h (Eq. 15, Figure 1). Finally, the influence of θ on ΔT is negative in every case, as expected from the arguments above concerning the fraction of sunlit vs. shaded canopy and soil visible to the sensor. The effect is rather large, with an increase of 3 °C for agricultural landcover or 4-4.5 °C for grasslands and forests as the sun-sensor angle shifts from 90° to 0°. Values of θ in the range of 90° are common in the composited AVHRR data, averaging around 60°, for a typical reduction from full-sun view surface temperatures of 2.5° C.

Table 10. Multiple regression results for ΔT vs. (R_n, NDVI, VPD, P_{pre}, θ), averaged by landcover type

Units in parentheses are for the regression coefficients. R^2 are the simple correlation coefficients

	R_n (°C/ Wm^{-2}) Coefficient		R^2		NDVI (°C) Coefficient		R^2	
Covertype	μ	σ	μ	σ	μ	σ	μ	σ
Evergreen for.	0.005	0.007	0.07	0.09	-9.4	10.0	0.10	0.11
Deciduous for.	0.004	0.006	0.04	0.05	-12.5	7.6	0.25	0.15
Agriculture	0.015	0.007	0.10	0.10	-22.0	8.1	0.22	0.16
Mixed ag/other	0.013	0.007	0.11	0.11	-18.7	9.8	0.14	0.14
Grassland	0.017	0.006	0.18	0.12	-18.8	9.2	0.08	0.08

Grass/shrub	0.021	0.006	0.27	0.13	-18.6	14.7	0.06	0.07
Barren	0.020	0.005	0.27	0.14	-20.1	24.3	0.27	0.14

	VPD (°C/Pa) Coefficient		R^2		P_{pre} (°C/cm) Coefficient		R^2	
Covertype	μ	σ	μ	σ	μ	σ	μ	σ
Evergreen for.	-4e-5	0.0011	0.06	0.07	-0.23	0.46	0.03	0.05
Deciduous for.	-5e-4	0.0001	0.13	0.12	-0.05	0.37	0.04	0.05
Agriculture	0.0004	0.0012	0.14	0.12	-0.48	0.78	0.07	0.07
Mixed ag/other	0.0006	0.0011	0.06	0.07	-0.52	0.60	0.05	0.06
Grassland	0.0010	0.0010	0.10	0.08	-0.73	0.60	0.05	0.06
Grass/shrub	0.0013	0.0010	0.17	0.11	-1.20	0.90	0.07	0.07
Barren	0.0006	0.0009	0.10	0.09	-1.47	0.70	0.08	0.07

	θ (°C/°) Coefficient		R^2		Intercept (°C)	
Covertype	μ	σ	μ	σ	μ	σ
Evergreen for.	-0.045	0.026	0.16	0.13	12.6	3.8
Deciduous for.	-0.035	0.022	0.12	0.11	13.0	3.0
Agriculture	-0.033	0.030	0.05	0.07	15.1	3.5
Mixed ag/other	-0.047	0.032	0.11	0.12	15.5	3.8
Grassland	-0.050	0.032	0.08	0.10	14.0	3.7
Grass/shrub	-0.050	0.032	0.06	0.07	15.0	3.9
Barren	-0.040	0.038	0.04	0.05	17.1	4.3

	Multiple R^2		Mean absolute error (°C)	
Covertype	μ	σ	μ	σ
Evergreen for.	0.42	0.15	1.71	0.43
Deciduous for.	0.47	0.15	1.58	0.40
Agriculture	0.54	0.14	2.44	0.54
Mixed ag/other	0.51	0.13	2.28	0.47
Grassland	0.55	0.11	2.36	0.42
Grass/shrub	0.59	0.12	2.54	0.47
Barren	0.57	0.13	2.25	0.41

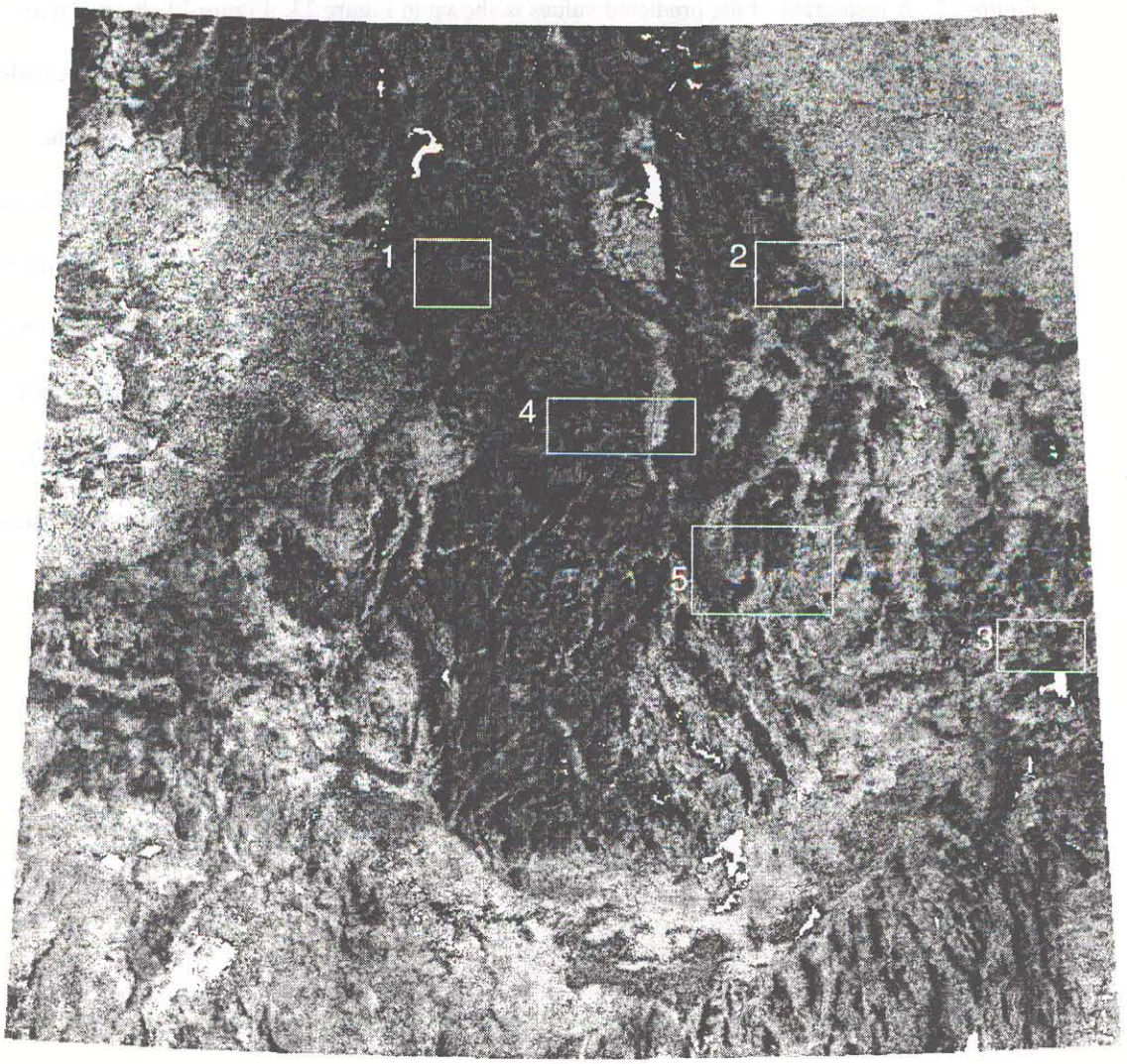
6.2 Predicted r_h: variation with canopy closure and NDVI

The means and standard deviations for predicted values of r_h and observed NDVI are presented in Table 11, stratified by covertype. A map of the predicted values of r_h over the study region is shown in Figure 22. A histogram of the predicted values is shown in Figure 23. Figure 24 shows a map of average NDVI, taken over all growing season composite periods from 1991-1994, and excluding periods for a given cell that were labeled as cloudy or as having excessively off-nadir viewing angles. The accompanying NDVI histogram is shown in Figure 25. Predicted negative values for r_h are frequent in the forested regions with the highest NDVI. Negative resistances have no physical meaning, but I retain them through the rest of this analysis in order to examine the relationship between predicted r_h and NDVI over the entire range of the predicted values. Since r_h is a function of the regression coefficient for R_n, the same patterns are observed in Table 11 as in Table 10. The more densely vegetated areas have low r_h and the more sparsely vegetated area have higher r_h, suggesting that the characteristic difference in values for this parameter between soils and vegetation is realized in these predictions.

Table 11. Predicted r_h and observed NDVI, average by covertype

	r_h		NDVI	
Covertype	μ	σ	μ	σ
Evergreen for.	6.0	8.0	0.391	0.058
Deciduous for.	4.4	7.3	0.456	0.040
Agriculture	17.9	8.1	0.337	0.067
Mixed ag/other	15.4	8.0	0.327	0.047
Grassland	20.5	7.0	0.262	0.051
Grass/shrub	25.2	7.3	0.186	0.037
Barren	24.7	6.5	0.134	0.016

Results from the stratification of predicted r_h with respect to canopy closure estimates from TM data are presented in Table 12, the format of which requires some explanation. Individual entries, delimited by horizontal bars, show means and standard deviations of predicted r_h and NDVI. Each entry is stratified by two or three canopy closure classes within a single landcover type from one of the five spatial subsets shown in Figure 22. Only combinations of covertype and canopy closure represented by 30 or more 1 km^2 grid cells in a given spatial subset are included in Table 12. Only covertypes with two canopy



Predicted r_h (s m^{-1}), also showing landcover subsets

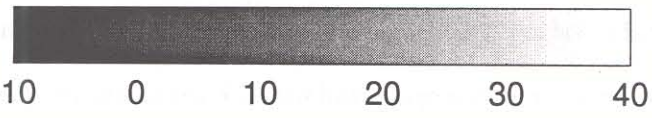


Figure 22

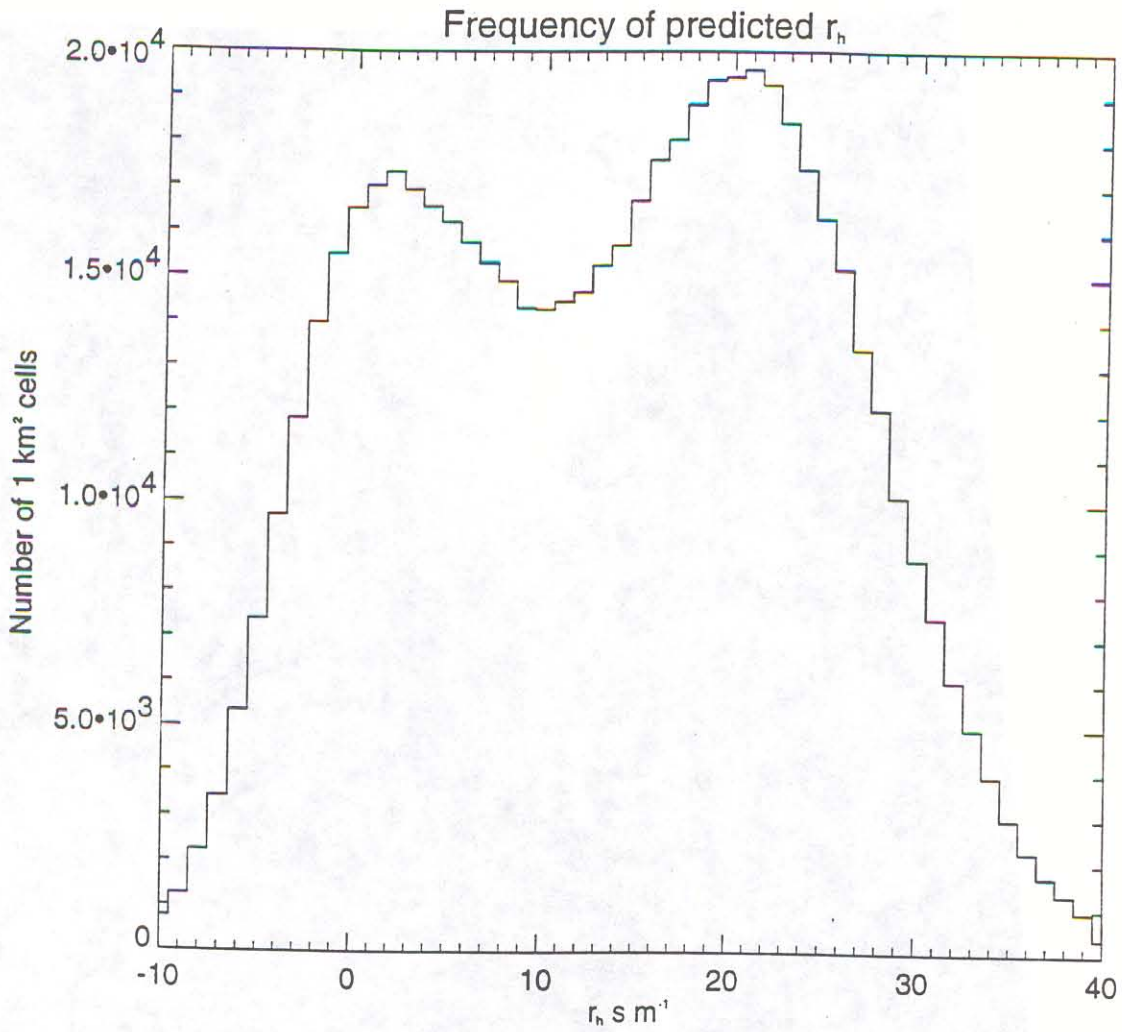
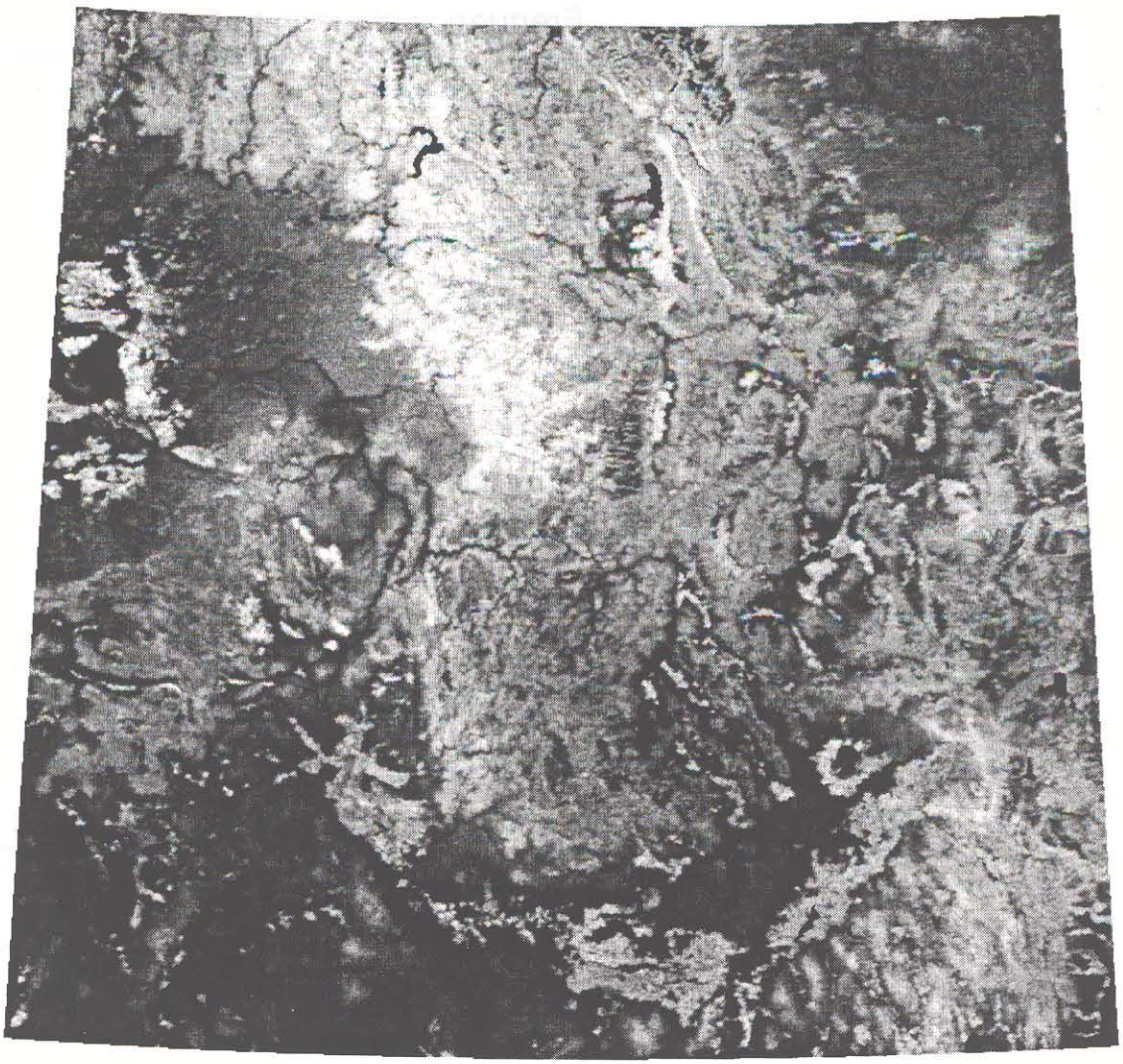

Frequency of predicted r_h
Number of 1 km² cells
2.0•10^4
1.5•10^4
1.0•10^4
5.0•10^3
0
-10
0
10
20
30
40
r_h s m^{-1}

Figure 23



Average growing-season NDVI, 1991-1994



Figure 24

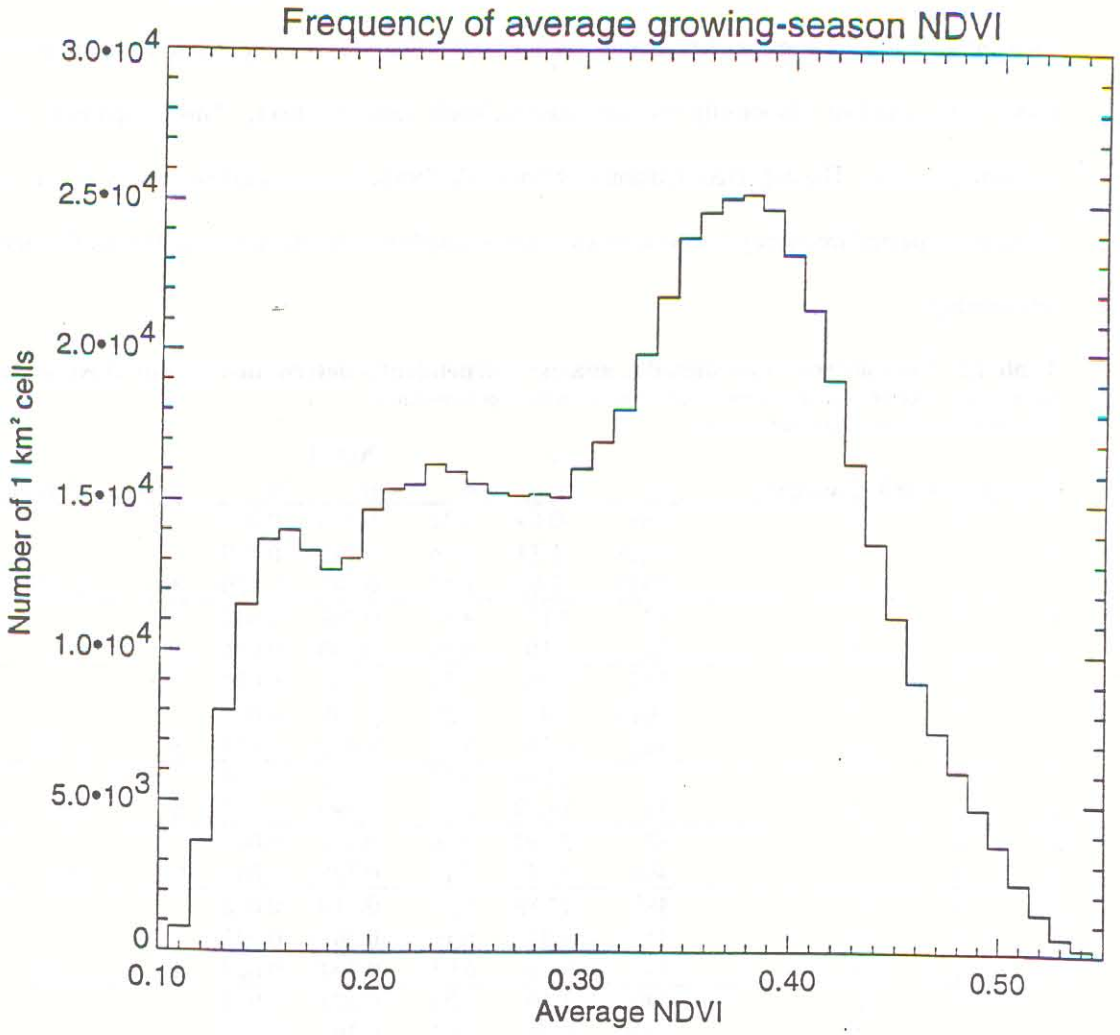

Frequency of average growing-season NDVI
Number of 1 km² cells
3.0•10^4
2.5•10^4
2.0•10^4
1.5•10^4
1.0•10^4
5.0•10^3
0
0.10
0.20
0.30
0.40
0.50
Average NDVI

Figure 25

are being assessed here, because the canopy classes are categorical, because they are not consistent between cover types, and because they are not necessarily consistent between spatial subsets. The maximum number of closure classes is three, and with three classes the number of possible independent comparisons is also three (1 vs. 2, 2 vs. 3, and 1 vs. 3). For cases with only two classes, a single comparison can be made (1 vs. 2). The number of rankings favorable to my hypothesis is recorded for each entry. The total number of favorable rankings over all entries is compared to the number expected under from a random relationship between canopy closure and predicted r_h. This comparison is made using a one-sided z-test. The expected frequency of favorable comparisons in each three-way case is 1.333, while the expected frequency for the two-way case is simply 0.5, for the null hypothesis of a random relationship.

Table 12. Comparison of predicted r_h against independently determined canopy closure classes.
Cover codes: 1=conifer forest, 2=xeric shrub, 3=mesic shrub, 4=grass/meadow
Closure codes: 1=low, 2=medium, 3=high

				r_h		NDVI			
Subset#	Cover#	Closure#	n	μ	σ	μ	σ	#favor	#expected
1	1	1	365	-0.09	3.31	0.475	0.031		
1	1	2	1428	-1.34	3.56	0.485	0.029		
1	1	3	1047	-3.02	3.53	0.481	0.030	3	1.333
1	3	2	32	2.17	4.43	0.464	0.043		
1	3	3	32	1.10	3.58	0.500	0.016	1	0.5
2	1	1	148	8.26	7.72	0.369	0.034		
2	1	2	1015	3.45	5.29	0.388	0.031		
2	1	3	442	0.53	3.71	0.402	0.031	3	1.333
2	3	1	51	12.35	8.32	0.383	0.038		
2	3	2	43	16.39	7.64	0.399	0.029	0	0.5
2	4	1	85	22.62	4.66	0.324	0.048		
2	4	2	496	22.77	5.12	0.348	0.037	0	0.5
3	1	1	482	12.82	5.25	0.310	0.032		
3	1	2	817	9.02	5.16	0.345	0.027		
3	1	3	53	8.79	6.91	0.362	0.027	3	1.333
4	1	1	640	4.10	5.29	0.374	0.042		
4	1	2	2220	0.96	4.03	0.398	0.045		
4	1	3	558	-1.27	3.82	0.452	0.045	3	1.333
4	4	1	143	13.60	7.89	0.354	0.052		
4	4	2	63	12.53	6.95	0.401	0.033		
4	4	3	38	11.29	8.02	0.415	0.041	3	1.333
5	1	1	171	8.52	6.50	0.329	0.036		
5	1	2	1911	8.26	4.85	0.369	0.027		
5	1	3	386	3.58	3.65	0.383	0.022	3	1.333
5	2	1	1686	18.35	6.63	0.268	0.068		
5	2	2	124	14.10	5.80	0.364	0.042	1	0.5

5	4	1	1062	21.96	5.30	0.261	0.081		
5	4	2	351	17.46	6.10	0.351	0.074		
5	4	3	135	12.73	7.60	0.354	0.045	3	1.333
						Totals		**23**	**11.333**
						z = (23-11.333)/2.5 =			**4.67**
									P<0.001

This evidence strongly contradicts the null hypothesis of a random relationship between r_h and canopy closure.

Values for NDVI have been included in Table 12 for the purpose of illustrating that within the canopy closure classes for a given entry there is an excellent agreement with rank of NDVI. This evidence leads to the new hypothesis that NDVI could be a good predictor of r_h, at least for certain covertypes or groups of covertypes. A plot of the average values of r_h and NDVI stratified by landcover (Loveland) over the entire study area (data from Table 11, plot in Figure 26) suggests the same hypothesis, although this plot contains no information on the form of the relationship within a particular landcover class. As a first look at the relationship between predicted r_h and observed NDVI within landcover types, I used the TM-based landcover and canopy closure data to stratify the r_h and NDVI data for a regression analysis. Two regressions were performed, one for conifer forest and another for grass and shrub cover types. Individual points in the regressions were based on the spatial average of r_h and NDVI within a given spatial subset, for a given covertype, and for a particular canopy closure class (Figure 27). Only partitions covering more than 30 gridcells are included in the regressions and in Figure 27. For reference, points from the agriculture covertypes are shown in Figure 27, although they were not included in the regressions. The resulting regression equations are:

Eq. 34 $r_h = 31.6 - 71.37\ NDVI \quad R^2 = 0.63$ (conifer forest)

$r_h = 49.2 - 90.73\ NDVI \quad R^2 = 0.69$ (other types)

For these regressions I aggregated a "standing burnt forest" class and a "barren" class into the conifer forest group. Standing burnt forest was assumed to be a case of very low fractional cover conifer forest, and subjective evaluation of the barren class indicated that it was mostly alpine rock environments along narrow ridges. These rocky patches are usually not fully resolved at the AVHRR sampling

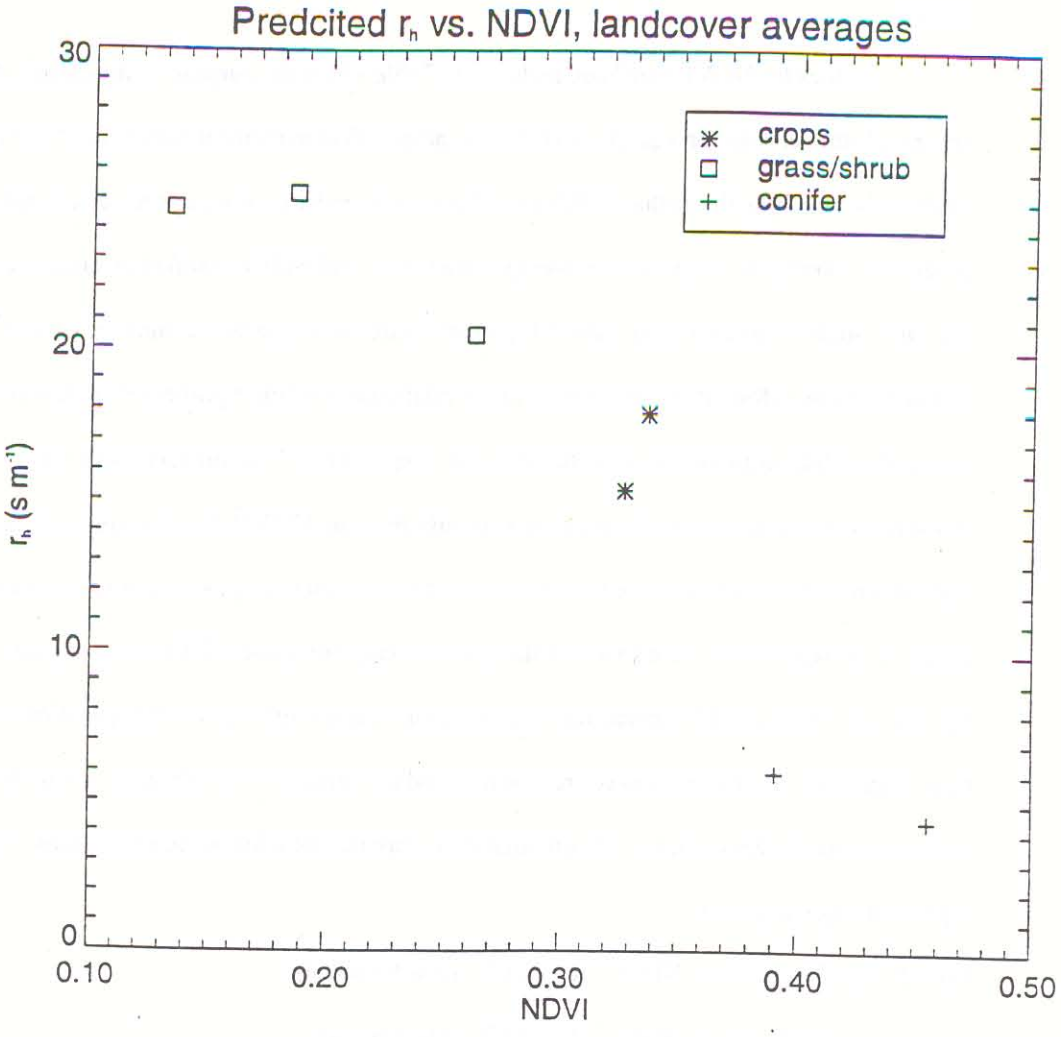

Predcited r_h vs. NDVI, landcover averages
30
20
10
0
r_h (s m^{-1})
0.10
0.20
0.30
0.40
0.50
NDVI
crops
grass/shrub
conifer

Figure 26

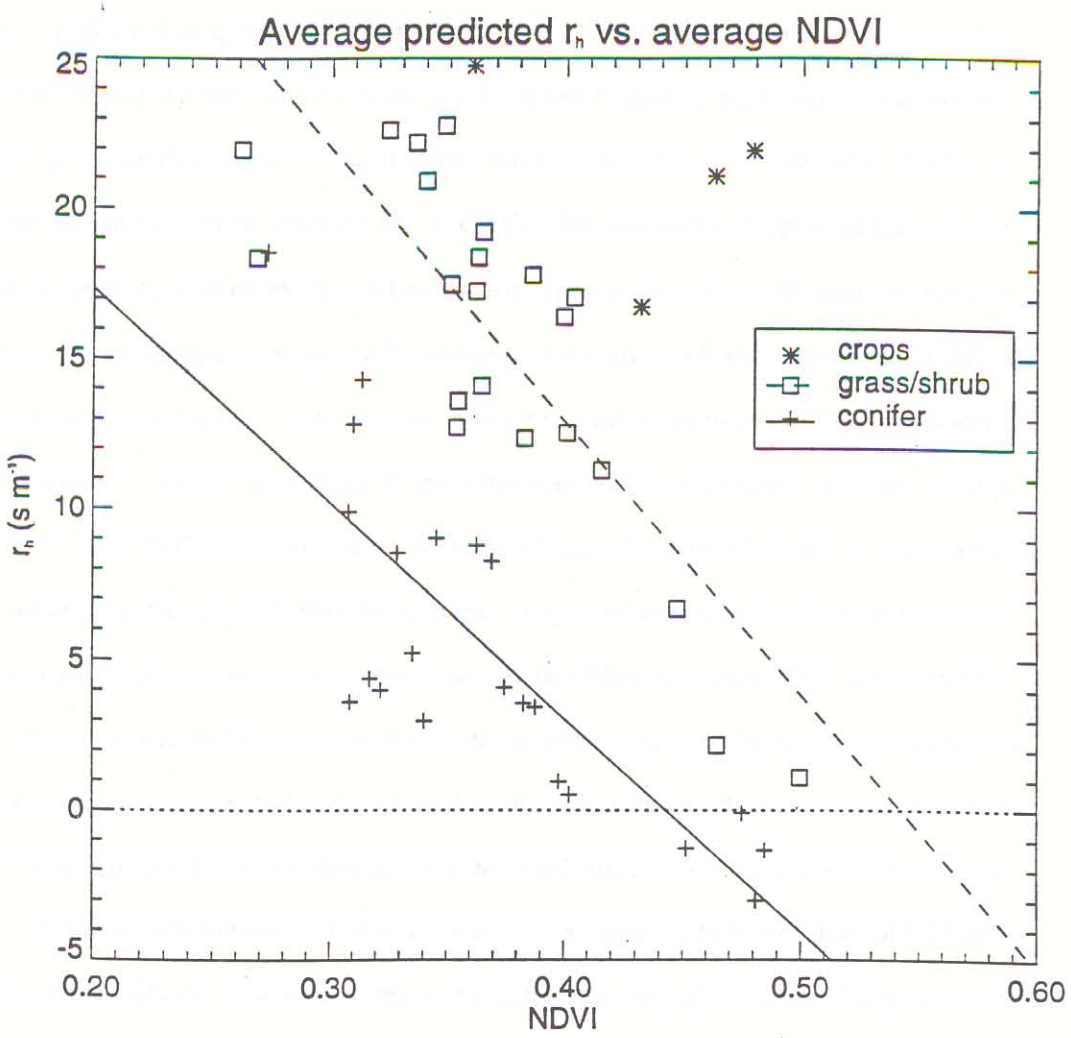

Average predicted r_h vs. average NDVI
25
20
15
10
5
0
-5
r_h (s m^{-1})
0.20
0.30
0.40
0.50
0.60
NDVI
crops
grass/shrub
conifer

Figure 27

resolution, and so I assume that they will be radiometrically mixed with the surrounding vegetation, which in this case is conifer forest.

The results plotted in Figure 27 suggest that there is in fact a unique relationship between r_h and NDVI for each of these two broad categories of landcover. The range of NDVI from the independent landcover subsets is not adequate to resolve the form of these relationships at low NDVI. If the theoretical framework that I have been building in this study is accurate, I would expect to find that the relationships for different vegetation types converge at the low end of the NDVI range, with increasing variation between vegetation types at moderate and high NDVI. As fractional cover decreases, the vegetation type becomes less important to the surface energy balance and the influence of the unvegetated surface increases, until at very low fractional cover, regardless of the vegetation type, the energy partitioning characteristics of the bare surface dominate the response of ΔT to R_n, and so will dominate the predicted r_h, using my method. Conversely, at high fractional cover the stronger atmospheric coupling (lower r_h) characteristic of leaves dominates the aggregate response in the sensor's field of view. For vegetation types, such as forests, that show an increase in vertical complexity with increasing fractional cover (Munro 1989, Hall et al. 1995) there is an additional downward influence on canopy-scale r_h due to increased turbulence that creates higher average wind speeds within the canopy (Domingo et al. 1995). Thus at high fractional cover in vertically complex types, the values of r_h should decrease, approaching zero. Some aspects of this assumption have recently been tested and supported by comparing the surface temperature from AVHRR with observed air temperature over dense grass and riparian forest canopies in Kansas (Prihodko and Goward 1997). They found that surface temperature was very close to, but slightly higher than, near-surface air temperature over a range of other conditions. This indicates a value for r_h close to zero. An exception to this pattern is expected for canopies that do not develop significant vertical complexity as fractional cover increases, such as in grasses, shrubs, and crops. For such canopies, the values of r_h characteristic of individual leaves should be maintained even at high fractional cover, although it is still expected that r_h would approach the value for the unvegetated ground surface with low fractional cover.

I explored this more explicit set of arguments using the graph-of-averages approach (Freedman et al. 1991). This method provides information similar to a regression analysis, but is better suited to the identification of nonlinear variations. Using the Loveland landcover classification, I divided the study region into five covertypes: forest (including both evergreen and deciduous forest), grassland, shrub/grassland, agriculture, and barren. Within each region I sorted the NDVI values, and then calculated the average predicted r_h and the average NDVI for each subset of 500 grid cells, progressing through the list of sorted NDVI values. By plotting these averages against each other, the shape of the average relationship between r_h and NDVI is revealed, including any potential nonlinearities. These plots are shown in Figure 28, and they indicate that the arguments presented above provide a plausible explanation for the general patterns in the data. Forest covertypes have a nearly linear decline in predicted r_h with increasing NDVI, reaching the zero-line for r_h at an NDVI of about 0.5, and remaining at this level through the rest of their NDVI range, up to about 0.55. The relationships for grasslands, shrub/grass mixtures, and agricultural areas are all distinctly different from the relationship for forests. These differences are consistent with the hypothesis that, for grass and shrub types, r_h should be characteristic of the leaf-scale value at high fractional cover, without showing the decline to lower values characteristic of rough canopies. Assuming a leaf width of 1 cm for a grass or crop, the value for r_h predicted from engineering principles is 16 s m^{-1}, for an average windspeed of 1 m s^{-1} (Nobel 1991), in good agreement with Figure 28. There is considerable variation between the curves for these non-forest covertypes, and I am not able to venture an explanation for it. There may be considerable confusion between the classification by Loveland and the reality of vegetation cover over the study area, especially with respect to the discrimination between grass, shrub, and dry agriculture. An explanation of differences between these types would rely on a more accurate assessment of landcover, perhaps as undertaken in the TM-based classification.

An additional observation from Figure 28 is that all the curves appear to converge at low NDVI, which is the expected result if the bare-ground has a relatively constant characteristic r_h for the different covertypes. Visual inspection of the curves for vegetated cover types suggests a value of around 30 s m^{-1} for the bare-surface r_h, and the small amount of data for the barren covertype supports this conclusion (the

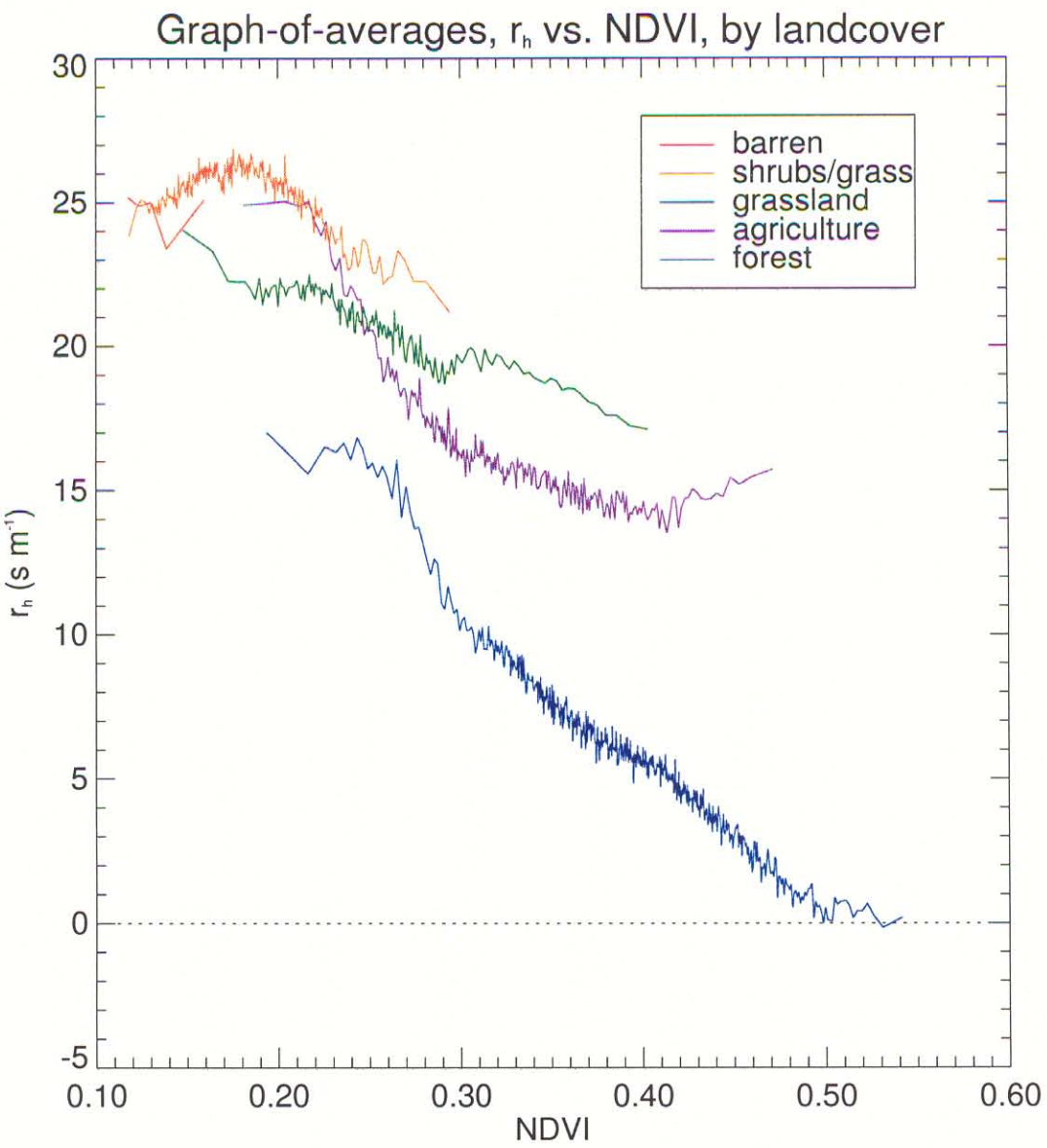

Graph-of-averages, r_h vs. NDVI, by landcover
barren
shrubs/grass
grassland
agriculture
forest
30
25
20
15
10
5
0
-5
r_h (s m^{-1})
0.10
0.20
0.30
0.40
0.50
0.60
NDVI

Figure 28

red line in the upper-left of Figure 28). This is a lower value than suggested by measurements made over very small areas of flat soil. For example, van de Griend and Owe (1994) find a value of about 50 s m^{-1} using similarity theory and observations over a bare soil with an area of 1 m^2. The difference may be due to the additional roughness at the 1 km^2 grid-cell scale that is not included in small-scale measurements.

7. Summary of results and discussion of applications

I found that the observations of radiometric surface temperature are broadly consistent with the surface energy balance theory summarized by Eq. 15. In particular, I found that the variation of ΔT with respect to meteorological parameters expected from the theory embodied in Eq. 15, with the exception of the sensitivity to VPD, is observed over a broad range of covertypes in a region with diverse climate and topography. An examination of the variation in ΔT with respect to variation in R_n was suggested by the form of Eq. 15 and implemented in terms of a predicted value for r_h. The variation over landcover type and fractional vegetation cover in values of r_h derived from observations of surface radiometric properties and estimates of near-surface meteorological conditions, according to Eq. 18, was consistent with: (a) the general principles known to govern r_h as observed for single canopy or soil elements; (b) considerations from a simple, qualitative theory of mixing for data from sensors with a heterogeneous field-of-view; and (c) observations from other studies.

None of this evidence alone is conclusive, but taken together I find that it supports the hypothesis that the radiometric observations and the surface meteorological data are in agreement with at least one part of the theory presented in Eq. 15, that is, with the theory embodied in the first term, showing the dependence of ΔT on R_n. The explicit purpose of this investigation is accomplished, then, but an important question remains: How are these results pertinent to the application of the model described in Chapter 3? I believe they are pertinent in the following two ways:

(1) They indicate that the algorithms presented in Chapter 3 for the treatment of the surface energy fluxes, although having been derived from theory developed mostly at very small spatial scales, may be applicable to predictions at large spatial scales. In other words, the results presented here support the

transfer of the "one dimensional BGC" algorithms to a "two dimensional" case. In particular, they support the idea that the basic ecophysiological theory embodied in Eq. 15 is consistent with observations at a much larger resolution (or level of aggregation) than the conditions under which it has been most studied, i.e., at the plant or plot scales. This conclusion is of practical importance for the development of global ecological monitoring strategies. Such strategies will necessarily rely on observations from satellite remote sensing platforms for information on spatial and temporal variability at the land surface, but in order to generate useful information on ecological processes from the radiometric data they will have to rely on an integrative numerical framework such as the one proposed in Chapter 3 (Running et al. 1994). My results here show a promising degree of correspondence between the radiometric observations from a satellite sensor with daily global sampling frequency and one piece of an integrative numerical ecological process simulation model.

(2) The 1D-BGC logic requires an estimate of r_h. By using radiometric data collected over several years, in conjunction with a database of surface meteorological conditions, an estimate of r_h can be derived for a range of landcover types and fractional vegetation coverage. The determination of this parameter has been problematic in the past, because of its scale-dependence, and because of its dependence on atmospheric conditions such as stability and turbulence, which are difficult to derive from the commonly available meteorological data.

An additional potential application of the work presented here is in the qualitative estimation from remotely sensed radiometric data of regional-scale drought. This has been a major objective of the recent thermal remote sensing literature, as reviewed earlier, and many methods have been proposed. These methods have generally relied on the radiometric surface temperature to provide a relative index of drought, but they have usually ignored the influence of incident radiation and near surface air temperature on the surface temperature. The value of r_h as predicted by the methods proposed here incorporates both of these influences, and represents the long-term average thermal response of the surface to variation in incident radiation. This long-term average response would be useful as a baseline for the detection of

unusually high (or low) values of ΔT under a particular incident radiation load, with unusually high values indicative of drought.

One liability of this method, in comparison to the method proposed by Nemani et al. (1993) and reviewed above, is that it requires an estimate of incident radiation and near-surface air temperature. The radiation estimate is not very problematic, since it can be derived from the time of data acquisition and the sun-slope geometry. Surface air temperature presents more of a problem, and the methods employed in the this study can not yet be applied operationally, since the majority of the surface air temperature observations are not available until at least months after they would be needed. However, an adequate approximation for the purpose of a qualitative drought index might be obtained from only the primary station reports, which are available almost immediately after they are taken. Additional information could be obtained, if necessary, from the thermal remote sensing data itself, using the observation, reported by Prihodko and Goward (1997), and also observed in this study, that canopy and air temperatures are nearly equal over dense forest vegetation.

An advantage of this method, in addition to the normalization for the influence of air temperature and radiation on the surface temperature, is that once the baseline values are determined, an estimate of drought severity can be made for each individual grid cell, provided it passes the cloud filtering procedure. The method proposed by Nemani et al. (1993) uses data from many grid cells in a region around the point of interest (they suggest a 400 km^2 region), and so results in a drought severity map with less spatial resolution. An improvement in the spatial resolution of such a map may be of little importance in the identification of regional trends in drought severity, but it could be of more importance to a land manager trying to assess the finer spatial details of drought severity.

8. Literature cited

Asrar, G., Fuchs, M., Kanemasu, E.T. and Hatfield, J.H., 1984. Estimating absorbed photosynthetic radiation and leaf area index from spectral reflectance in wheat. Agronomy Journal, 76: 300-306.

Asrar, G., Myneni, R.B. and Choudhury, B.J., 1992. Spatial heterogeneity in vegetation canopies and remote sensing of absorbed photosynthetically active radiation: a modeling study. Remote Sensing of Environment, 41: 85-103.

Bartholic, J.F., Namken, L.N. and Wiegand, C.L., 1972. Aerial thermal scanner to determine temperatures of soils and of crop canopies differing in water stress. Agronomy Journal, 64: 603-608.

Beaumont, P., 1989. Drylands: Environmental Management and Development. Routledge, London.

Becker, F., 1987. The impact of spectral emissivity on the measure of land surface temperature from a satellite. International Journal of Remote Sensing, 8(10): 1509-1522.

Bristow, K.L. and Campbell, G.S., 1984. On the relationship between incoming solar radiation and daily maximum and minimum temperature. Agricultural and Forest Meteorology, 31: 159-166.

Brutsaert, W. and Sugita, M., 1992. Regional surface fluxes under nonuniform soil moisture conditions during drying. Water Resources Research, 28(6): 1669-1674.

Byrne, G.F., Begg, J.E., Fleming, P.M. and Dunin, F.X., 1979. Remotely sensed land cover temperature and soil water status - a brief review. Remote Sensing of Environment, 8: 291-305.

Carlson, T.N., Belles, J.E. and Gillies, R.R., 1991. Transient water stress in a vegetation canopy: simulations and measurements. Remote Sensing of Environment, 35: 175-186.

Carlson, T.N. and Buffum, M.J., 1989. On estimating total daily evapotranspiration from remote surface temperature measurements. Remote Sensing of Environment, 29: 197-207.

Carlson, T.N., Capehart, W.J. and Gillies, R.G., 1995a. A new look at the simplified method for remote sensing of daily evapotranspiration. Remote Sensing of Environment, 54: 161-167.

Carlson, T.N., Perry, E.M. and Schmugge, T.J., 1990. Remote estimation of soil moisture availability and fractional vegetation cover for agricultural fields. Agricultural and Forest Meteorology, 52: 45-69.

Carlson, T.N. et al., 1995b. An overview of the workshop on thermal remote sensing held at La Londe les Maures, France, September 20-24, 1993. Agricultural and Forest Meteorology, 77: 141-151.

Caselles, V., Coll, C. and Valor, E., 1997. Land surface emissivity and temperature determination in the whole HAPEX-Sahel area from AVHRR data. International Journal of Remote Sensing, 18(5): 1009-1027.

Caselles, V., Sobrino, J.A. and Coll, C., 1992. A physical model for interpreting the land surface temperature obtained by remote sensors over incomplete canopies. Remote Sensing of Environment, 39: 203-211.

Chanzy, A. and Bruckler, L., 1993. Significance of soil surface moisture with respect to daily bare soil evaporation. Water Resources Research, 29(4): 1113-1125.

Choudhury, B.J., 1987. Relationships between vegetation indices, radiation absorption, and net photosynthesis evaluated by a sensitivity analysis. Remote Sensing of Environment, 22: 209-233.

Choudhury, B.J., 1991. Multispectral satellite data in the context of land surface heat balance. Reviews of Geophysics, 29(2): 217-236.

Choudhury, B.J., Idso, S.B. and Reginato, R.J., 1986. Analysis of a resistance-energy balance method for estimating daily evaporation from wheat plots using one-time-of-day infrared temperature observations. Remote Sensing of Environment, 19: 253-268.

Cihlar, J., 1996. Identification of contaminated pixels in AVHRR composite images for studies of land biosphere. Remote Sensing of Environment, 56: 149-163.

Clum, H.H., 1926a. The effect of transpiration and environmental factors on leaf temperatures I. Transpiration. American Journal of Botany, 13: 194-216.

Clum, H.H., 1926b. The effects of transpiration and environmental factors on leaf temperatures II. Light intensity and the relation of transpiration to the thermal death point. American Journal of Botany, 13(4): 217-230.

Coll, C., Caselles, V. and Schmugge, T.J., 1994a. Estimation of land surface emissivity differences in the split-window channels of AVHRR. Remote Sensing of Environment, 48: 127-134.

Coll, C., Caselles, V., Sobrino, J.A. and Valor, E., 1994b. On the temperature dependence of the split-window equation for land surface temperature. International Journal of Remote Sensing, 15(1): 105-122.

Crago, R.D., 1996. Mixed layer convective turbulence theory with First International Satellite-Land Surface Climatology Project Field Experiment data. Water Resources Research, 32(9): 2767-2774.

Curtis, O.F., 1936. Leaf temperature and the cooling of leaves by radiation. Plant Physiology, 11: 343-364.

Daamen, C.J. and Simmonds, L.P., 1996. Measurement of evaporation from bare soil and its estimation using surface resistance. Water Resources Research, 32(5): 1393-1402.

Daly, C., Neilson, R.P. and Phillips, D.L., 1994. A statistical-topographical model for mapping climatological precipitation over mountainous terrain. Journal of Applied Meteorology, 33: 140-158.

Diak, G.R. and Whipple, M.S., 1993. Improvements to models and methods for evaluating the land-surface energy balance and 'effective' roughness using radiosonde reports and satellite-measured 'skin' temperature. Agricultural and Forest Meteorology, 63: 189-218.

Domingo, F., van Gardingen, P.R. and Brenner, A.J., 1996. Leaf boundary layer conductance of two native species in southeast Spain. Agricultural and Forest Meteorology, 81: 179-199.

Eck, T.F. and Kalb, V.L., 1991. Cloud-screening for Africa using geographically and seasonally variable infrared threshold. International Journal of Remote Sensing, 12(6): 1205-1221.

Eidenshink, J.C., 1992. The 1990 conterminous U.S. AVHRR data set. Photogrammetric Engineering and Remote Sensing, 58(6): 809-813.

Eklundh, L.R., 1995. Noise estimation in NOAA AVHRR maximum-value composite NDVI images. International Journal of Remote Sensing, 16(15): 2955-2962.

Forsythe, W.E. and Christison, F.L., 1930. The absorption of radiation from different sources by water and by body tissue. Journal of the Optical Society of America, 20: 693-700.

Gates, D.M., 1964. Leaf temperature and transpiration. Agronomy Journal, 56: 273-277.

Gates, D.M., 1980. Biophysical Ecology. Springer-Verlag, New York, New York.

Gillespie, A.R., 1992. Spectral mixture analysis of multispectral thermal infrared images. Remote Sensing of Environment, 42: 137-145.

Glantz, M.H., 1994. Drought Follows the Plow. Cambridge University Press, Cambridge, 197 pp.

Goward, S.N., Cruickshanks, G.D. and Hope, A.S., 1985. Observed relation between thermal emission and reflected spectral radiance of a complex vegetated landscape. Remote Sensing of Environment, 18: 137-146.

Gutman, G.G., Ignatov, A.M. and Olson, S., 1994. Towards better quality of AVHRR composite images over land: reduction in cloud contamination. Remote Sensing of Environment, 50: 134-148.

Hall, F.G. et al., 1991. Satellite remote sensing of surface energy and mass balance: results from FIFE. Remote Sensing of Environment, 35: 187-199.

Hall, F.G., Shimabukuro, Y.E. and Huemmrich, K.F., 1995. Remote sensing of forest biophysical structure using mixture decomposition and geometric reflectance models. Ecological Applications, 5(4): 993-1013.

Heilman, J.L., Kanemasu, E.T., Rosenberg, N.J. and Blad, B.L., 1976. Thermal scanner measurement of canopy temperatures to estimate evapotranspiration. Remote Sensing of Environment, 5: 137-145.

Hoffer, R.M. and Johannsen, C.J., 1969. Ecological potentials in spectral signature analysis. In: P.J. Johnson (Editor), Remote Sensing in Ecology. University of Georgia Press, Athens.

Hope, A.S., 1988. Estimation of wheat canopy resistance using combined remotely sensed spectral reflectance and thermal observations. Remote Sensing of Environment, 24: 369-383.

Hope, A.S., Petzold, D.E., Goward, S.N. and Ragan, R.M., 1986. Simulated relationships between spectral reflectance, thermal emissions, and evapotranspiration of a soybean canopy. Water Resources Research, 22(6): 1011-1019.

Hsiao, T.C., 1973. Plant response to water stress. Annual Reviews of Plant Physiology, 24: 519-570.

Humes, K.S., Kustas, W.P. and Moran, M.S., 1994. Use of remote sensing and reference site measurements to estimate instantaneous surface energy balance components over a semiarid rangeland watershed. Water Resources Research, 30(5): 1363-1373.

Hungerford, R.D., Nemani, R.R., Running, S.W. and Coughlan, J.C., 1989. MTCLIM: A mountain microclimate simulation model. U.S. Forest Service Research Paper, INT-414.

Hunt, E.R., Jr., Rock, B.N. and Nobel, P.S., 1987. Measurement of leaf relative water content by infrared reflectance. Remote Sensing of Environment, 22: 429-435.

Idso, S.B., 1981. A set of equations for full spectrum and 8- to 14-μm and 10.5- to 12.5-μm thermal radiation from cloudless skies. Water Resources Research, 17(2): 295-304.

Idso, S.B. and Baker, D.G., 1967. Relative importance of reradiation, convection, and transpiration in heat transfer from plants. Plant Physiology, 42: 631-640.

Idso, S.B., Jackson, R.D. and Reginato, R.J., 1977. Remote-sensing of crop yields. Science, 196: 19-24.

Idso, S.B., Reginato, R.J., Reicosky, D.C. and Hatfield, J.L., 1981. Determining soil-induced plant water potential depressions in alfalfa by means of infrared thermometry. Agronomy Journal, 73: 826-830.

Idso, S.B., Schmugge, T.J., Jackson, R.D. and Reginato, R.J., 1975. The utility of surface temperature measurements for the remote sensing of surface soil water status. Journal of Geophysical Research, 80(21): 3044-3049.

Irbane, J.V. and Godson, W.L., 1981. Atmospheric Thermodynamics. D. Reidel Publishing Company, Dordrecht, The Netherlands.

Jackson, R.D., 1982. Canopy temperature and crop water stress. In: D. Hillel (Editor), Advances in Irrigation. Academic Press, Inc., New York, pp. 43-85.

Jackson, R.D., Idso, S.B., Reginato, R.J. and Pinter, P.J., Jr., 1981. Canopy temperature as a crop water stress indicator. Water Resources Research, 17(4): 1133-1138.

Jackson, R.D., Reginato, R.J. and Idso, S.B., 1977. Wheat canopy temperature: A practical tool for evaluating water requirements. Water Resources Research, 13(3): 651-656.

Jackson, R.D., Slater, P.N. and Pinter, P.J., Jr., 1983. Discrimination of growth and water stress in wheat by various vegetation indices through clear and turbid atmospheres. Remote Sensing of Environment, 13: 187-208.

Jarvis, P.G., 1976. The interpretation of the variations in leaf water potential and stomatal conductance found in canopies in the field. Philosophical Transactions of the Royal Society of London, B. 273: 593-610.

Jones, H.G., 1992. Plants and Microclimate. University Press, Cambridge, 428 pp.

Jordan, C.F., 1969. Derivation of leaf area index from quality of light on the forest floor. Ecology, 50: 663-666.

Kidwell, K.B., 1995. NOAA Polar Orbital Data User Guide (TIROS-N, NOAA-6, NOAA-7, NOAA-8, NOAA-9, NOAA-10, NOAA-12, NOAA-13, and NOAA-14), National Oceanic and Atmospheric Administration, Washington, D.C.

Kimes, D.S., 1980. Effects of vegetation canopy structure on remotely sensed canopy temperatures. Remote Sensing of Environment, 10: 165-174.

Kimes, D.S., Idso, S.B., Pinter, P.J., Jr., Reginato, R.J. and Jackson, R.D., 1980. View angle effects in the radiometric measurements of plant canopy temperatures. Remote Sensing of Environment, 10: 273-284.

Knipling, E.B., 1969. Leaf reflectance and image formation on color infrared film. In: P.N. Johnson (Editor), Remote Sensing in Ecology. University of Georgia Press, Athens.

Körner, C., 1995. Leaf diffusive conductances in the major vegetation types of the globe. In: E.-D. Schulze and M.M. Caldwell (Editors), Ecophysiology of Photosynthesis. Springer-Verlag, Berlin Heidelberg New York, pp. 463-490.

Kustas, W.P., Daughtry, C.S.T. and Van Oevelen, P.J., 1993. Analytical treatment of the relationships between soil heat flux / net radiation ratio and vegetation indices. Remote Sensing of Environment, 46: 319-330.

Kustas, W.P., Humes, K.S., Norman, J.M. and Moran, M.S., 1996. Single- and dual-source modeling of surface energy fluxes with radiometric surface temperature. Journal of Applied Meteorology, 35: 110-121.

Kustas, W.P. et al., 1994a. Surface energy balance estimates at local and regional scales using optical remote sensing from an aircraft platform and atmospheric data collected over semiarid rangelands. Water Resources Research, 30(5): 1241-1259.

Kustas, W.P. et al., 1990. Instantaneous and daily values of the surface energy balance over agricultural fields using remote sensing and a reference field in an arid environment. Remote Sensing of Environment, 32: 125-141.

Kustas, W.P., Perry, E.M., Doraiswamy, P.C. and Moran, M.S., 1994b. Using satellite remote sensing to extrapolate evapotranspiration estimates in time and space over a semiarid rangeland basin. Remote Sensing of Environment, 49: 275-286.

Loveland, T.R., Merchant, J.W., Ohlen, D.O. and Brown, J., 1991. Development of a land-cover database for the conterminous U.S. Photogrammetric Engineering and Remote Sensing, 57: 1453-1463.

Monteith, J.L. and Szeicz, G., 1962. Radiative temperature in the heat balance of natural surfaces. Quarterly Journal of the Royal Meteorological Society, 88: 496-507.

Moody, A. and Strahler, A.H., 1994. Characteristics of composited AVHRR data and problems in their classification. International Journal of Remote Sensing, 15(17): 3473-3491.

Moran, M.S., Clarke, T.R., Inoue, Y. and Vidal, A., 1994. Estimating crop water deficit using the relation between surface-air temperature and spectral vegetation index. Remote Sensing of Environment, 49: 246-263.

Moran, M.S., Jackson, R.D., Raymond, L.H., Gay, L.W. and Slater, P.N., 1989. Mapping surface energy balance components by combining Landsat Thematic Mapper and ground-based meteorological data. Remote Sensing of Environment, 30: 77-87.

Moran, M.S. et al., 1996. Combining the Penman-Monteith equation with measurements of surface temperature and reflectance to estimate evaporation rates of semiarid grassland. Agricultural and Forest Meteorology, 80: 87-109.

Munro, D.S., 1989. Stomatal conductance and surface conductance modelling in a mixed wetland forest. Agricultural and Forest Meteorology, 48: 235-249.

Myneni, R.B. and Ganapol, B.D., 1992. Remote sensing of vegetation canopy photosynthetic and stomatal conductance. Remote Sensing of Environment, 42: 217-238.

Nemani, R., Pierce, L., Running, S. and Goward, S., 1993. Developing satellite-derived estimates of surface moisture status. Journal of Applied Meteorology, 32(3): 548-557.

Nemani, R.R. and Running, S.W., 1989. Estimation of regional surface resistance to evapotranspiration from NDVI and thermal-IR AVHRR data. Journal of Applied Meteorology, 28(4): 276-284.

Nobel, P.S., 1991. Physiochemical and Environmental Plant Physiology. Academic Press, San Diego, CA.

Norman, J.M., Kustas, W.P. and Humes, K.S., 1995. Two-source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature. Agricultural and Forest Meteorology, 77: 263-293.

Oke, T.R., 1987. Boundary Layer Climates. Routledge, London, 435 pp.

Ottlé, C. and Vidal-Madjar, D., 1992. Estimation of land surface temperature with NOAA9 data. Remote Sensing of Environment, 40: 27-41.

Ou, S.C., Liou, K.N. and Baum, B.A., 1996. Detection of multilayer cirrus cloud systems using AVHRR data: verification based on FIRE II IFO composite measurements. Journal of Applied Meteorology, 35: 178-191.

Pierce, L.L. and Congalton, R.G., 1988. A methodology for mapping forest latent heat flux densities using remote sensing. Remote Sensing of Environment, 24: 405-418.

Poff, N.L. and Ward, J.V., 1989. Implications of streamflow variability and predictability for lotic community structure: a regional analysis of streamflow patterns. Canadian Journal of Fisheries and Aquatic Science, 46: 1805-1818.

Prabhakara, C., Dalu, G. and Kunde, V.G., 1974. Estimation of sea surface temperature from remote sensing in the 11- to 13-μm window region. Journal of Geophysical Research, 79(33): 5039-5044.

Price, J.C., 1984. Land surface temperature measurements from the split window channels of the NOAA 7 Advanced Very High Resolution Radiometer. Journal of Geophysical Research, 89: 7231-7237.

Priestley, C.H.B. and Taylor, R.J., 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. Monthly Weather Review, 100(2): 81-92.

Prihodko, L. and Goward, S.N., 1997. Estimation of air temperature from remotely sensed surface observations. Remote Sensing of Environment, 60: 335-346.

Privette, J.L., Fowler, C., Wick, G.A., Baldwin, D. and Emery, W.J., 1995. Effects of orbital drift on advanced very high resolution radiometer products: normalized differenced vegetation index and sea surface temperature. Remote Sensing of Environment, 53: 164-171.

Qualls, R.J. and Brutsaert, W., 1996. Evaluation of spatially distributed ground-based and remotely sensed data to estimate spatially distributed sensible heat fluxes. Water Resources Research, 32(8): 2489-2495.

Reginato, R.J., Jackson, R.D. and Pinter, P.J., Jr., 1985. Evapotranspiration calculated from remote multispectral and ground station meteorological data. Remote Sensing of Environment, 18: 75-89.

Research Systems, 1995. Interactive Data Language. Research Systems, Inc., Boulder, CO.

Rubio, E., Caselles, V. and Badenas, C., 1997. Emissivity measurements of several soils and vegetation types in the 8-14 μm wave band: analysis of two field methods. Remote Sensing of Environment, 59: 490-521.

Running, S.W. et al., 1994. Terrestrial remote sensing science and algorithms planned for EOS/MODIS. International Journal of Remote Sensing, 15(17): 3587-3620.

Running, S.W., Nemani, R.R. and Hungerford, R.D., 1987. Extrapolation of synoptic meteorological data in mountainous terrain and its use for simulating forest evaporation and photosynthesis. Canadian Journal of Forest Research, 17: 472-483.

Salisbury, J.W. and D'Aria, D.M., 1992. Emissivity of terrestrial materials in the 8-14 μm atmospheric window. Remote Sensing of Environment, 42: 83-106.

Sandholt, I. and Andersen, H.S., 1993. Derivation of actual evapotranspiration in the Senegalese Sahel, using NOAA-AVHRR data during the 1987 growing season. Remote Sensing of Environment, 46: 164-172.

Saunders, R.W. and Kriebel, K.T., 1988. An improved method for detecting clear sky and cloudy radiances from AVHRR data. International Journal of Remote Sensing, 9(1): 123-150.

Schmugge, T.J., Becker, F. and Li, Z.-L., 1991. Spectral emissivity variations observed in airborne surface temperature measurements. Remote Sensing of Environment, 35: 95-104.

Seguin, B., Courault, D. and Guérif, M., 1994. Surface temperature and evapotranspiration: application of local scale methods to regional scales using satellite data. Remote Sensing of Environment, 49: 287-295.

Seguin, B. and Itier, B., 1983. Using midday surface temperature to estimate daily evaporation from satellite thermal IR data. International Journal of Remote Sensing, 4(2): 371-383.

Seguin, B., Lagouarde, J.-P. and Savane, M., 1991. The assessment of regional crop water conditions from meteorological satellite thermal infrared data. Remote Sensing of Environment, 35: 141-148.

Shuttleworth, W.J. and Wallace, J.S., 1985. Evaporation from sparse crops - an energy combination theory. Quarterly Journal of the Royal Meteorological Society, 111: 839-855.

Simpson, J.J. and Gobat, J.I., 1996. Improved cloud detection for daytime AVHRR scenes over land. Remote Sensing of Environment, 55: 21-49.

Smith, R.C.G. and Choudhury, B.J., 1991. Analysis of normalized difference and surface temperature observations over southeastern Australia. International Journal of Remote Sensing, 12(10): 2021-2044.

Sobrino, J. and Caselles, V., 1990. Thermal infrared radiance model for interpreting the directional radiometric temperature of a vegetative surface. Remote Sensing of Environment, 33: 193-199.

Sobrino, J., Coll, C. and Caselles, V., 1991. Atmospheric correction for land surface temperature using NOAA-11 AVHRR channels 4 and 5. Remote Sensing of Environment, 38: 19-34.

Soer, G.J.R., 1980. Estimation of regional evapotranspiration and soil moisture conditions using remotely sensed crop surface temperatures. Remote Sensing of Environment, 9: 27-45.

Stoms, D.M., Bueno, M.J. and Davis, F.W., 1997. Viewing geometry of AVHRR image composites derived using multiple criteria. Photogrammetric Engineering and Remote Sensing, 63(6): 681-689.

Stone, L.R. and Kanemasu, E.T., 1975. Grain sorghum canopy temperature as influenced by clouds. Remote Sensing of Environment, 4: 177-181.

Suits, G.H., 1972. The calculations of the directional reflectance of a vegetative canopy. Remote Sensing of Environment, 2: 117-125.

Taconet, O., Bernard, R. and Vidal-Madjar, D., 1986. Evapotranspiration over an agricultural region using a surface flux / temperature model based on NOAA-AVHRR data. Journal of Climate and Applied Meteorology, 25: 284-307.

Thunnissen, H.A.M. and Nieuwenhuis, G.J.A., 1990. A simplified method to estimate regional 24-h evapotranspiration from thermal infrared data. Remote Sensing of Environment, 31: 211-225.

Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sensing of Environment, 8: 127-150.

Valor, E. and Caselles, V., 1996. Mapping land surface emissivity from NDVI: application to European, African, and South American areas. Remote Sensing of Environment, 57: 167-184.

van de Griend, A.A. and Owe, M., 1994. Bare soil surface resistance to evaporation by vapor diffusion under semiarid conditions. Water Resources Research, 30(2): 181-188.

Verma, S.B. et al., 1993. Photosynthesis and stomatal conductance related to reflectance on the canopy scale. Remote Sensing of Environment, 44: 103-116.

Vidal, A., Pinglo, F., Durand, H., Devaux-Ros, C. and Maillet, A., 1994. Evaluation of a temporal fire risk index in Mediterranean forests from NOAA thermal IR. Remote Sensing of Environment, 49: 296-303.

Walker, G.K. and Hatfield, J.L., 1979. Test of the stress-degree-day concept using multiple planting dates of red kidney beans. Agronomy Journal, 71: 967-971.

Weaver, D.K., Butler, W.E. and Olson, C.E., Jr., 1969. Observations on interpretation of vegetation from infrared imagery. In: P.L. Johnson (Editor), Remote Sensing in Ecology. University of Georgia Press, Athens.

Wiegand, C.L. and Namken, L.N., 1966. Influences of plant moisture stress, solar radiation, and air temperature on cotton leaf temperature. Agronomy Journal, 58: 582-586.

Yost, E. and Wenderoth, S., 1967. Multispectral color aerial photography. Photogrammetric Engineering, 33: 1020-1033.