

4. REGIONÁLIS KLÍMAPARAMÉTEREK BECSLÉSE

Ebben a fejezetben a két kiválasztott mérsékeltövi régióban a havi csapadékösszegre és a Palmer-féle szárazság-index havi értékeire vonatkozó modellbecsléseket elemezzük. Vizsgálataink során kétféle modell eredményességét hasonlítjuk össze. A hagyományos matematikai statisztikai módszerek közül a többváltozós lineáris regressziót alkalmazzuk, s az itt kapott eredményeinket vetjük össze a fuzzy-szabályokból felépülő modellek felhasználásával nyert eredményekkel. Mindkét modell típusban a nagy térséget jellemző makrocirkulációs helyzeteket, valamint az éghajlati oszcillációs jelenségeket (ENSO, NAO) jellemző karakterisztikákat tekintjük meghatározó tényezőnek. A kétféle módszer ismertetése után kerül sor a különböző alkalmazások bemutatására, a modellek érzékenységének vizsgálatára, s az eredmények összehasonlítására.

4.1. Hagományos és új fejlesztésű becslési módszerek

A hagyományos matematikai statisztikai becslési módszerek közül a szakirodalomban leggyakrabban használt lineáris regressziós modellek alapjait tárgyaljuk röviden összefoglalva. Majd ezt követi a fuzzy-szabályokból álló modellek felépítésének bemutatása.

4.1.1. A lineáris regresszió alapuló statisztikai modellek

Az egyik legtöbbször alkalmazott matematikai statisztikai módszer a lineáris regresszió (Neter et al., 1996), ahol egy prediktandusz változót (Y) $n-1$ darab prediktor változó (X_i) lineáris kombinációjából és egy normális eloszlású hibatagból állítunk elő a megfigyelt adatsorok felhasználásával. A többváltozós lineáris regressziós egyenlet a következő alakban írható fel:

$$Y_j = \beta_0 + \sum_{i=1}^{n-1} \beta_i \cdot X_{ij} + \varepsilon_j ,$$

ahol Y_j a prediktandusz változó j -edik értéke; X_{ij} az i -edik prediktor változó j -edik értéke; β_i regressziós együtthatók az egyes X_i prediktor egységnyi megváltozásához tartozó Y -beli változásokat fejezik ki, amennyiben a többi prediktor értéke állandó marad; β_0 regressziós együttható a prediktor változók zérus értéke esetén adja meg a prediktanduszt, amennyiben ez az eset értelmezhető; az ε_j a hibatagot kifejező valószínűségi változó j -edik értéke, melynek

várható értéke 0, szórásnégyzete σ^2 , és egymás közötti kovarianciájuk 0; $(n-1)$ pedig a prediktorok száma.

A lineáris regressziós egyenletet mátrix alakban is felírhatjuk:

$$\mathbf{Y} = \mathbf{X} \cdot \boldsymbol{\beta} + \boldsymbol{\varepsilon} ,$$

ahol \mathbf{Y} a prediktandusz változó értékeit tartalmazza $k \times 1$ -es oszlopmátrix formájában (k az idősor hossza); \mathbf{X} a prediktor változók értékeit tartalmazza $k \times n$ -es mátrix formájában (az 1. oszlopban 1-esek állnak); $\boldsymbol{\beta}$ a regressziós együtthatókat tartalmazza $n \times 1$ -es oszlopmátrix formájában (β_0 -tól indulva β_{n-1} -ig); $\boldsymbol{\varepsilon}$ pedig a véletlen hiba értékeit tartalmazza $k \times 1$ -es oszlopmátrix formájában.

A regressziós együtthatók meghatározásának szokásos módja a legkisebb négyzetek felhasználásával történik, vagyis az alábbi Q érték minimumát kell megkeresni (Neter et al., 1996):

$$Q = \sum_{j=1}^k \left(Y_j - \beta_0 - \sum_{i=1}^{n-1} \beta_i \cdot X_{ij} \right)^2 .$$

A szükséges számítások elvégzése után előáll a regressziós együtthatók torzítatlan becslését tartalmazó $\mathbf{b} = \hat{\boldsymbol{\beta}}$, $n \times 1$ -es oszlopmátrix, mely az alábbi módon fejezhető ki (Neter et al., 1996):

$$\mathbf{b} = (\mathbf{X}' \cdot \mathbf{X})^{-1} \cdot (\mathbf{X}' \cdot \mathbf{Y}) .$$

Ekkor a regressziós egyenletből a prediktandusz változó ismeretlen értékeinek becslését a következő mátrixegyenlettel adhatjuk meg:

$$\hat{\mathbf{Y}} = \mathbf{X} \cdot \mathbf{b} ,$$

ahol $\hat{\mathbf{Y}}$ a prediktandusz változó ismeretlen értékeit tartalmazza $l \times 1$ -es oszlopmátrix formájában (l az idősor hossza); \mathbf{X} a prediktor változók értékeit tartalmazza $l \times n$ -es mátrix formájában (az 1. oszlopban 1-esek állnak); \mathbf{b} pedig a regressziós együtthatók torzítatlan becsléseit tartalmazza $n \times 1$ -es oszlopmátrix formájában (b_0 -tól indulva b_{n-1} -ig).

Vizsgálataink során a lineáris regressziós modellek prediktorai az MCP osztályok havi relatív gyakoriságai, valamint az ENSO-t illetve a NAO-t jellemző indexek időben eltolt idősorai. A modellekben szereplő prediktandusz az alkalmazás célja szerint változik: így a Palmer-féle szárazsági index (PDSI illetve PMDI) havi értékeit, a havi csapadékösszeget különböző klímaállomásokon, illetve a Balatonra vonatkozó hidrológiai mérleg egyes tagjait (csapadék, területi párolgás, felszíni hozzáfolyás, vízkészletváltozás) modellezzük a

továbbiakban. A lineáris regressziós modellekben szereplő fenti prediktorokat és prediktanduszokat a **4-1. táblázat**ban foglaljuk össze (a következő alfejezetben ismertetésre kerülő fuzzy-szabályokból felépülő modellek is ugyanezeket a változókat vizsgálják).

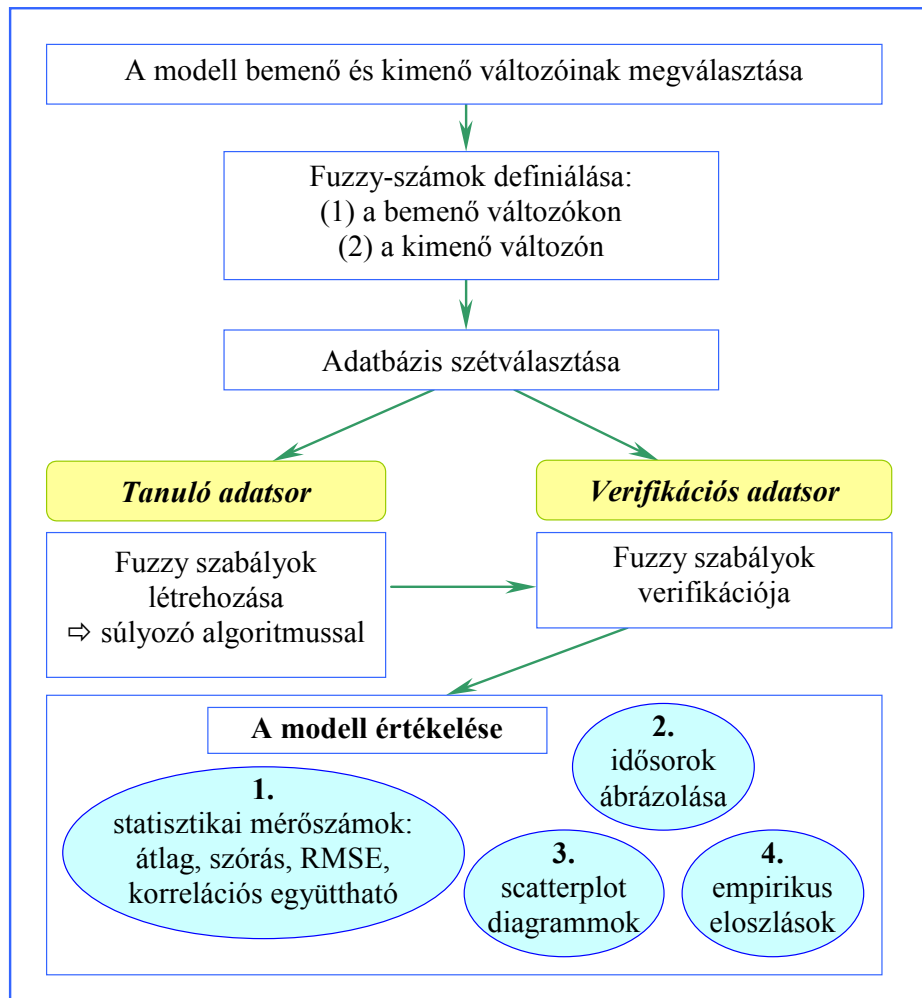
4-1. táblázat: A lineáris regressziós modellek prediktanduszai és prediktorai

Prediktandusz	Prediktorok
Havi PMDI Nebraska állam régióiban (1946-94)	MCP osztályok havi relatív gyakoriságai SOI havi idősorok, a késleltetett hatás figyelembevételével
Havi PDSI hazai klímaállomásokon (1901-99)	MCP osztályok havi relatív gyakoriságai SOI havi idősorok, a késleltetett hatás figyelembevételével
Havi PDSI hazai klímaállomásokon (1950-99)	MCP osztályok havi relatív gyakoriságai A NAO-t jellemző SST-index havi idősorai, a késleltetett hatás figyelembevételével
Magyarországon mért havi csapadékösszegek (1901-99)	MCP osztályok havi relatív gyakoriságai SOI havi idősorok, a késleltetett hatás figyelembevételével
Magyarországon mért havi csapadékösszegek (1950-99)	MCP osztályok havi relatív gyakoriságai A NAO-t jellemző SST-index havi idősorai, a késleltetett hatás figyelembevételével
A Balaton vízháztartási elemeinek havi idősorai (1950-95)	MCP osztályok havi relatív gyakoriságai A NAO-t jellemző SST-index havi idősorai, a késleltetett hatás figyelembevételével

4.1.2. Fuzzy-szabályokból felépülő modellek

A hagyományos matematikai statisztikai módszerek mellett fuzzy-szabályokon alapuló modellezést is alkalmaztunk az MCP relatív gyakoriságok, az oszcillációs indexek és a regionális éghajlati változók közötti kapcsolat kimutatására. A fuzzy-szabályok lehetővé teszik, hogy hasonló bemenő feltételek mellett többféle lehetséges kimenetelt figyelembe vegyünk különböző súllyal, s ezek eredőjéből határozzuk meg a bekövetkező végső választ (Zadeh, 1965). A fuzzy-szabályokból felépülő módszer a hagyományos modellekkel szemben lényegesen egyszerűbb, nem követeli meg a vizsgált idősorok függetlenségét, az elemzendő folyamatok linearitását, továbbá rövidebb idősorokra is alkalmazható (Galambosi et al., 1999).

Ebben az alfejezetben a fuzzy-szabályokból felépített modelleket (Pongrácz et al., 1999a) mutatjuk be, melyek vázlatos szerkezetét a **4-1. ábrán** láthatjuk.



4-1. ábra: A fuzzy-szabályokból felépülő modellek vázlatos szerkezete.

A modellben szereplő X_i meghatározó tényezők (MCP osztályok havi relatív gyakorisága, illetve a különféle oszcillációs indexek havi értékei) és regionális klímaoutputok (Y) megválasztása után minden változóra fuzzy-számokat definiálunk. Egy adott A fuzzy-számot $(x, \mu(A, x))$ számpárok alkotnak, ahol x valamilyen alaphalmazon (pl. a valós számok \mathbf{R} halmazán) értelmezett folytonos változó, $\mu(A, x)$ pedig a tagsági függvény, mely megmondja, hogy adott x helyen a fuzzy-számmal jelölt tulajdonság milyen mértékben teljesül. Definíció szerint a tagsági függvény értéke 0 és 1 között változik, továbbá nincs lokális minimuma

(Dubois és Prade, 1980). Az 1 tagsági függvényértékhez tartozó alaphalmazbeli pontok alkotják az A fuzzy-szám magját, melyet $core(A)$ -val jelölünk:

$$core(A) = \{x \mid \mu(A,x) = 1\}.$$

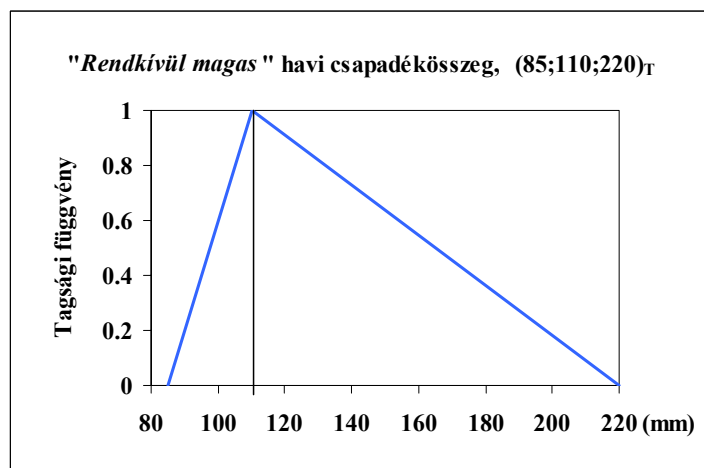
Az alaphalmaz azon pontjainak összességét, amelyek tagsági függvény értéke 0-nál nagyobb, az A fuzzy-szám tartójának nevezzük, s $supp(A)$ -val jelöljük:

$$supp(A) = \{x \mid \mu(A,x) > 0\}.$$

Az általunk felépített modellekben az egyik legegyszerűbb fuzzy-szám típust, az ún. háromszög fuzzy-számot használtuk (Kóczy & Tikk, 2000), mely a tagsági függvény alakjáról kapta a nevét. Az $A = (a_1; a_2; a_3)_\Gamma$ háromszög fuzzy-szám tagsági függvényének definíciója a következőképpen írható fel:

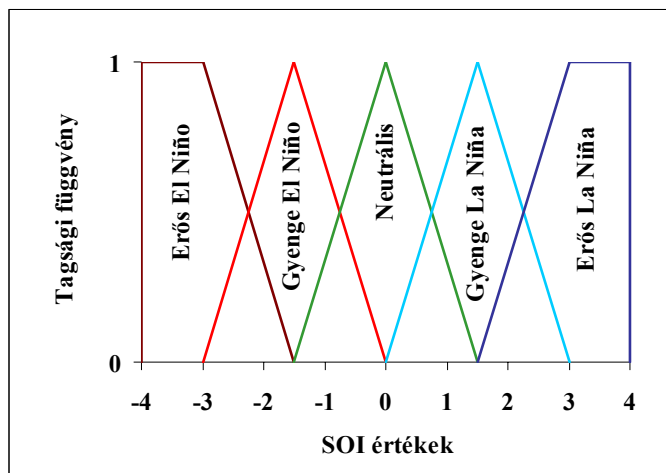
$$\mu(A,x) = \begin{cases} 0 & , \text{ ha } x \leq a_1 \\ \frac{x - a_1}{a_2 - a_1} & , \text{ ha } a_1 < x < a_2 \\ 1 & , \text{ ha } x = a_2 \\ \frac{x - a_3}{a_2 - a_3} & , \text{ ha } a_2 < x < a_3 \\ 0 & , \text{ ha } x \geq a_3 \end{cases}$$

ahol $a_2 = core(A)$ és $[a_1; a_3] = supp(A)$.



4-2. ábra: Háromszög fuzzy-szám a rendkívül csapadékos hónapok jellemzésére Magyarország területén.

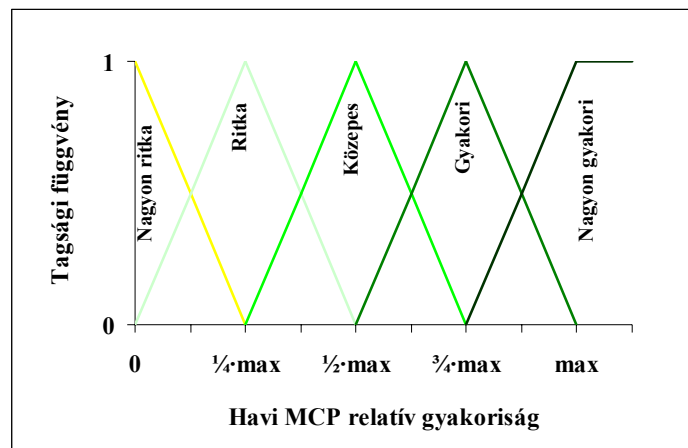
Tekintsünk egy meteorológiai példát a háromszög fuzzy-szám tagsági függvényének definíciójára. A **4-2. ábrán** látható a "rendkívül magas" havi csapadékösszeg tagsági függvénnyel történő értelmezése. Ha ugyanezt csupán egy határérték kiválasztásával próbálnánk elérni, az nem lenne megfelelő – hiszen egy éles határvonallal nem dönthetjük el, hogy például a 85 mm nagyságú havi csapadékösszeg még nem rendkívül magas, de a 86 mm már az. Hasonlóképpen nem elégedhetünk meg egy intervallummal sem, mely szerint a 86 mm-es és a 115 mm-es havi csapadékmennyiséget egyformán rendkívül magasnak tekintenénk, holott óriási köztük a különbség. A fuzzy-számok alkalmazásának lényeges előnye, hogy az összes lehetséges csapadékösszeg érték valamekkora tagsági függvény értékkel "rendkívül magas" mennyiségnek tekinthető, természetesen a nagyon kis havi csapadékértékek tagsági függvénye 0. Válasszuk például a $(85;110;220)_T$ háromszög fuzzy-számot a "rendkívül magas" havi csapadékösszeg jellemzésére (**4-2. ábra**). Ekkor a 86 mm-es és a 115 mm-es havi csapadékmennyiség egyaránt "rendkívül magas", de különböző mértékben: az előbbihez 0,08-as, az utóbbihoz viszont 0,96-os tagsági függvényérték tartozik. Így elérhetjük, hogy az egymástól csupán kis mértékben eltérő csapadékösszegek fuzzy jellege is csak kevésbé tér el.



4-3. ábra: A SOI havi értékein definiált fuzzy-számok ötelemű rendszere.

Többféle módon is megválaszthatjuk a modellek bemenő és kimenő változóiin definiált fuzzy-számok rendszerét ($A(l_i)$, $B(m)$, ahol l_i a fuzzy-számok elemszáma az i -dik bemenő változón, m pedig a kimenő változón definiált fuzzy-szám rendszer elemszáma). Vizsgálataink során végeztünk erre vonatkozó összehasonlításokat, melyekről az eredmények

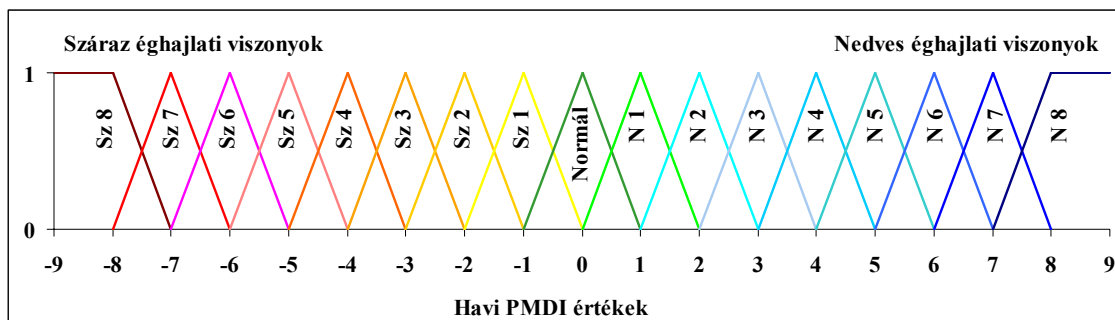
tárgyalásánál ejtünk szót. A 4-3. és 4-4. ábrákon a SOI havi értékein, illetve az MCP osztályok havi relatív gyakoriságain definiált fuzzy-szám rendszereket mutatjuk be. Mindkét esetben öt fuzzy-számból álló rendszert tekintve az MCP osztályoknál a „Nagyon ritka”, „Ritka”, „Közepes”, „Gyakori” és „Nagyon gyakori” fuzzy-számokat definiálhatjuk (4-4. ábra), míg a SOI értékekre az „Erős El Niño”, „Gyenge El Niño”, „Neutrális”, „Gyenge La Niña”, „Erős La Niña” fázist reprezentáló fuzzy-számokat (4-3. ábra). Az egyes MCP osztályok havi előfordulási valószínűsége egyáltalán nem azonos, vannak olyan cirkulációs viszonyok, melyek sokkal gyakoribbak, mint más jóval ritkábban kialakuló időjárási helyzetek. A fuzzy-számok definiálásakor ezt figyelembe vettük, s a maximálisan lehetséges havi relatív gyakoriságok [0;1] intervallumát minden MCP osztály esetén leszűkítettük a valójában előfordult leggyakoribb esetszámra (ezeket jelöltük *max*-szal az ábrán) – az alsó határ 0 értékét nem változtattuk meg, mivel szinte az összes MCP osztályra létezett olyan hónap, amikor az adott cirkulációs helyzet egyáltalán nem fordult elő.



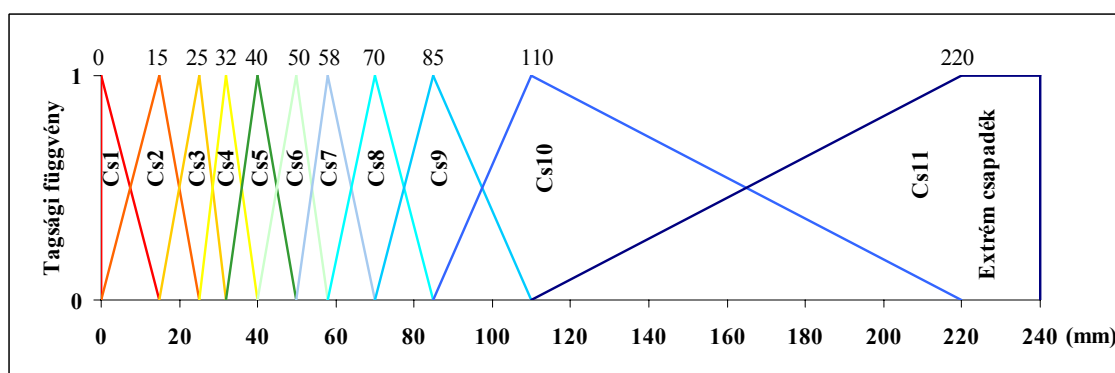
4-4. ábra: Az MCP osztályok havi relatív gyakoriságain definiált fuzzy-számok ötelemű rendszere. (*max*: az adott MCP osztályra a valaha előfordult maximális havi relatív gyakoriság értéke)

Modelljeink regionális klímaoutputjának megválasztása nagymértékben befolyásolja, hogy milyen fuzzy-szám rendszert definiálhatunk a kimenő változókra. Szimmetrikus eloszlást követő klímaméter esetén végezhetünk lineáris típusú skála felbontást (pl. a havi PMDI értékekre a 4-5. ábra szerinti rendszer), míg a ferde eloszlású éghajlati változók esetén a fuzzy-számok rendszere is követi az eloszlás aszimmetriáját, s így a decilisek jelölik ki a fuzzy-számok magját (pl. a havi csapadékmennyiségre a 4-6. ábra szerinti fuzzy-szám

rendszer). A PMDI értékeken definiált rendszerben a mag 1-esével növekedve az „*Extrém száraz*”-tól az „*Extrém nedves*” fuzzy-számig jut el, a csapadékösszegezen értelmezett fuzzy-számok rendszere pedig a „*Csapadékmentes*”-tól az „*Extrém csapadékos*” viszonyokig terjed.



4-5. ábra: A PMDI havi értékein definiált fuzzy-számok rendszere.
 (Sz 8: „*Extrém száraz*”, Sz 7: „*Rendkívül száraz*”, Sz 6: „*Nagyon száraz*”, Sz 5: „*Erősen száraz*”, Sz 4: „*Közepesen száraz*”, Sz 3: „*Eléggé száraz*”, Sz 2: „*Kicsit száraz*”, Sz 1: „*Enyhén száraz*”, N 1: „*Enyhén nedves*”, N 2: „*Kicsit nedves*”, N 3: „*Eléggé nedves*”, N 4: „*Közepesen nedves*”, N 5: „*Erősen nedves*”, N 6: „*Nagyon nedves*”, N 7: „*Rendkívül nedves*”, N 8: „*Extrém nedves*”)



4-6. ábra: A havi csapadékösszegezen (Keszthely) definiált fuzzy-számok rendszere.
 (Cs1: „*Csapadékmentes*”, Cs2: „*Száraz*”, Cs3: „*Kevéssé csapadékos*”, Cs4: „*Enyhén csapadékos*”, Cs5: „*Kicsit csapadékos*”, Cs6: „*Eléggé csapadékos*”, Cs7: „*Közepesen csapadékos*”, Cs8: „*Erősen csapadékos*”, Cs9: „*Nagyon csapadékos*”, Cs10: „*Rendkívül csapadékos*”, Cs11: „*Extrém csapadékos*”)

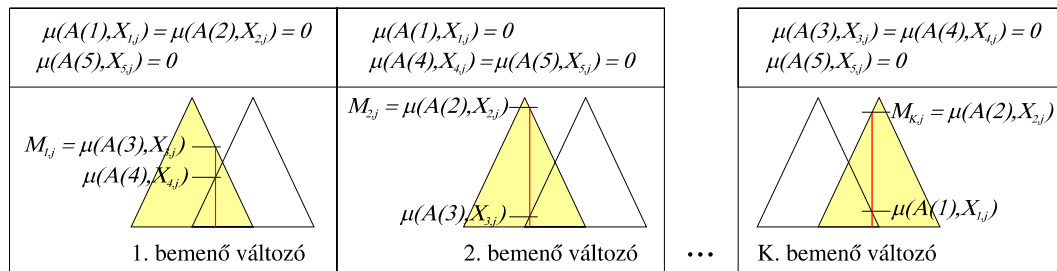
A fuzzy-számok definiálása után a teljes adatbázist két részre osztjuk: a tanuló τ és a verifikációs ν időszakra. A tanuló adatsor szolgál a fuzzy-szabályok létrehozására, míg a

verifikációs adatsor segítségével értékeljük a megalkotott modell "jóságát". A teljes idősor felosztásánál általában 2:1 arányt alkalmaztunk. Mindenegyes output típusra többféle konkrét szétválasztást is kipróbáltunk, de a modellek által előállított becslések és a tényleges mért adatok közötti illeszkedésben a felosztások változtatása nem okozott lényeges különbséget.

A modellek legfontosabb része a fuzzy-szabályok megalkotása, melyet az ún. *súlyozó algoritmussal* (Bárdossy & Duckstein, 1995) végzünk el a tanuló időszak felhasználásával. Az algoritmus lépéseit az alábbiakban foglaljuk össze.

1. Adott j időpontban meghatározzuk a bemenő változókat tartalmazó adatvektor összes $\mu(A(l_i), X_{i,j})$ tagsági függvényét.
2. Minden bemenő változóra az $M_{i,j}$ maximális tagsági függvényértékkel rendelkező fuzzy-számot vesszük figyelembe (4-7. ábra):

$$M_{i,j} = \max_{l_i=1,\dots,5} (\mu(A(l_i), X_{i,j}))$$



4-7. ábra: A maximális tagsági függvény értékű fuzzy-szám meghatározása a bemenő változókon definiált 5 elemű fuzzy-szám rendszerek esetén

3. A j időpontban meghatározzuk a kimenő változó összes $\mu(B(m), Y_j)$ tagsági függvényét.
4. A kimenő változóra is az $M_{0,j}$ maximális tagsági függvény értékkel rendelkező fuzzy-számot vesszük figyelembe.

$$M_{0,j} = \max_{m=1,\dots,18} (\mu(B(m), Y_j))$$

5. A bemenő változók együttes hatásának meghatározása az **ÉS** operátor alkalmazásával (Zimmermann, 1985). A kiválasztott fuzzy-számok, valamint az **ÉS** operátor felhasználásával egy fuzzy-szabály általános alakja a következő:

$$\mathbf{HA} (X_{1,j} \in A(l_1) \mathbf{ÉS} X_{2,j} \in A(l_2) \mathbf{ÉS} \dots \mathbf{ÉS} X_{K,j} \in A(l_K)) \mathbf{AKKOR} (Y_j \in B(m))$$

A szabály együttes hatását az ún. illeszkedési vagy hasonlósági mérték (degree of fulfillment, DOF) adja meg, melyet az $M_{i,j}$ maximális tagsági függvény értékek szorzatával állíthatunk elő.

$$DOF_j = \prod_{i=1}^K M_{i,j}$$

6. A súlyozási értékek hozzárendelése az adott szabályhoz. A tanuló időszak alapján létrehozott fuzzy-szabályok mindegyike esetén a hozzárendelt súlytényező adja meg, hogy az adott szabály milyen mértékben ad magyarázatot a kimenő változó viselkedésére. A súlytényezőt a szabály előfordulási aránya határozza meg, mely a szabály megjelenésekor kiszámolt ω_j segédsúlytényezők összegeként (a 7. lépésben) definiálunk.

$$\omega_j = DOF_j \cdot M_{0,j}.$$

7. Megvizsgáljuk, hogy a vizsgált adatvektorból adódó fuzzy-szabály szerepel-e már a szabályok között. Ha még nem fordult elő eddig, akkor felvesszük a 6. lépés szerint kiszámolt ω_j súllyal a szabályrendszerbe. Ha viszont már szerepel a szabályrendszerben, akkor az adott fuzzy-szabály eddigi súlytényezőjét módosítjuk úgy, hogy hozzáadjuk az ω_j segédsúlyértéket az eddigi súlyértékhez. A τ tanuló időszakon végig haladva végső soron az r -edik fuzzy-szabály ω_r súlytényezője az alábbi módon áll elő:

$$\omega_r = \sum_{\tau} DOF_j \cdot M_{0,j}.$$

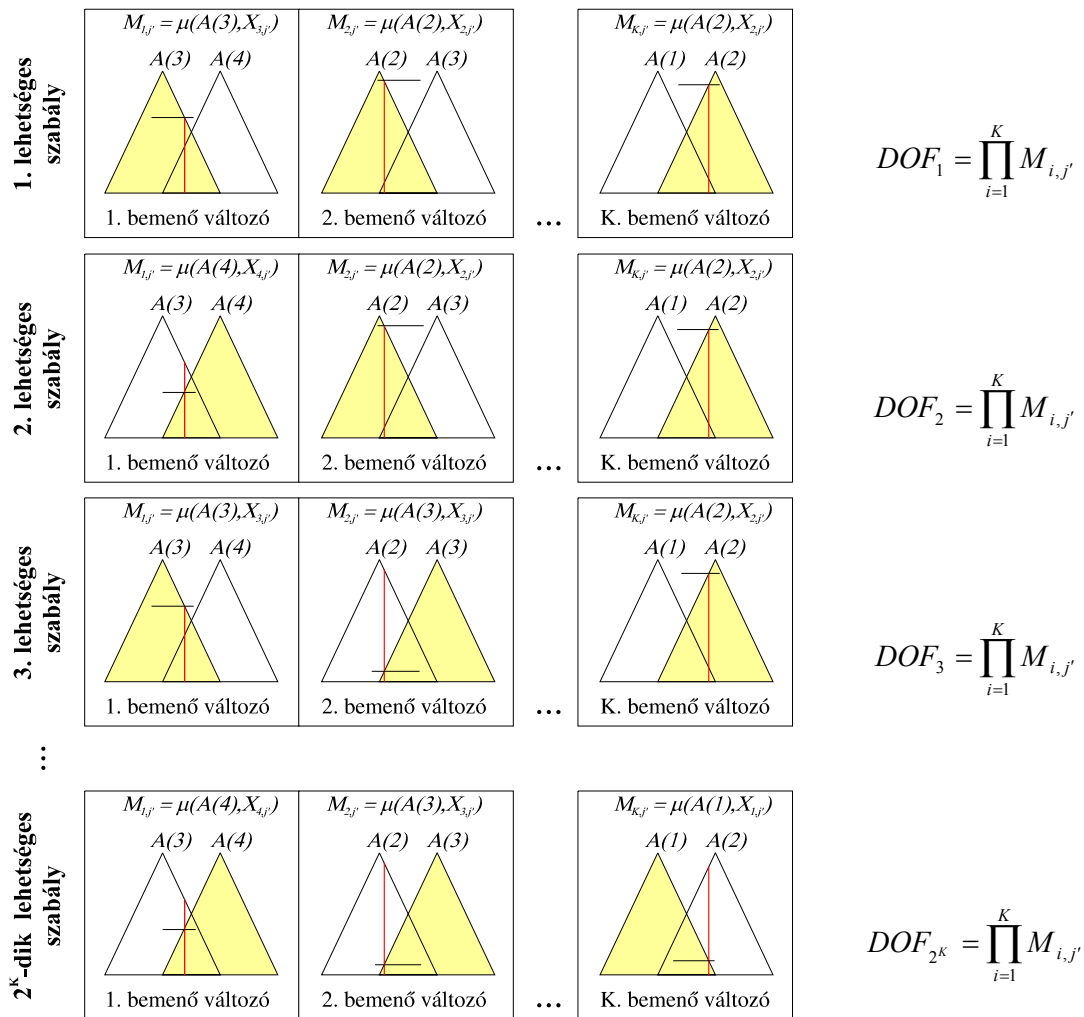
8. Visszatérünk az 1. lépéshez, amíg a teljes tanuló időszakot fel nem dolgoztuk.

Miután a tanuló időszakból meghatároztuk a modell fuzzy-szabályainak rendszerét, minden fuzzy-szabályhoz tartozik egy súlytényező. Ezt követi a verifikációs eljárás, melynek során a szabályok felhasználásával a kimenő változóra becslést adunk, majd értékeljük a létrehozott modellt. A verifikáció lépéseit az alábbiakban foglaljuk össze.

1. A ν verifikációs időszak j' adatvektorához tartozó valamennyi lehetséges DOF-et kiszámítjuk a tagsági függvényértékek felhasználásával (**4-8. ábra**).

Míthogy a fuzzy-számok rendszerét minden bemenő változón úgy definiáltuk, hogy egy adat pontosan két fuzzy-szám esetén ad nem-nulla tagsági függvényértéket, ebből adódóan a lehetséges fuzzy-szabályok (és így a hozzájuk tartozó DOF-ek)

száma 2^K , ahol K a bemenő paraméterek száma. Ha például tíz bemenő változóval számolunk, akkor $2^{10} = 1024$ különböző szabályt kellene figyelembe vennünk egy-egy adatvektor esetén. Ehelyett azonban csupán néhány jön számításba: azok, melyek a tanuló időszak alatt ténylegesen előfordultak (s így súlytényezőjük 0-nál nagyobb).



4-8. ábra: Az összes lehetséges DOF meghatározása a verifikációs eljárás során.

2. Az ún. defuzzifikáció során a modell által adott fuzzy-válaszok kombinációjaként állítjuk elő a kimenő változó szimulált értékét ($\hat{Y}_{j'}$). A teljes szabályrendszerből alkalmazható l fuzzy-szabályok együttes figyelembevételével a súlypontot meghatározó defuzzifikációs módszert (Van Leekwijck & Kerre, 1999) felhasználva kapjuk, hogy:

$$\hat{Y}_j = \frac{\sum_{l \in \tau} DOF_l \cdot \omega_l \cdot core(B_l)}{\sum_{l \in \tau} DOF_l \cdot \omega_l}.$$

3. Visszatérünk az 1. lépéshez, amíg a teljes időszakra elő nem állítjuk a kimenő változó szimulált idősorát.
4. Végül értékeljük a fuzzy-szabályokból alkotott modell eredményességét. Az értékeléshez különféle statisztikai vizsgálatokat végzünk el. Összehasonlítjuk a szimulált és a tényleges idősorokat a statisztikai jellemzők segítségével (pl.: átlag, szórás, kvantilisek), meghatározzuk a két idősor közötti korrelációs együtthatót. A grafikus értékelés részeként ábrázolhatjuk magukat az idősorokat, a hisztogramjukat, elkészíthetjük a tényleges és a becült idősor szórásdiagramját, valamint összehasonlíthatjuk a két idősor eloszlását az empirikus eloszlásfüggvények illeszkedési vizsgálatával (*Dévényi és Gulyás, 1988*). A modell becslési hibáját meghatározhatjuk a közepes hibával (*ME*), a közepes abszolút hibával (*MAE*), valamint a közepes négyzetes hiba négyzetgyökével (*RMSE*):

$$ME = \frac{1}{J} \cdot \sum_{j=1}^J (\hat{Y}_j - Y_j), \quad MAE = \frac{1}{J} \cdot \sum_{j=1}^J |\hat{Y}_j - Y_j|, \quad RMSE = \sqrt{\frac{1}{J} \cdot \sum_{j=1}^J (\hat{Y}_j - Y_j)^2}.$$

A fuzzy-szabályokból felépülő modellezési módszer alkalmazásával készített szimulációk eredményeit a következő alfejezetben mutatjuk be. Ezen modellek felhasználhatóságát vetjük össze a többváltozós lineáris regressziós modellekből kapott eredményeinkkel.

4.2. A Palmer-féle szárazsági indexre készített becslés

Mindkét kiválasztott régió esetén a fuzzy-szabályokból álló modellek egyik bemenő paramétere az ENSO-t jellemző havi bontású SOI idősor. Az adott hónapra érvényes SOI értéken kívül az ENSO késleltetett távhatását is figyelembe vettük, melyhez kiszámítottuk a havi Palmer-index (hazánk esetén a PDSI, Nebraska esetén a PMDI) idősorai és a SOI havonkénti elcsúsztatásával létrehozott adatsorok közötti korrelációs együtthatókat (*Pongrácz et al., 1999a*). A modellekbe minden magyar klímaállomás illetve Nebraska állambeli éghajlati régió esetén a legerősebb kapcsolatú elcsúsztatások kerültek be. A modellek másik típusú bemenő paramétere a makrocirkulációs helyzetek havi relatív gyakoriságai.

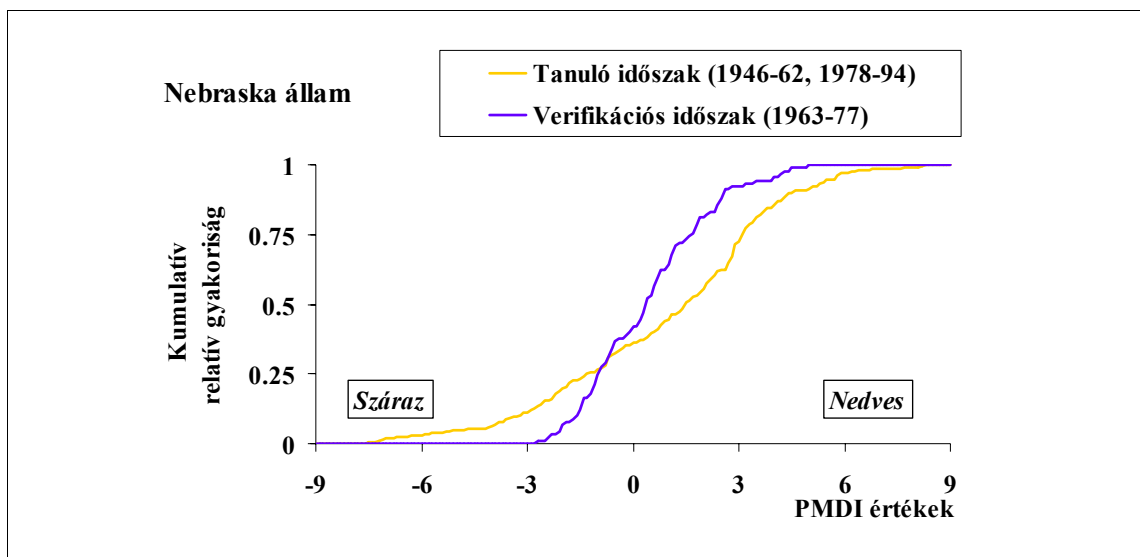
Értelemszerűen Nebraska állam és a Kárpát-medence nagy távolsága miatt a két területre különböző MCP-osztályozást használtunk, melyeket a 2.1. alfejezetben ismertettünk. Az Atlanti-Európai térségre a nemzetközi szakirodalomban is elterjedt Hess-Brezowsky-féle MCP típusok osztályait alkalmaztuk (melyekről az említett alfejezet A./ részében írtunk részletesen). Nebraska esetében pedig a cirkulációs mezők változékonyságának nagy hányadáért felelős első 9 EOF módust felhasználva definiáltuk a különböző MCP típusokat (Matyasovszky et al., 1993).

Modelleztünk csak az ENSO illetve csak a légköri makrocirkulációs feltételekre támaszkodva is (Pongrácz et al., 1999a; 2000; 2002b), melyek eredményeit a következőkben összehasonlítjuk a kétféle információ együttes alkalmazása esetén kapott szimulált idősorokkal.

4.2.1. Az amerikai Nebraska állam területére készített modellezési eredmények

A nyolc Nebraska állambeli éghajlati régió PMDI idősorát elemző modellekben a bemenő változók egyrészt az ENSO jelenséget és annak késleltetett távhatását jellemző havi bontású SOI idősorok, másrészt pedig a nagy-térségű cirkulációt jellemző hat MCP-osztály havi relatív gyakoriságai voltak (Pongrácz et al., 1998a; 1998b; 1999b; 1999c). A kétféle meghatározó tényező közötti kapcsolatot a 3.1. alfejezetben már bemutattuk.

Mivel Nebraska államban elsősorban a nyári félévben válhat kritikussá a szárazság, ezért az elemzéseinkben is az április-szeptember közötti hat hónapot vizsgáltuk. A fuzzy-szabályok megalkotásához felhasznált tanuló idősor az 1946-62. és 1978-94. időszakokat tartalmazza, míg a verifikációs idősor az 1963-77. közötti időszakot öleli fel. A Nebraska állam teljes területére megadott két idősor empirikus eloszlásfüggvényét hasonlítja össze a **4-9. ábra**. Kolmogorov-Szmirnov-féle illeszkedési vizsgálatot elvégezve, a két eloszlás egyezése a 0,01-es szignifikancia szinten elvethető. Ehhez hasonlóan Nebraska éghajlati kistérségeire is igaz, hogy a tanuló és a verifikációs idősorok eloszlása szignifikánsan különbözik egymástól.



4-9. ábra: A tanuló és a verifikációs PMDI idősor empirikus eloszlásfüggvénye Nebraskában.

A súlyozó algoritmussal kapott számos fuzzy-szabály közül egy példát mutatunk be itt. A Közép-Dél-Nebraska éghajlati régiót vizsgáló modell egyik szabálya a következő:

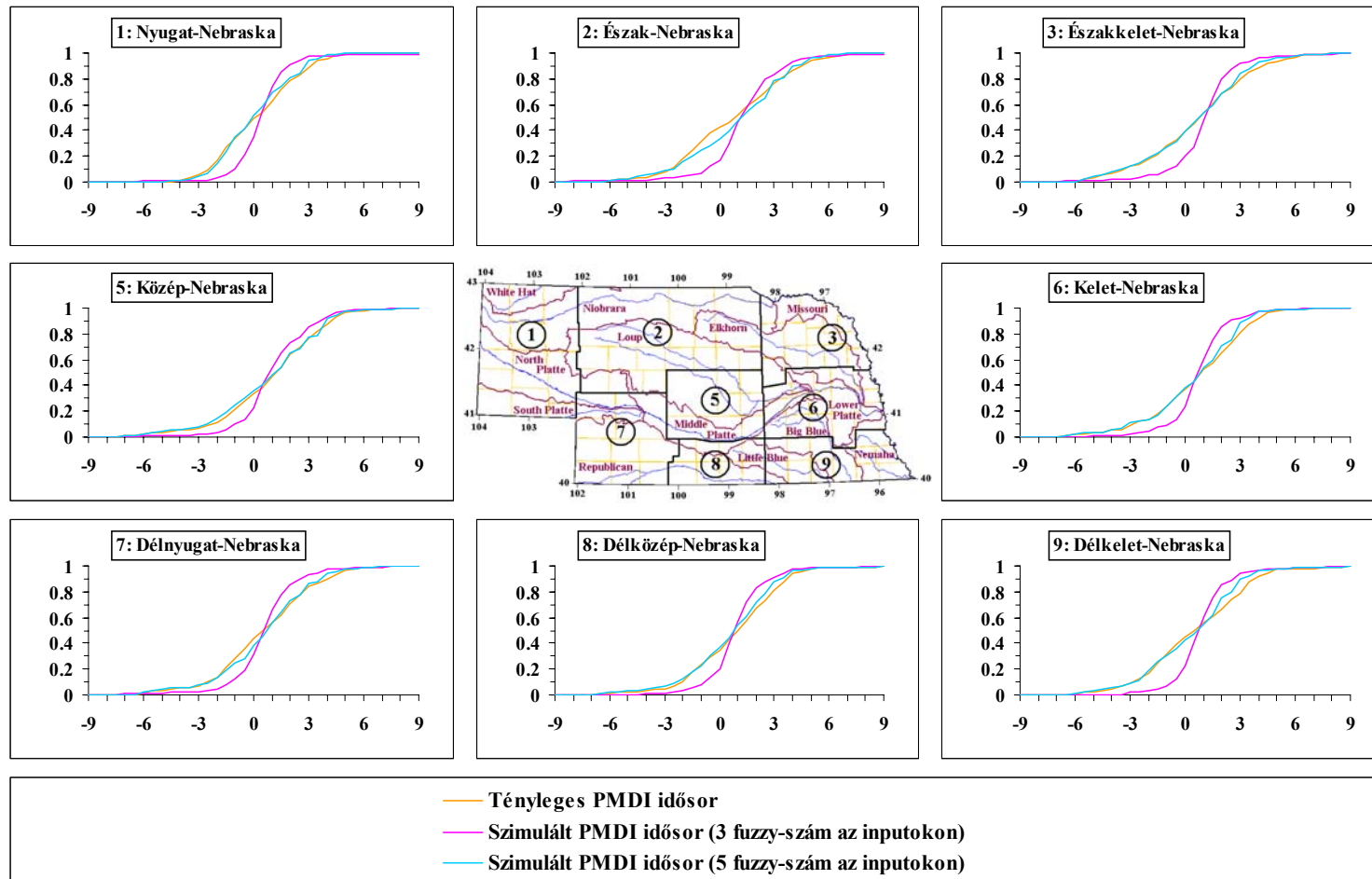
HA

(Közepes MCP1 havi relatív gyakoriság **ÉS** Nagyon ritka MCP2 havi relatív gyakoriság **ÉS**
 Ritka MCP3 havi relatív gyakoriság **ÉS** Nagyon ritka MCP4 havi relatív gyakoriság **ÉS**
 Nagyon ritka MCP5 havi relatív gyakoriság **ÉS** Gyakori MCP6 havi relatív gyakoriság **ÉS**
 Gyenge El Niño fázis az adott hónapban **ÉS** Neutrális fázis 2 hónappal korábban **ÉS**
 Neutrális fázis 4 hónappal korábban **ÉS** Neutrális fázis 6 hónappal korábban)

AKKOR

(Kicsit száraz PMDI-vel jellemzett éghajlati viszonyok)

A fuzzy-szabályokból felépülő modellek értékelésekor megvizsgáltuk, hogy a bemenő paraméterekre vonatkozó fuzzy-számok rendszereiben az elemszám miként befolyásolja a szimuláció eredményességét. Az optimális 5-5 fuzzy-szám helyett 3-3 fuzzy-számot definiálva a szimulált és a tényleges PMDI idősorok empirikus eloszlásfüggvényét hasonlítjuk össze a 4-10. ábrán. Jól látható, hogy minden éghajlati régióban kevesebb fuzzy-szám esetén a szélsőséges éghajlati viszonyokat nem tudják a modellek jól reprodukálni, míg 5-5 fuzzy-szám esetén a régiónkénti két eloszlás szignifikánsan (0,01-es szinten) nem különbözik egymástól.



4-10. ábra: A PMDI szimulált eloszlásainak összevetése a tényleges empirikus eloszlással különböző elemszámú (3 illetve 5) fuzzy-szám rendszereket definiálva a bemenő paramétereken (1963-77)

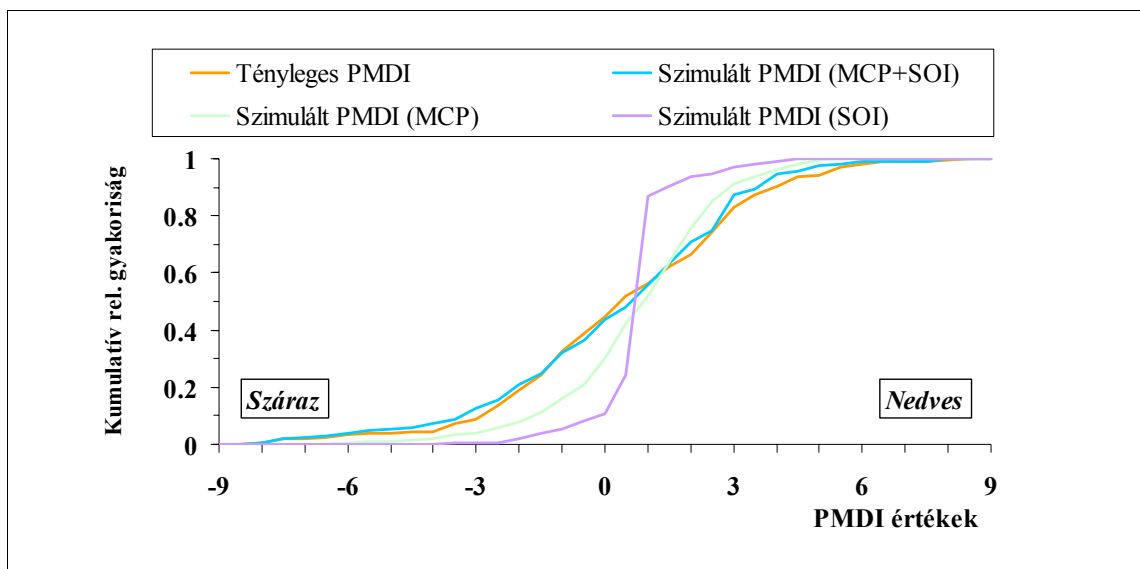
A különböző éghajlati kistérségekre elvégzett elemzéseket a **4-2. táblázat** hasonlítja össze a modellhibák szempontjából. Mind a MAE-t, mind az RMSE-t tekintve jelentősen javulnak a modellek 5 elemből álló fuzzy-szám rendszerek alkalmazása esetén, s a tényleges PMDI idősorok szórásánál (2,1-2,9 közé esnek a szórásértékek, melyeket a **4-12. ábrán** is bemutatunk) jelentősen kisebbek ezek a modellhibák. Néhány kistérséget kivéve általában az ME is csökken, azonban meg kell jegyeznünk, hogy ennél a hibafajtánál a szimulációk alá- és fölébecslései kiejthetik egymást. Végezetül a szimulált és a tényleges PMDI idősorok közötti kapcsolat rendre 0,1-0,5-ös illetve 0,7-0,8-as korrelációs együtthatóval jellemezhető a bemenő paraméterekre vonatkozó 3 illetve 5 elemszámú fuzzy-szám rendszerek alkalmazása esetén.

4-2. táblázat: A modellek hibái a bemenő paramétereken definiált 3 illetve 5 elemszámú fuzzy-szám rendszerek esetén. Összehasonlításképp a tényleges PMDI idősor szórása látható.

Régió		1	2	3	5	6	7	8	9	NE
ME	(3)	0,91	0,83	0,76	-0,01	-0,04	-0,03	0,07	0,31	0,26
	(5)	-0,05	0,18	-0,04	-0,09	-0,17	0,05	-0,17	-0,15	-0,12
MAE	(3)	2,44	2,50	2,48	2,10	1,84	2,06	1,75	1,90	2,08
	(5)	0,90	1,07	0,94	1,00	0,93	1,02	0,95	0,96	1,00
RMSE	(3)	7,59	5,28	7,11	3,64	2,59	3,48	2,26	2,46	2,73
	(5)	1,54	1,82	1,63	1,74	1,57	1,73	1,64	1,65	1,77
Tényleges szórás		2,13	2,87	2,84	2,70	2,59	2,60	2,47	2,75	2,91

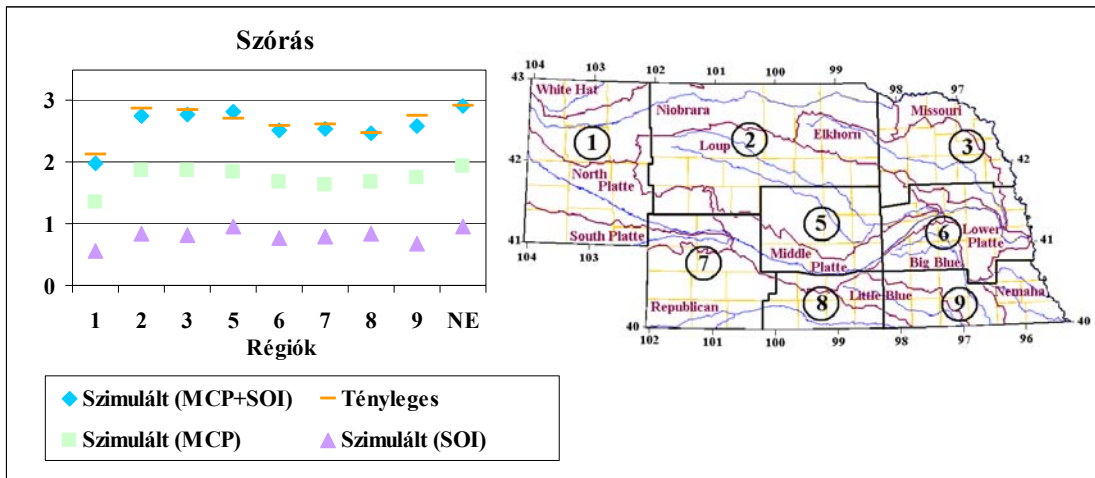
További vizsgálatokat végeztünk arra vonatkozóan, hogy a kétféle információ típus – MCP illetve ENSO – alkalmazásával hogyan változik a klímaoutputra kapott szimulált idősor a ténylegeshez viszonyítva. A **4-11. ábrából** kitűnik, hogy ha akár kizárólag az MCP havi relatív gyakoriságokat, akár csak az ENSO-t jellemző SOI idősorokat vesszük figyelembe, a modellek nem képesek megfelelően reprodukálni a tényleges PMDI idősort. Az ezekkel a modellekkel szimulált idősorok jóval szűkebb PMDI-tartományba esnek, mint az együttes

információval dolgozó modellekből kapott PMDI értékek. Megerősíti ezt a megállapítást a **4-12. ábra** is, ahol a szimulált és a tényleges PMDI idősorok szórásait vethetjük össze a különböző Nebraska-i kistérségek esetén.

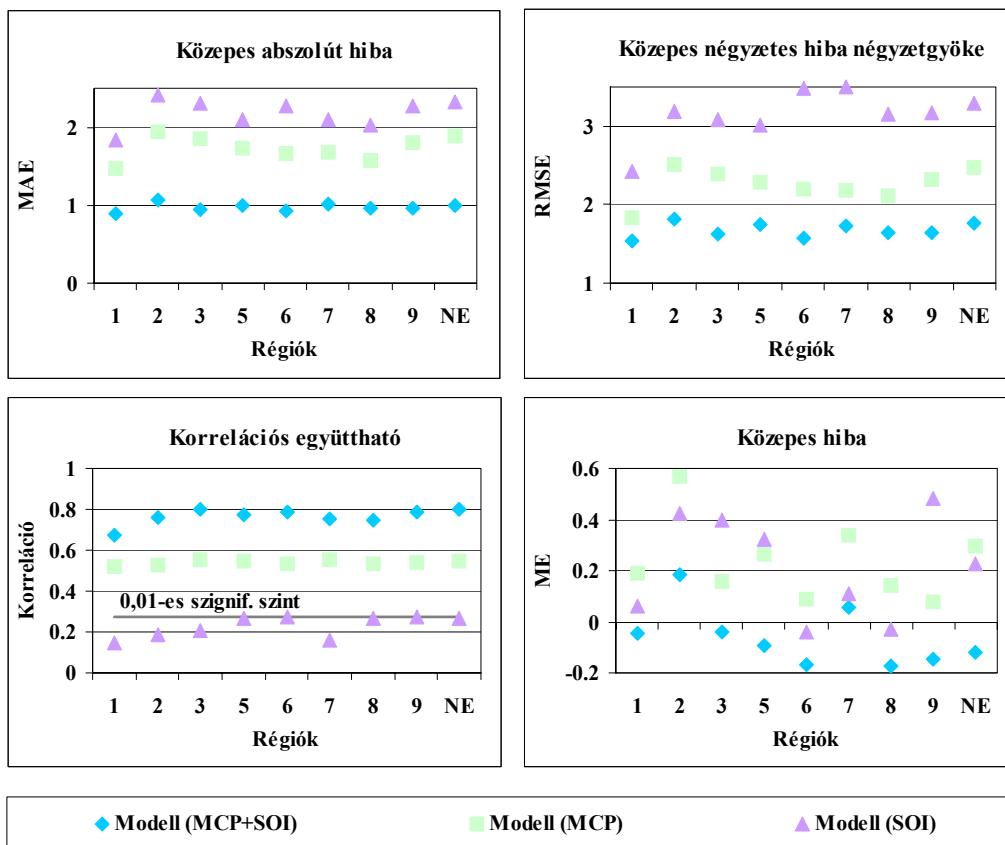


4-11. ábra: A PMDI szimulációk eloszlásainak összevetése a megfigyelt empirikus eloszlással különböző bemenő változókat felhasználó modellek esetén (Nebraska teljes területére, 1963-77)

A modellek ME, MAE, RMSE hibáit, valamint a szimulált és a tényleges PMDI idősorok közötti korrelációs együtthatókat a **4-13. ábrán** látható grafikonon hasonlítjuk össze. Az együttes nagytérségű klimatológiai információ jelentősen csökkenti a modellek hibáit (főként a MAE-t és az RMSE-t). A szimulált és a tényleges idősorok közötti statisztikai kapcsolat ebben az esetben eléri a 0,7-0,8-as erősséget, míg csupán az MCP-t figyelembe vevő modellszimulációkkal minden kistérségben mindössze 0,5-0,6 erősségű korrelációt érhetünk el, ennél is gyengébb (0,1-0,3-as) korrelációkat eredményeznek a kizárólag a SOI-t tekintő fuzzy-szabályokból felépülő modellek. Ez utóbbiak 0,01-es szinten legtöbbször nem szignifikánsak (a szignifikáns korrelációs együtthatók 0,27-nél nagyobbak). A Palmer-index tehetetlensége miatt erős az autokorreláció is (0,8-0,9 közé esnek az 1 hónapos elcsúsztatással előállított autokorrelációs együttható értékek), vagyis rövidebb távra közvetlen előrejelzésre nem érdemes alkalmazni a PMDI-re kifejlesztett modelleket. Elsősorban olyan alkalmazások esetén érhetünk el jelentős eredményt, amikor hosszabb időszakra szükséges regionális skálájú éghajlati információt generálnunk.

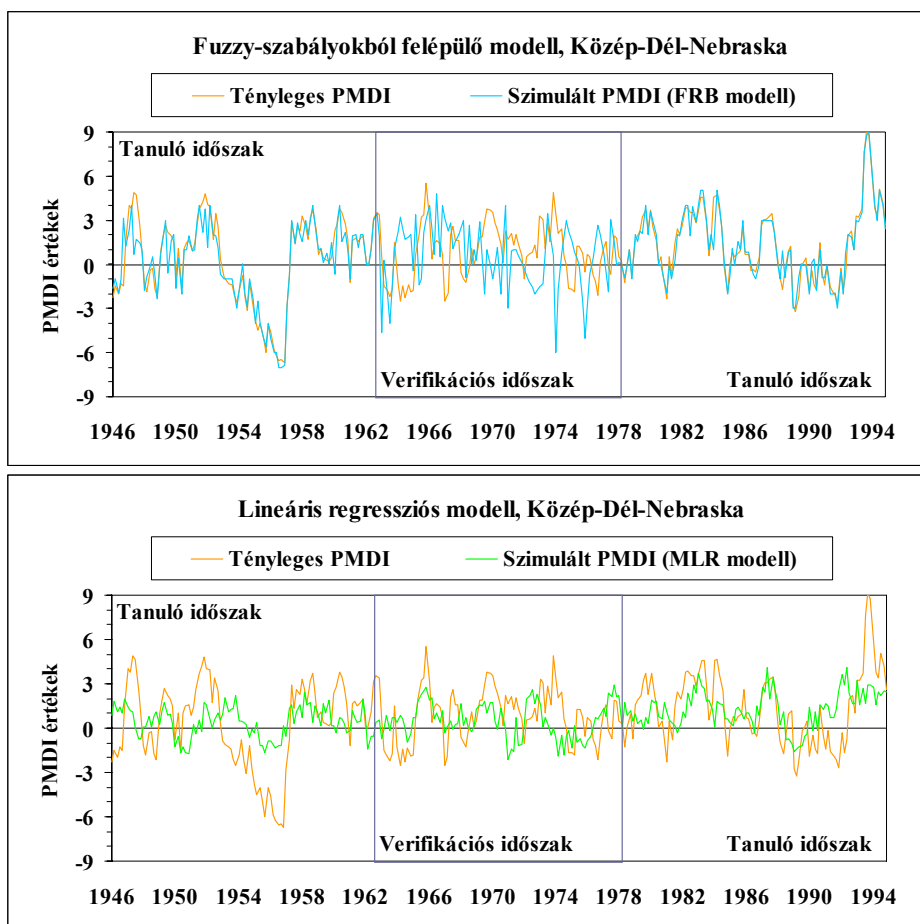


4-12. ábra: A szimulált PMDI idősorok szórásainak összevetése a tényleges PMDI idősor szórásával különböző bemenő változókat figyelembe vevő modellekre (1963-77)



4-13. ábra: A modellhibák és a korrelációk különböző bemenő változókat figyelembe vevő modellek esetén (1963-77)

Végezetül a fuzzy-szabályokból felépülő modellek eredményeit összehasonlítottuk a többváltozós lineáris regressziós modellek adta becslésekkel. Mindkét modelltípusnál azonos bemenő változókat (MCP relatív gyakoriságok és SOI együttesen) alkalmazva a **4-14. ábra** mutatja a teljes 1946-94-ig terjedő tényleges PMDI-idősor és a szimulált idősorok összevetését. Míg a lineáris regressziós modell a szélsőségesebb éghajlati viszonyokat (a +3-nál nedvesebb illetve -3-nál szárazabb PMDI értékeket) egyáltalán nem képes rekonstruálni – még a tanuló időszak alatt sem –, addig a fuzzy-szabályok alkalmazásával ezek a viszonyok is előállíthatók. A szórásokat tekintve is hasonló következtetésre juthatunk (a tényleges PMDI idősor 2,47-es szórással rendelkezik, ehhez képest a fuzzy-szabályokból felépülő modell 2,46-os szórású szimulált PMDI idősort hoz létre, lineáris regresszióval viszont csupán ennek kb. fele, 1,20-as szórásérték érhető el).



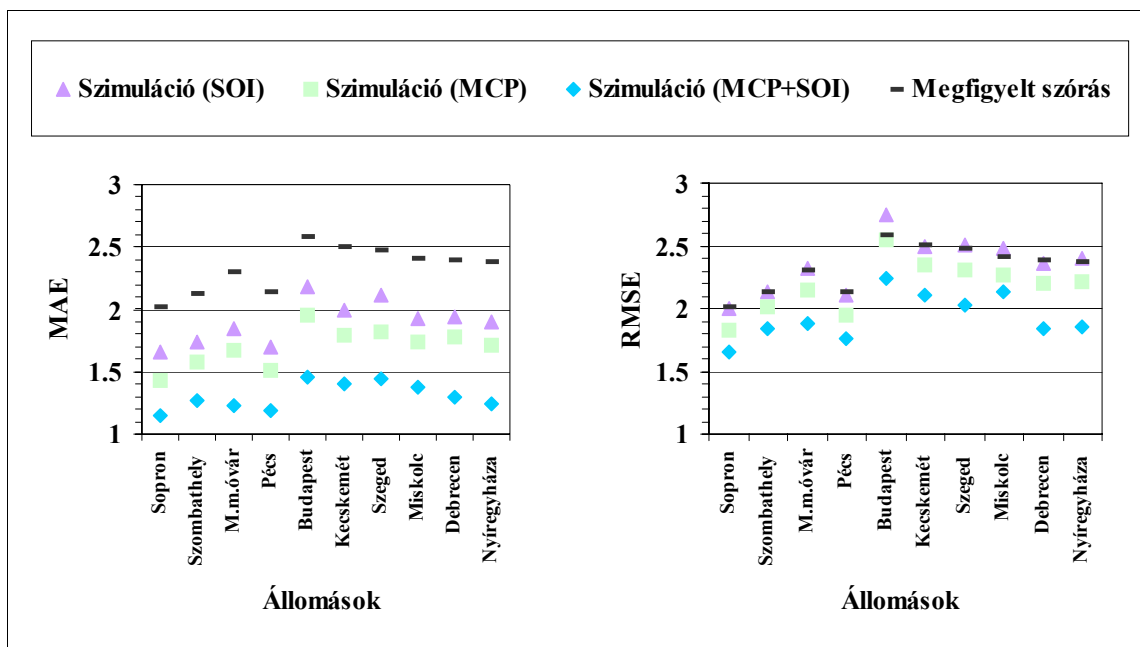
4-14. ábra: A lineáris regressziós (MLR) és a fuzzy-szabályokból felépülő (FRB) modell PMDI idősor becslései összevetve a tényleges PMDI idősorral (Közép-Dél-Nebraska, 1946-94)

A fentebb bemutatott eredmények arra utalnak, hogy a fuzzy-szabályokon alapuló modellek statisztikailag megfelelően reprodukálják Nebraska szárazsági viszonyait, amennyiben a nagytérségű makroszinoptikus helyzetet és az ENSO-ra vonatkozó információkat egyaránt felhasználjuk.

4.2.2. Magyarországi klímaállomásokra végzett vizsgálatok eredményei

A tíz magyar klímaállomásra meghatározott havi PDSI értékek vizsgálatánál (Pongrácz et al., 2000; 2002b) a bemenő változók egyrészt az ENSO-t és késleltetett távhatását jellemző havi bontású SOI idősorok, másrészt pedig a Hess-Brezowsky féle makrocirkulációs típusok 2.1. alfejezetben tárgyalt csoportosításából adódó MCP-osztályok havi relatív gyakoriságai voltak. Magyarországon a jelentősebb aszályveszély a nyári

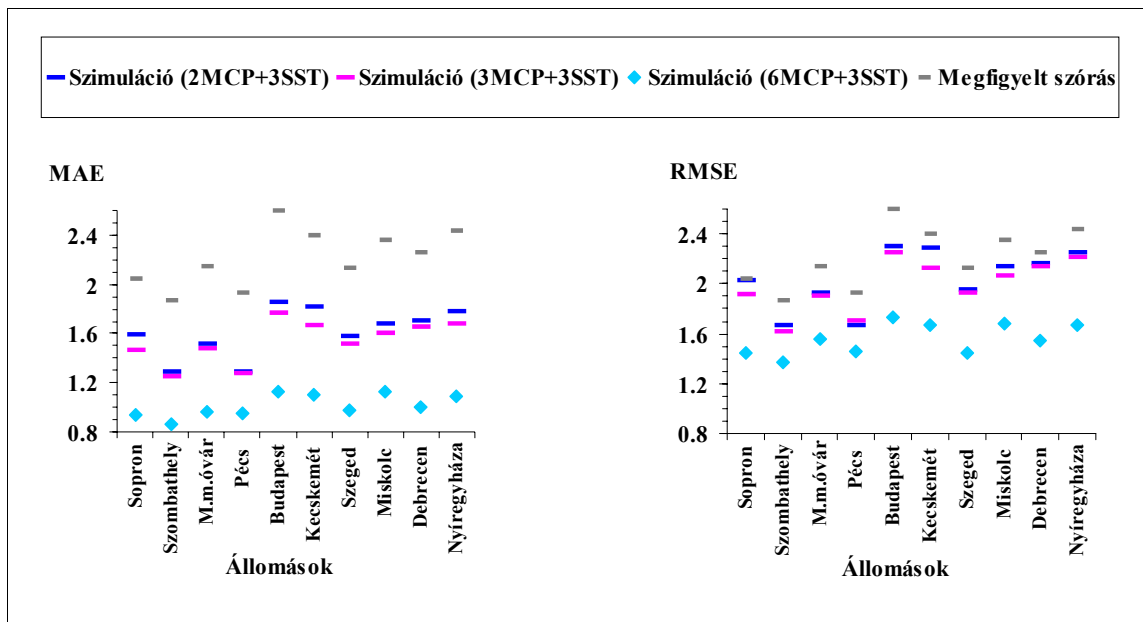
időszakhoz kapcsolódik, így elsősorban a tenyészidőszak hat hónapjára (április-szeptember) végeztük a modell-futtatásokat. A tanuló időszak az 1901-1970 közötti hetven évet tartalmazza, a verifikációs időszak pedig az 1971-1999 éveket öleli fel.



4-15. ábra: A modellhibák összehasonlítása különböző bemenő változókat figyelembe vevő modellek esetén (1971-99)

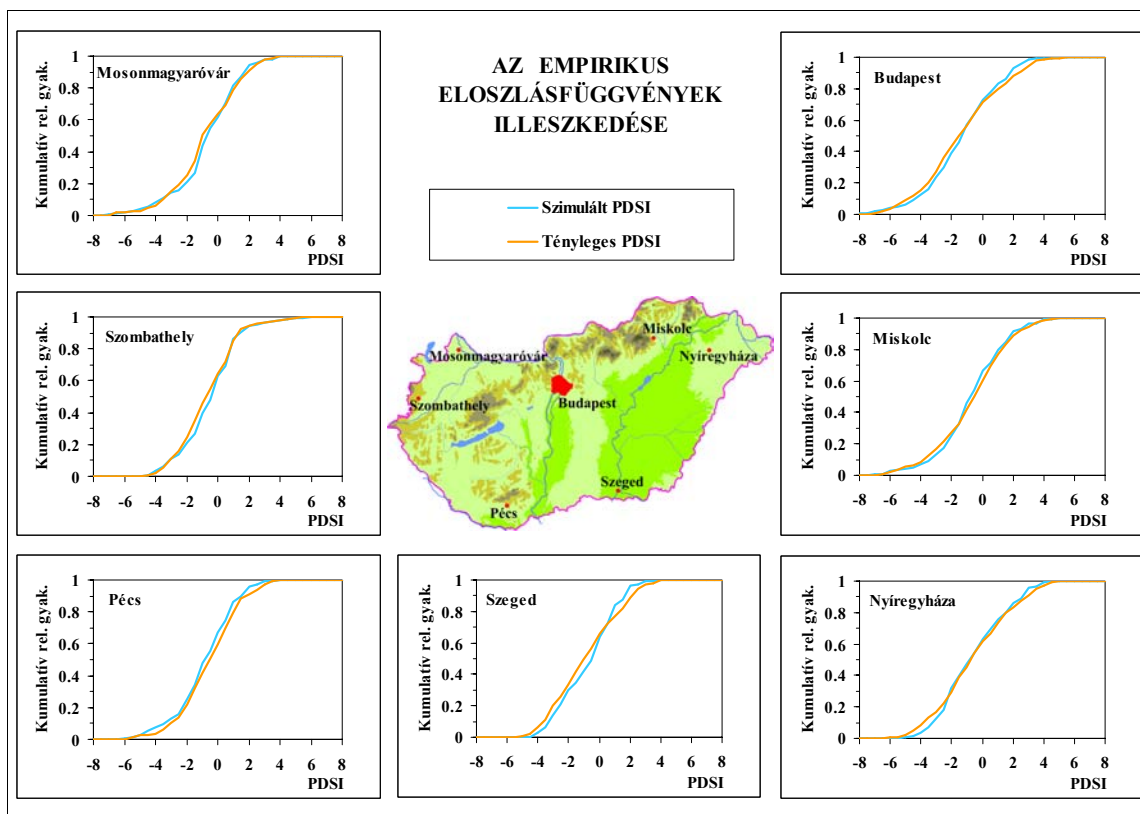
Elsőként a kétféle bemenő paraméter csoportot külön-külön vettük figyelembe a fuzzy-szabályokból álló modellek felépítésekor. A MAE és az RMSE modellhibákat a 4-15. ábrán hasonlítjuk össze. Mind a tíz állomás esetén jelentősen javul a szimulált PDSI értékek illeszkedése a tényleges PDSI idősorhoz viszonyítva, amennyiben a kétféle információt együttesen alkalmazzuk a modellezés során, s a szórásnál is jóval kisebbé válnak a modellhibák. Ekkor a szimulált és a tényleges idősorok közötti korrelációs kapcsolat is erősödik: 0,6-0,7-es (0,01-es szinten) szignifikáns korrelációs együtthatót kapunk, míg az információk elkülönült felhasználásával az ENSO illetve az MCP esetén rendre csupán 0,1-0,2 illetve 0,3-0,4 erősségű korrelációt tapasztalunk (a 0,01-es szinten szignifikáns korreláció 0,23-nál nagyobb együtthatóra áll fenn). A PDSI definíciójából következik, hogy az idősorok 1 hónapos elcsúsztatásával kapott autokorrelációs együttható is viszonylag nagy (0,7-0,8). Az időbeli elcsúsztatás növelésével csökken az autokorreláció erőssége: 2 hónapos elcsúsztatással 0,5-0,7 az autokorrelációs együttható értéke, 3 hónap esetén 0,3-0,6 közé esik, 4 hónapos elcsúsztatást végezve pedig már csupán 0,2-0,5.

A magyar klímaállomások esetén a távolabbi akcióközpontú ENSO helyett a NAO figyelembevételével is végeztünk modellezési vizsgálatokat. Ezekben a modellekben a fuzzy-szabályok megtanulására az 1950-70 és 1986-99 időszakok szolgálnak, míg a verifikációs időszakot az 1971-85 közötti tizenöt év tartalmazza. A NAO-t a 2.3. alfejezetben bemutatott SST-index-szel jellemeztük, melynek néhány hónapos késleltetett hatását is tekintetbe vettük a modell-futtatások során.



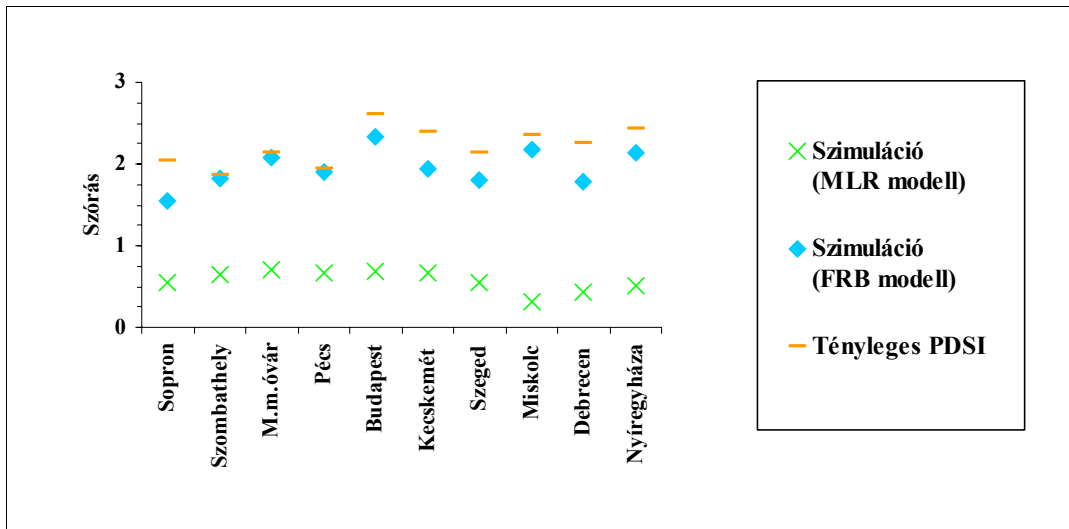
4-16. ábra: A modellhibák összehasonlítása különböző modellek esetén (1971-85)

Vizsgálataink során az MCP típusok többféle lehetséges csoportosítását (ld.: a 2.1. alfejezetben) alkalmaztuk a modellekben. Az így futtatott modellek MAE és RMSE hibáinak összevetése látható a 4-16. ábrán. Minden állomás esetén akkor kapjuk a legkisebb modellhibákat, ha az európai ciklonális/anticiklonális dominanciát és a cirkulációs jelleg zonalitását is figyelembe véve összesen 6 MCP osztály havi relatív gyakoriságait tekintjük bemenő paraméternek a havi SST-index idősorok mellett. A korrelációs együtthatók ekkor érik el a 0,7-0,8-as erősséget, mely 0,01-es szinten szignifikáns. A 4-17. ábra mutatja a szimulált és a tényleges PDSI értékek idősorához tartozó empirikus eloszlásfüggvények illeszkedését. A Kolmogorov-Szmirnov-féle kétmintás próba alapján az eloszlások 0,1-es szignifikancia szinten nem különböznek egymástól.



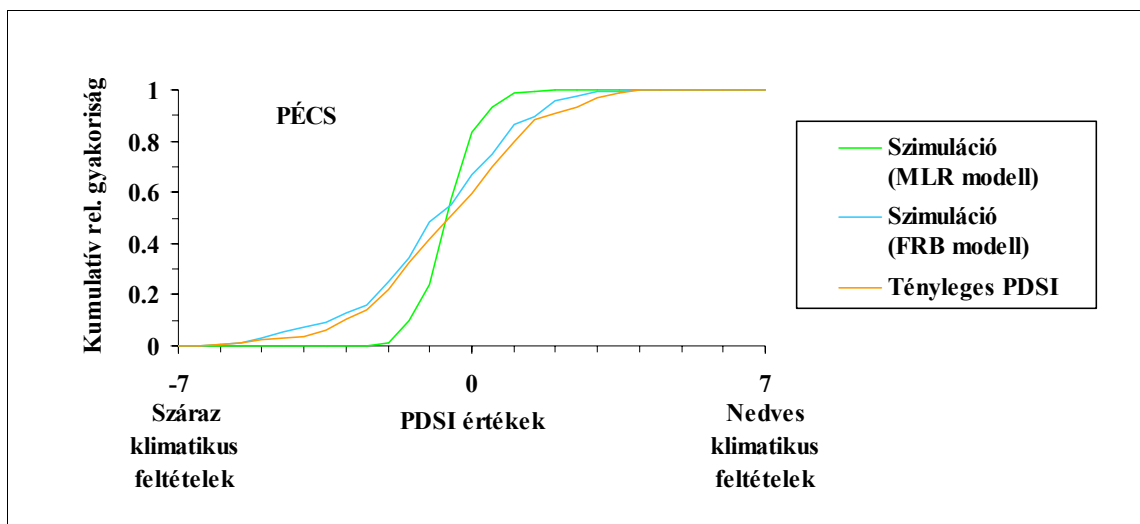
4-17. ábra: A szimulált és a tényleges PDSI idősorok eloszlásának összehasonlítása (1971-85)
 (A modell bemenő paramétereit: 6 MCP osztály havi relatív gyakoriságai és a NAO-t jellemző SST-index idősorok a késleltetett hatást is figyelembe véve)

Végül a kétféle alkalmazott módszert hasonlítjuk össze. A fuzzy-szabályokból felépülő modellekkel illetve a többváltozós lineáris regresszióval előállított PDSI idősorok szórásait vethetjük össze a tényleges PDSI idősorokéval a **4-18. ábrán**. Mind a tíz magyar állomásra szignifikánsan jobban közelíthetjük a tényleges adatsor statisztikai tulajdonságait a fuzzy-szabályokon alapuló modellezéssel. Míg a lineáris regressziós modellek felhasználásával adódó szimulált PDSI idősorok szórása csupán 0,3-0,7 között változik, addig a fuzzy-szabályokból álló modellekkel 1,5-2,3 szórású PDSI idősorokat állítottunk elő (a tényleges idősorok szórása 1,8-2,5 között változik a különböző állomásokra).



4-18. ábra: A szimulált és a tényleges PDSI idősorok szórásainak összehasonlítása a fuzzy-szabályokból álló (FRB) modellek és a lineáris regressziós (MLR) módszer esetén (1971-85)
 (A modellek bemenő paraméterei azonosak: 6 MCP osztály havi relatív gyakoriságai és a NAO-t jellemző SST-index idősorok a késleltetett hatást is figyelembe véve)

A fentieket erősíti meg a **4-19. ábra** is, ahol Pécs példáján az empirikus eloszlásfüggvények illeszkedéseit láthatjuk a különböző típusú modellekben azonos bemenő paramétereket alkalmazva. A többi állomás esetén is hasonló eredményre juthatunk: többváltozós lineáris regresszióval a szélsőséges éghajlati viszonyok modellezése egyik állomás esetén sem lehetséges, míg a fuzzy-szabályok képesek reprodukálni az extrém klimatikus feltételeket is.



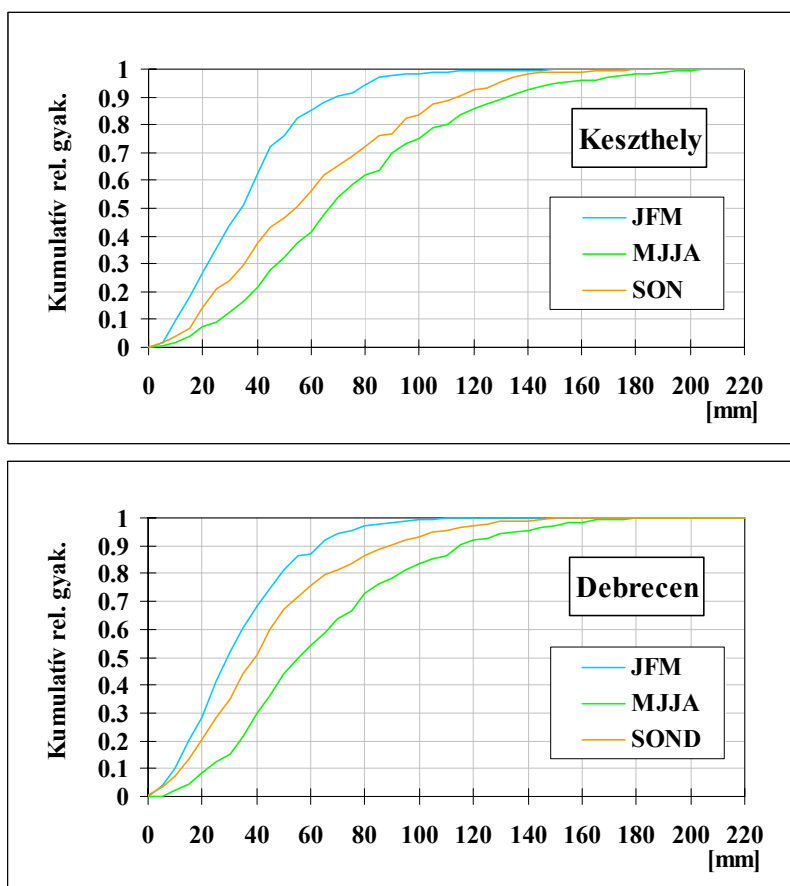
4-19. ábra: A fuzzy-szabályokból álló (FRB) modellel és a lineáris regresszióval (MLR) szimulált eloszlások összehasonlítása a tényleges PDSI idősor eloszlásával (Pécs, 1971-85).

4.3. Hazai havi csapadékösszegre készített becslés fuzzy-modellekkel

A havi csapadékmennyiségeket is megvizsgáltuk mindkét térségre, ebben az alfejezetben két magyar állomásra kapott eredményt mutatunk be, az egyik az Alföldet, a másik a Dunántúlt reprezentálja. Fuzzy-szabályokból felépített modelleket készítettünk s alkalmaztunk a nagy-térségű Európai/Atlanti cirkulációt reprezentáló havi MCP relatív gyakoriságok és az ENSO-ra illetve a NAO-ra vonatkozó klímaindexek havi idősorainak figyelembevételével. Mivel a csapadékösszeg eloszlása erősen aszimmetrikus, így a modellek éghajlati outputjára vonatkozó fuzzy-szám rendszerek definiálásakor a ferdeséget számításba vevő decilisek szerinti fuzzy-szám kijelölést alkalmaztunk (4-6. ábra).

A havi csapadékösszeg jellegzetes évi menete és az évszázados csapadékidősorok elemzése (Bartholy & Pongrácz, 1998c) indokolta a hagyományos évszakok helyett hónap-csoportok szerinti modell-futtatások elvégzését. Példaként Debrecen és Keszthely esetén három-három hónap-csoportot (éveleji, nyári, őszi csoport) definiálhatunk, melyek havi csapadékösszegeire meghatároztuk az empirikus eloszlásfüggvényeket. Az ezekben megjelenő szignifikáns különbségek a 4-20. ábrán láthatóak. Az egyes hónapok besorolásakor előfordult, hogy valamelyik hónap az előtte és az utána következőktől is jelentősen eltérő statisztikai tulajdonságokat mutatott (az átlagérték, a szórás, valamint az évszázad során mért értékek alapján meghatározott hisztogramok alapján), s emiatt egyik csoporthoz sem osztottuk

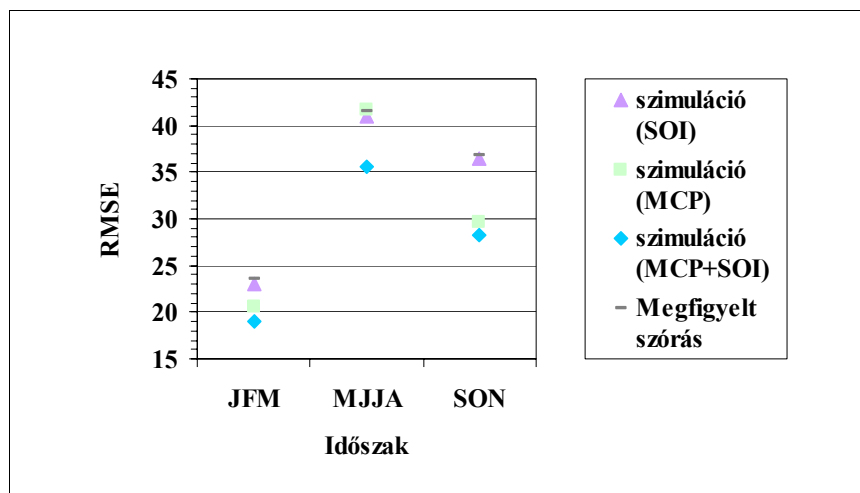
be. Így például Keszthely esetén áprilist és decembert, Debrecen esetén pedig csak az áprilist hagytuk ki a modellezési vizsgálatokból.



4-20. ábra: A hónap-csoportok havi csapadék eloszlásainak összehasonlítása Keszthely és Debrecen esetén (1901-99)

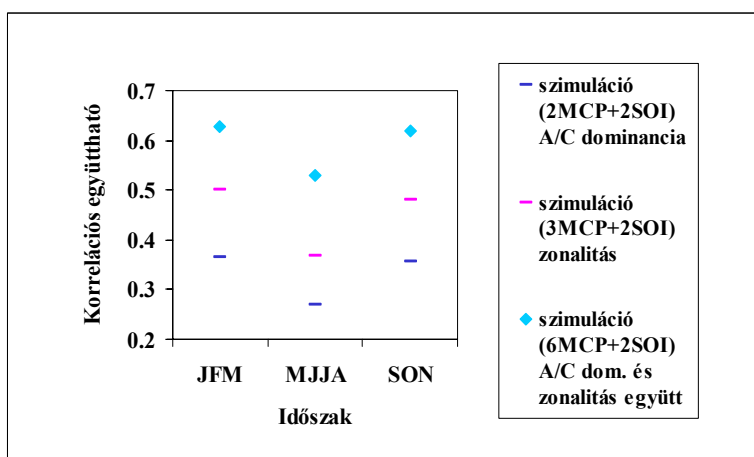
4.3.1. Magyarországi klímaállomásokra végzett modellezési eredmények

Elsőként az ENSO illetve a makrocirkulációs információ szerepét vizsgáltuk (Pongrácz et al., 2000; 2001; Bogardi et al., 2001). A fuzzy-szabályok megalkotására szolgáló tanuló időszakot 1901-63 között jelöltük ki, míg a verifikációs időszak az ezt követő harmincegy évet tartalmazza (1964-94 között).



4-21. ábra: Az RMSE összehasonlítása különböző modellek esetén (Keszthely, 1964-94).

A **4-21. ábrán** összehasonlítjuk az RMSE modellhibákat a bemenő paraméterek segítségével figyelembe vett információk függvényében. Mindhárom időszakban kisebb hibákat kaphatunk, ha a rendelkezésre álló információkat együttesen használjuk fel a fuzzy-szabályok előállítására. Ebben az esetben a megfigyelt szóráshoz viszonyítva is jelentősen csökken az RMSE értéke.

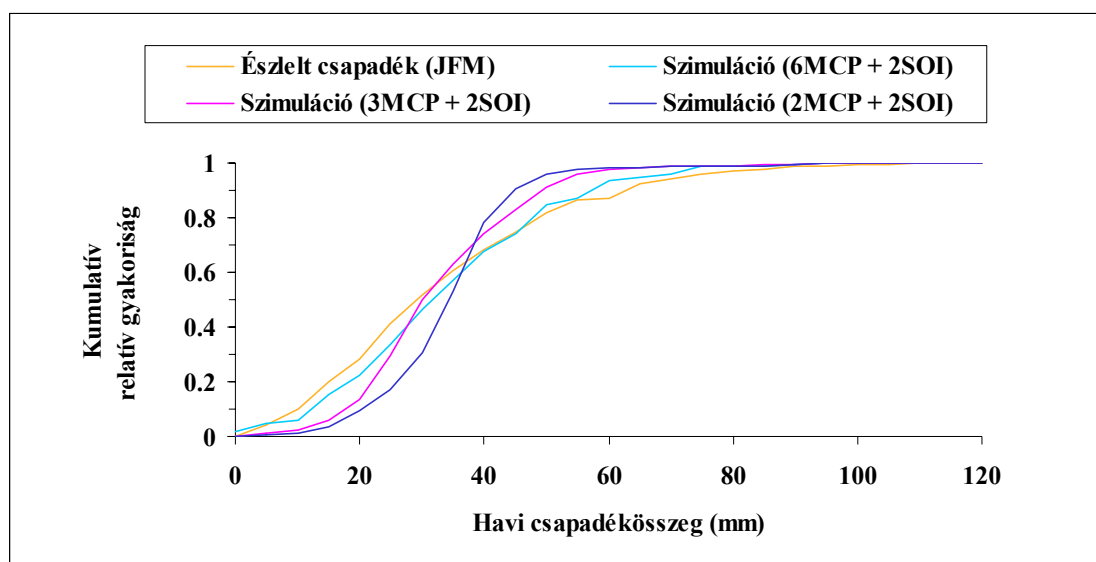


4-22. ábra: A szimulált és a tényleges csapadékösszegek (Keszthely) közötti korrelációs kapcsolatok összehasonlítása különböző modellek esetén (1964-94)

A makrocirkulációs helyzetek többféle lehetséges csoportosítását (ld.: a 2.1. alfejezetben) alkalmazva a modellekben a szimulált és a tényleges havi csapadékösszegek idősorai között a legerősebb korreláció akkor áll elő (**4-22. ábra**), ha a ciklon/anticiklon

dominanciát és a zonalitást is figyelembe véve 6 MCP osztály havi relatív gyakoriságait tekintjük bemenő paraméternek a SOI idősorok mellett. Meghatároztuk a 0,01-es szinten szignifikáns korrelációs együtthatókat, melynek értéke 0,27 a háromhónapos időszakok esetén, s 0,23 a négyhónapos időszakokban. A csapadékösszeg idősoraiban a Palmer-indexhez viszonyítva elhanyagolható a megmaradási hajlam, már az 1 hónapos időbeli csúsztatással számított autokorrelációs együttható sem szignifikáns, értéke 0,20-nál kisebb.

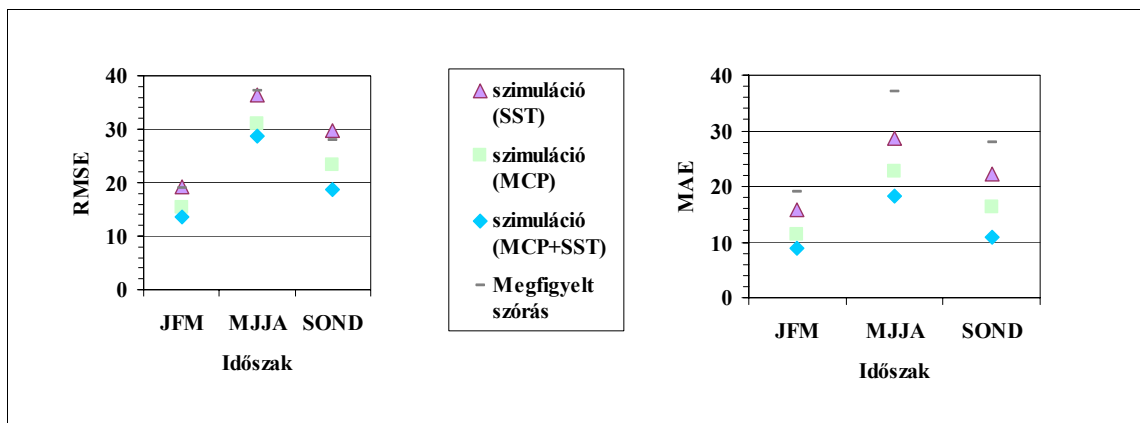
A modellel előállított és a ténylegesen regisztrált csapadék idősorok eloszlásait összevetve megerősíthetjük azt a következtetésünket, hogy a szimuláció sikeressége abban az esetben a legjobb, amennyiben a ciklonális illetve anticiklonális dominanciát és a zonalitást is figyelembe vesszük (4-23. ábra).



4-23. ábra: A szimulált és a tényleges csapadékösszegek (Debrecen) empirikus eloszlásfüggvényeinek összehasonlítása különböző modellek esetén (január-február-március hónapokra, 1964-94).

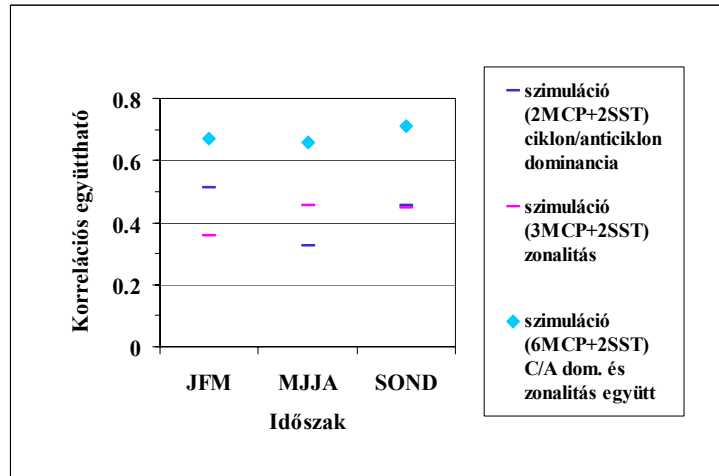
Az Európához földrajzilag jóval közelebb észlelhető NAO jelenség tekintetbe vételével végzett modellezési vizsgálatokban a fuzzy-szabályok megalkotására szolgáló tanuló időszakot 1950-84 között jelöltük ki, míg a verifikációs időszak az 1985-99 közötti másfél évtizedet tartalmazza. A NAO-t a 2.3 alfejezetben definiált SST-index-szel jellemeztük, s az ENSO-hoz hasonlóan néhány hónapos késleltetést is figyelembe vettünk a modell-futtatások során.

Vizsgálatainkban összevetettük a modellhibákat, melyek közül a MAE és az RMSE összehasonlítását mutatjuk be példaként Debrecenre a **4-24. ábrán**. Minden állomás esetén a legkisebb hibák akkor keletkeznek a havi csapadék szimulációjában, ha a kétféle információt (MCP havi relatív gyakoriságok, NAO-t jellemző SST-index féléven belüli két legerősebb korrelációjú havi értéke) együttesen használjuk fel bemenő változóként. Továbbá ekkor lesz a csapadékidősor megfigyelt szórásánál jelentősen kisebb a modellhiba.



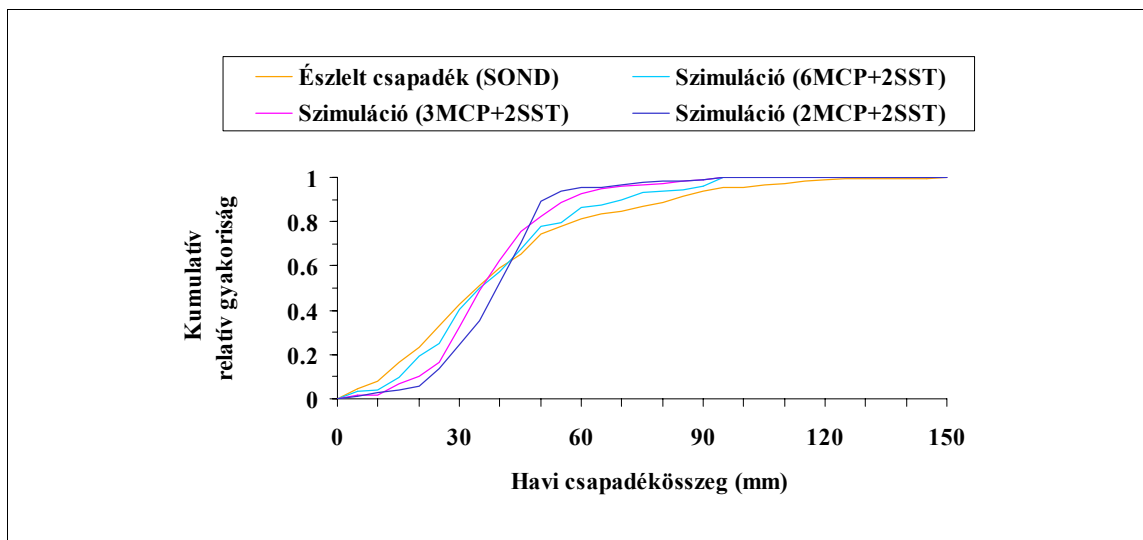
4-24. ábra: Az RMSE és a MAE összehasonlítása különböző modellek esetén (Debrecen, 1964-94).

A szimulált és a ténylegesen regisztrált havi csapadékösszegek közötti korrelációs kapcsolatokat vizsgálva az MCP ciklon/anticiklon dominanciája szerinti, vagy a cirkulációs jelleget (azaz zonális, meridionális, kevert MCP osztályokat elkülönítve) tekintő felosztásokat alkalmazó modellekre 0,3-0,5 közötti korrelációs együttható a jellemző, mely 0,01-es szinten szignifikáns. Ezzel szemben a mindkét szempontot figyelembe vevő modellek 0,6-0,8-as erősségű szignifikáns korrelációs kapcsolatot érhetnek el a szimulált és a mért havi csapadéksorok között. Debrecen példáját láthatjuk a **4-25. ábrán**, ahol a NAO hatással mindhárom modellben azonos módon számítottunk.



4-25. ábra: A szimulált és a tényleges csapadékösszegek (Debrecen) közötti korrelációs kapcsolatok összehasonlítása különböző modellek esetén (1964-94)

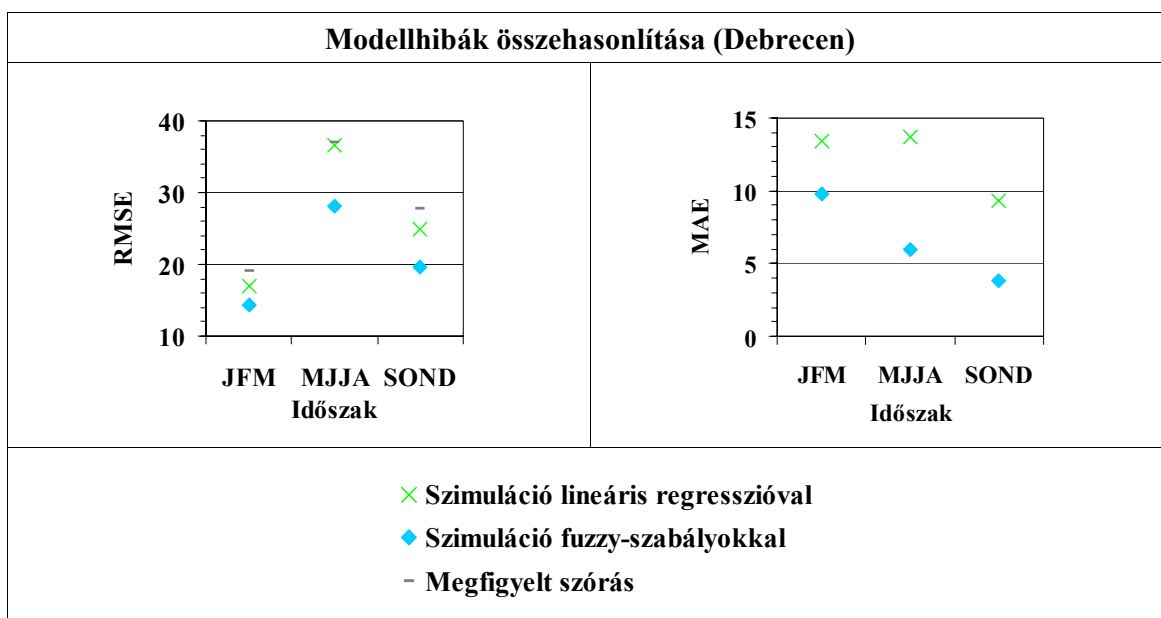
Az észlelt és szimulált csapadékidősorok eloszlásait értékelve ugyancsak arra juthatunk, hogy az empirikus eloszlásfüggvények szignifikánsan különböznek, ha az MCP típusokban csak a ciklon/anticiklon dominanciát vagy csak a cirkulációs jelleget vesszük figyelembe, ám a mindkét szempontot tartalmazó fuzzy-szabályokon alapuló modellek eredményei 0,05-ös szignifikancia szinten már nem tekinthetők eltérőnek. Példaként a **4-26. ábrán** a Debrecenre vonatkozó összehasonlítást mutatjuk be az őszi időszakra.



4-26. ábra: A szimulált és a tényleges csapadékösszegek (Debrecen) empirikus eloszlásfüggvényeinek összehasonlítása különböző modellek esetén (szeptember-október-november-december időszakra, 1964-94).

Végül vizsgálatainkban kitértünk a fuzzy-szabályokból felépített modellek és a többváltozós lineáris regresszió módszerének összehasonlítására. A havi csapadékösszegekre adódó szimulációs modellhibákat hasonlítjuk össze a **4-27. ábrán**. Mind a MAE-t, mind az RMSE-t tekintve a fuzzy-szabályok alkalmazásával kisebb hibákat kaptunk, mint lineáris regressziós modellekkel. A tényleges és a szimulált csapadékidősorok közötti korrelációs kapcsolat lineáris regresszióval általában 0,4-0,5 erősségű, míg a fuzzy-módszerrel 0,7-0,8-as korrelációs együtthatókat érhetünk el azonos bemenő változók felhasználásával.

Az ebben az alfejezetben bemutatott eredmények arra utalnak, hogy a fuzzy-szabályokon alapuló modellek statisztikailag jól reprodukálják a csapadékviszonyokat Magyarországon, amennyiben a nagytérségű makroszinoptikus helyzetet és a légköri oszcillációkra vonatkozó információkat egyaránt felhasználjuk.



4-27. ábra: Az RMSE és a MAE összehasonlítása különböző módszerek alkalmazása esetén (1964-94).

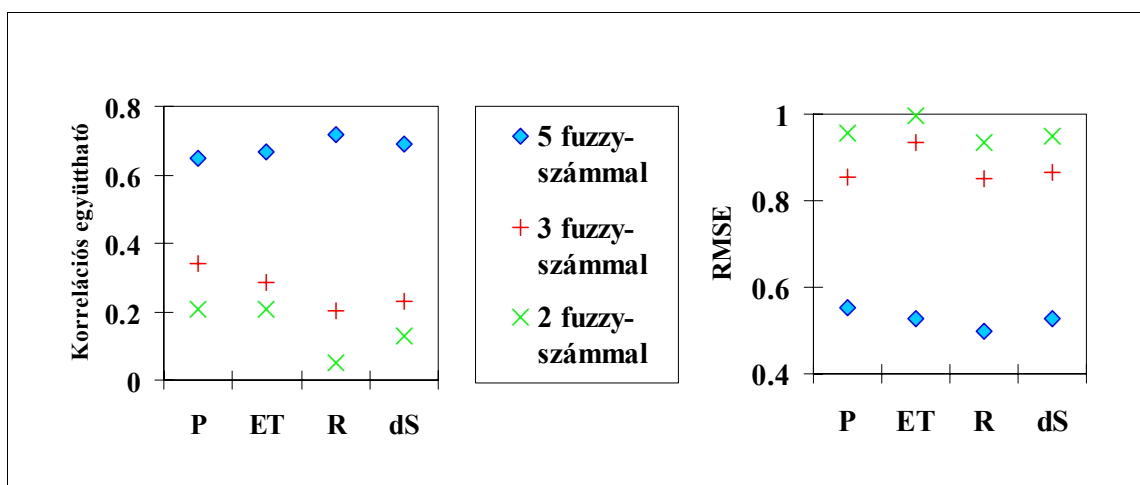
4.4. Balatoni hidrológiai elemekre készített becslés

Ebben a vizsgálatban az általunk alkotott modellek segítségével a Balaton területére vonatkozó vízháztartási mérleg négy elemére készítettünk szimulációs becsléseket (Pongrácz et al., 2002a; Bogardi et al., 2002), melyhez a csapadékösszeg, a területi párolgás, a felszíni

hozzáfolyás és a vízkészletváltozás havi felbontású idősorait használtuk fel. Az adatok a *Vituki Rt.* (1996) adatbázisából származnak és az 1921 és a 1995 közötti időszakra vonatkoznak. Mivel ezek a változók lényegesen eltérő éves menetűek és eloszlásúak (*Dévényi & Gulyás, 1988*), ezért indokoltnak tartottuk az eredeti értékek standardizálását. Mind a négy idősort átalakítottuk a kiszámított havi átlagok és a havi becült szórások segítségével, s így már nem szükséges az egyes hónapokat külön kezelnünk, továbbá a kapott eredmények is összevethetőek lettek egymással.

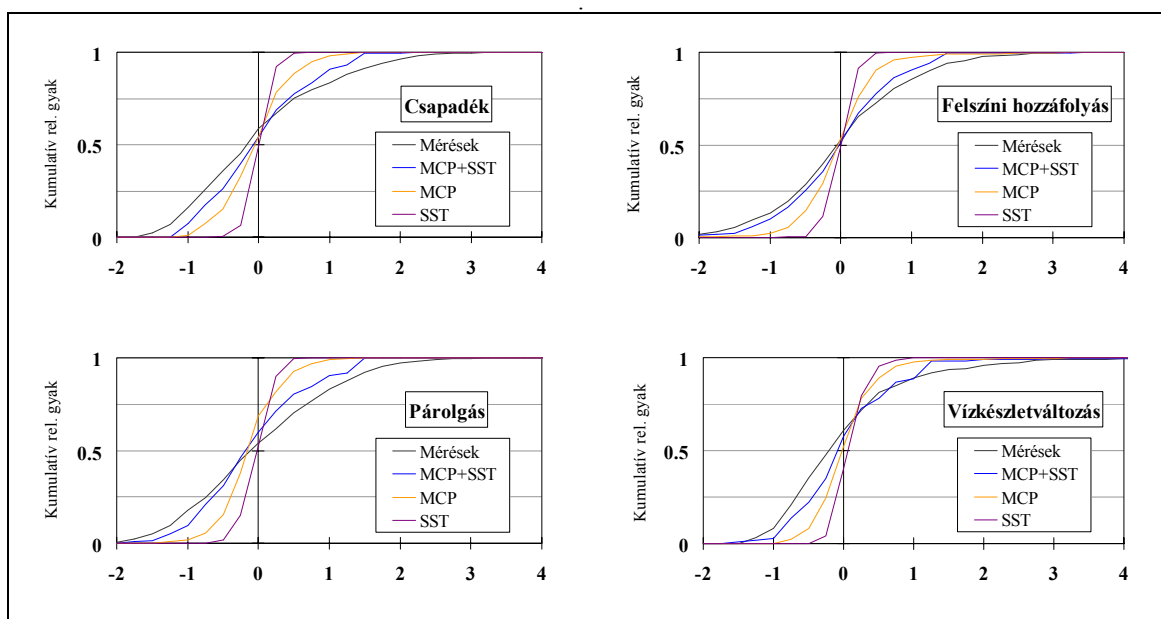
A fuzzy-szabályokból felépülő modellekhez kétféle információt használunk fel: a Hess-Brezowsky-féle tipizálás alapján kategorizált makrocirkulációs helyzeteket (s ezeknek a 2.1 alfejezetben ismertetett csoportosításait) és a NAO jelenséget, melyet az SST-ből származtatott indexértékekkel jellemeztünk. A tanuló időszak 1950-től 1980-ig tart, míg a verifikációs időszak 1981-től 1995-ig terjed.

Elsőként azt tekintjük át, hogy miként változik a modellbecslés eredményessége az egyes változókra definiált fuzzy-számok mennyiségének függvényében. Az input változókon 2 (*alacsony* és *magas* indexértékek), 3 (*alacsony, közepes, magas* indexértékek) illetve 5 (*nagyon alacsony, alacsony, közepes, magas, nagyon magas* indexértékek) fuzzy-számot határoztunk meg, az outputok esetében pedig 6-ot illetve 11-et. Az inputokon definiált fuzzy-számok hatását a **4-28. ábrán** látjuk.



4-28. ábra: A modellek összehasonlítása a korrelációs együttható és az RMSE-hiba segítségével (1981-95). Mindhárom modellben ugyanazok a meghatározó tényezők szerepelnek (3 SST-index havi értékei, 6 MCP-osztály havi relatív gyakorisága). Az input-változókon 2, 3, illetve 5 fuzzy-számot definiáltunk az egyes modellekben.

Ha 2 vagy 3 fuzzy-számot alkalmazunk, a korreláció igen gyenge lesz mind a négy vízháztartási elem esetében – a korrelációs együttható értéke 0,4-nél kisebb. Az RMSE viszont nagyon magas – 0,8-nál nagyobb. Eredményeink jelentős mértékben javulnak 5 fuzzy-szám használatával: a korrelációs együttható 0,6-nál nagyobb, az RMSE pedig a felére csökken. (A standardizálás következtében mind a négy változó szórása 1-től csupán kevéssé tér el, a 0,01-es szinten szignifikáns korrelációs együttható értéke 0,23.) A mért és a szimulált adatsorok közötti korreláció a felszíni hozzáfolyás esetében erősödik a legnagyobb mértékben, s az RMSE is erre a vízháztartási elemre adódik a legkisebbnek. Az output változókon a 6 és a 11 fuzzy-szám alkalmazása csupán egészen kis eltéréseket eredményez a modelleredményekben.



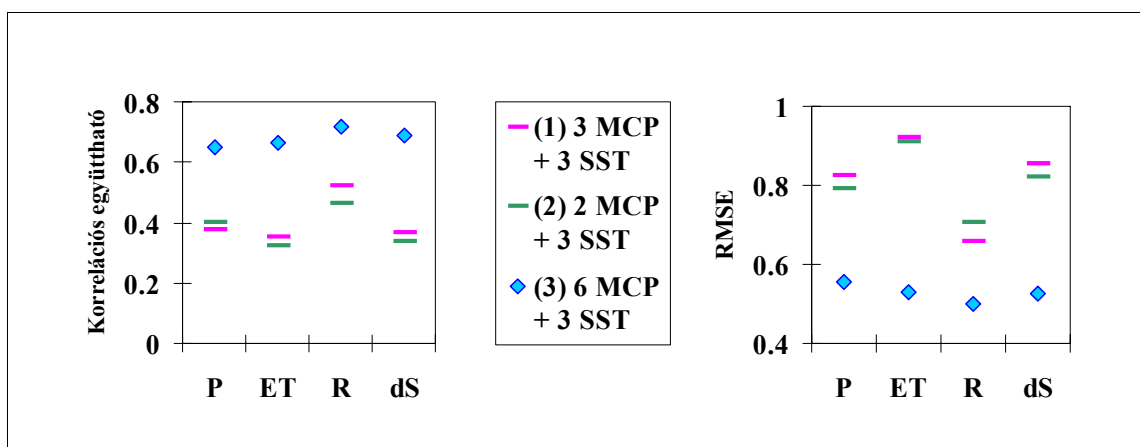
4-29. ábra: A mért és a szimulált standardizált adatsorok empirikus eloszlásainak összehasonlítása a Balaton vízháztartási mérlegének elemeire (1981-95).

Vizsgálataink következő lépésében a kétféle (NAO és nagy-térségű cirkuláció) információ felhasználását elemezzük. A **4-29. ábrán** az empirikus eloszlásfüggvények összehasonlításával azt mutatjuk be, hogy a csak NAO-hoz kapcsolódó, illetve a csak makrocirkulációs információt figyelembe vevő modellekkel készült szimulációk mennyivel javulnak, amennyiben a két tényezőt együtt tekintjük. Az eloszlások illeszkedését két-mintás Kolmogorov-Szmirnov-próbával ellenőriztük. A csapadék, a párolgás és a felszíni hozzáfolyás esetén a 6 MCP osztály havi relatív gyakoriságait és a késleltetett NAO-hatásokat

is figyelembe vevő modellek eredményeiből meghatározott empirikus eloszlások 0,01-es szignifikancia szinten a mérések eloszlásaival azonosnak tekinthetők. Ha azonban csak a 6 MCP osztály relatív gyakoriságait vagy csupán az SST alapú NAO-indexeket vesszük számításba, a modellekkel becsült és a mért adatsorok eloszlásai szignifikánsan eltérnek egymástól.

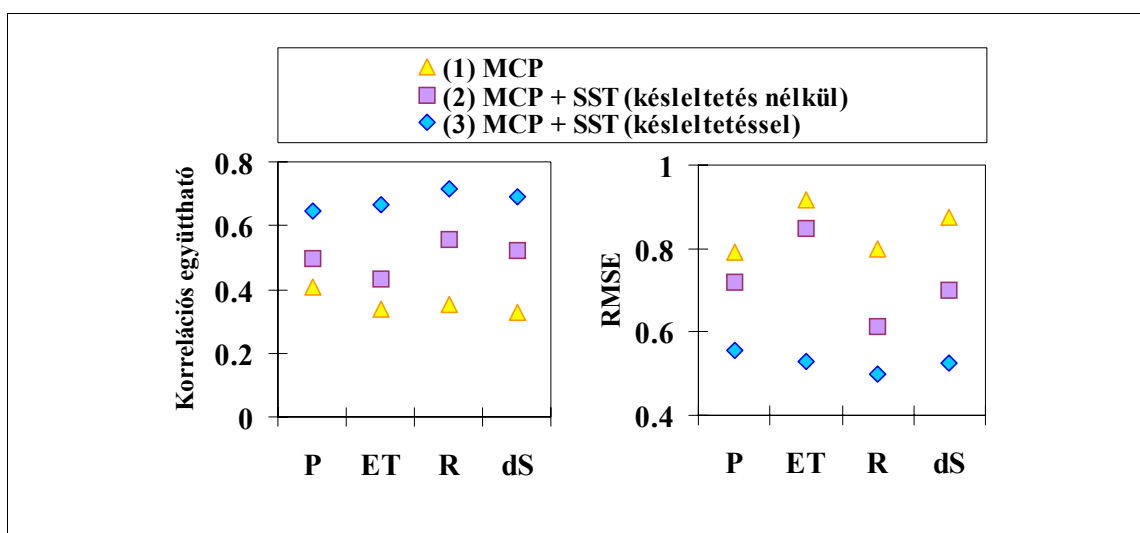
Az érzékenységi vizsgálatok további eredményei a **4-30.** és **4-31. ábrán** láthatók. Először csak a makrocirkulációs információra vonatkozóan változtattuk a modellek meghatározó tényezőit, s a NAO-t maximálisan figyelembe vettük. Majd a makrocirkulációt változatlanul hagyva az SST-idősorokból alkalmaztunk egyre többet.

A mérések és a szimulált adatsorok összevetéséből kapott korrelációs együtthatókat, valamint a modelltől származó RMSE hibaértékeket foglaljuk össze 2, 3 illetve 6 MCP osztály alkalmazása esetén a **4-30. ábrán**. (Ezekben a modellekben 5 fuzzy-számot definiáltunk minden input változón.) Amennyiben csak a ciklonális/anticiklonális dominanciát vesszük figyelembe, vagy csupán a cirkulációs jellegeket különítjük el, a korrelációs együttható általában nem éri el a 0,5-es értéket – míg a kétféle felbontás egyesítésével kapott MCP osztályok használatával a korreláció jelentősen erősödik, s értéke 0,6-0,8-re növekedik. Figyelembe véve a korrelációs együtthatóban és az RMSE-ben bekövetkezett változásokat, a legnagyobb javulást a párolgás esetében értük el.



4-30. ábra: A modellek összehasonlítása a korrelációs együttható és az RMSE-hiba segítségével (1981-95). Mindhárom modellben figyelembe vettük a késleltett NAO-hatást. A makrocirkuláció szerepe az egyes modellekben: (1) 3 MCP-osztály relatív gyakoriságát tekintve csak a cirkulációs jelleget vesszük figyelembe. (2) 2 MCP-osztály esetén csak a ciklon/anticiklon dominancia szerepel. (3) 6 MCP-osztály mindkét szempontot egyesíti.

A **4-31. ábra** arról nyújt tájékoztatást, hogy az SST-index használata nélküli illetve a csupán az egyidejű SST-index alkalmazásával működő modellekhez képest a két legerősebb kapcsolatú hónap késleltetett hatását is figyelembe vevő modellek milyen mértékben javítják a vízháztartási elemek szimulált értékeit. Ha egyáltalán nem tekintjük meghatározó tényezőnek a NAO-jelenséget, akkor a korrelációs együttható minden esetben 0,4 alatt marad. Amennyiben az SST-index egyidejű értékét figyelembe vesszük, a korreláció valamelyest erősödik, de 0,6 fölé csak akkor emelkedik, ha késleltető hatásokat is bevonunk modelljeinkbe. Hasonló mértékű javulások tapasztalhatók az RMSE modellhibák értékeiben is. Különösen a párolgás esetében javulnak nagymértékben a modelleredmények.



4-31. ábra: A modellek összehasonlítása a korrelációs együttható és az RMSE-hiba segítségével (1981-95). Mindhárom modellben 6 MCP-osztály gyakoriságával számoltunk. A modellekben (1) a NAO-t egyáltalán nem vettük figyelembe, (2) csak az egyidejű SST-indexet alkalmaztuk, (3) a késleltetett NAO-hatást is figyelembe vettük.