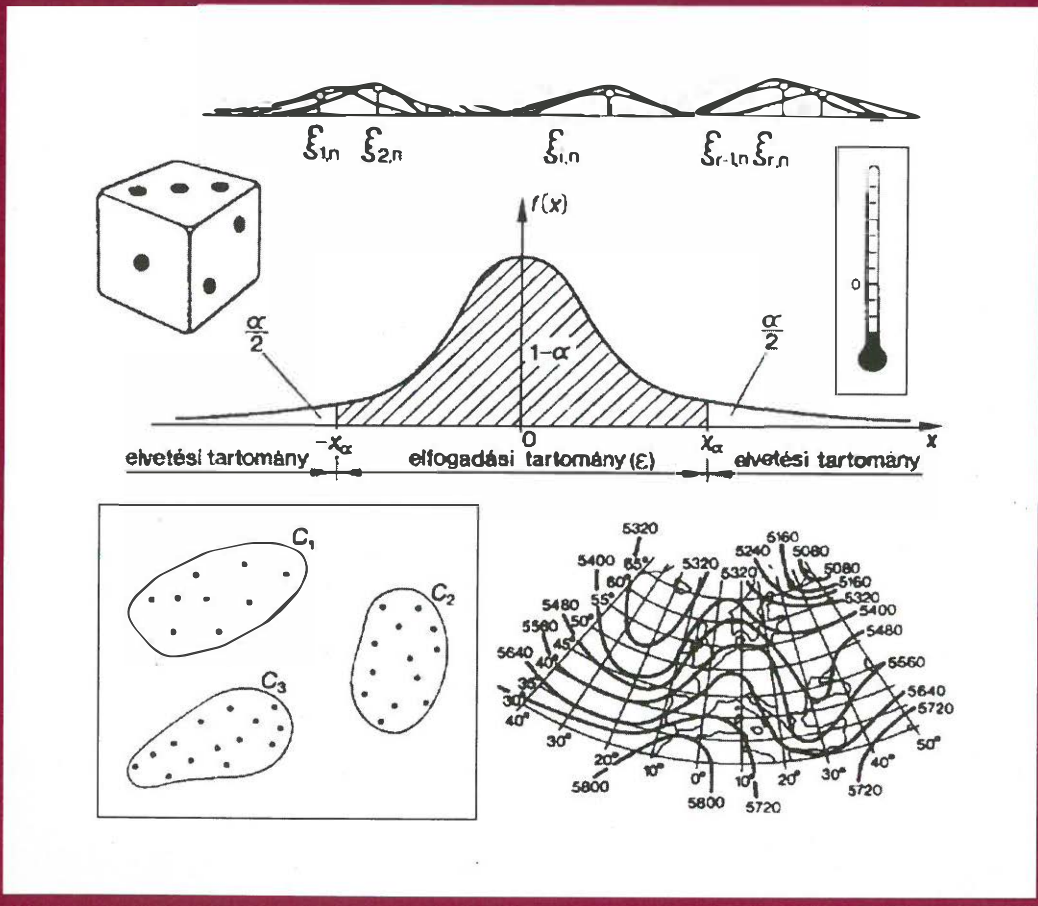
EMLÉK-KÖTET

# GULYÁS OTTÓ

HALÁLÁNAK 10. ÉVFORDULÓJÁRA



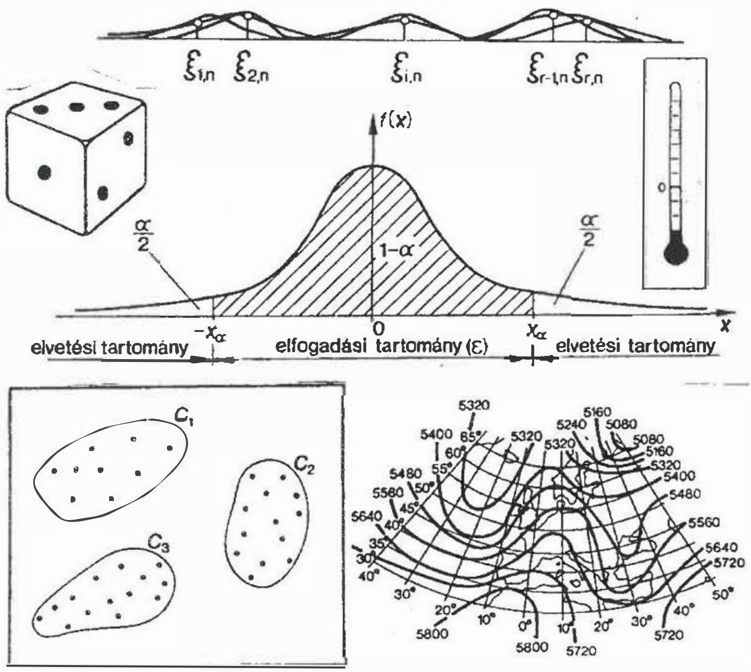

Országos Meteorológiai Szolgálat

Budapest, 1998

EMLÉK-KÖTET

# GULYÁS OTTÓ

HALÁLÁNAK 10. ÉVFORDULÓJÁRA




Országos Meteorológiai Szolgálat

A Gulyás Ottó Emlékülést rendezték:

**ORSZÁGOS METEOROLÓGIAI SZOLGÁLAT**

**EÖTVÖS LÓRÁND TUDOMÁNYEGYETEM**
**Meteorológiai Tanszéke**
**Valószínűségelméleti és Matematikai Statisztika Tanszéke**

**MAGYAR TUDOMÁNYOS AKADÉMIA**
**Meteorológiai Tudományos Bizottsága**

**MAGYAR METEOROLÓGIAI TÁRSASÁG**

*Szerkesztette:* Matyasovszky István
és Mika János

*Technikai szerkesztő:* Sáhó Ágnes

ISBN 963 7202 78 4

Kiadja az Országos Meteorológiai Szolgálat
A címlapon látható montázs a
Dévényi D. - Gulyás O. : Matematikai módszerek a meteorológiában c. könyv ábráiból készült
Kiadásért felel: dr. Mersich Iván az OMSZ elnöke
Sokszorosította a SZÍN&EL Kft.

# TARTALOMJEGYZÉK

*Előszó* 7

Götz Gusztáv:

*Megnyitó* 9

Czelnai Rudolf:

*Gulyás Ottó találkozása a meteorológiával* 11

Ambrózy Pál:

*Gulyás Ottó tevékenysége a Központi Meteorológiai Intézetben* 15

Szeidl László:

*Gulyás Ottó munkássága matematikus szemmel* 19

Szentimrey Tamás:

*Éghajlati idősorok statisztikai elemzése* 23

Bartholy Judit:

*Meteorológiai mezők objektív osztályozása* 33

Dévényi Dezső:

*Bayes becslés a meteorológiai adatasszimilációban* 41

*Gulyás Ottó publikációi*

(szerk. Szeidl László és Wantuchné Dobi Ildikó) 57

*Gulyás Ottó*

*1932 - 1988*

# ELŐSZÓ

Az éghajlat és az időjárás, egyszóval a légkörtudomány, a kutatás formai jegyei alapján is több száz évre tekint vissza. Ez idő alatt sokan alkottak maradandót akár csak hazánk szülöttei közül is. Bár törekszünk rá, mégsem tudunk nagy elődeink minden kerek évfordulójáról megemlékezni. Az idő múlásával Ők egyre inkább a könyvtárak polcain és az egyetemi tananyagokban élnek tovább, míg személyük lassanként feledésbe merül.

Csupán néhányan vannak, akik tevékenysége a munkatársak, tanítványok jóvoltából egyenértékű színvonalon folytatódik a mester kiválása után is. Ilyen kollégánk volt Gulyás Ottó matematikus, akit szívbemarkolóan korán ért utol az elmúlás. Köztünk létének a meteorológia nem kevesebbet köszönhet, mint a matematikai statisztika korrekt és igényes alkalmazásának, mint tudományos követelménynek az általánossá válását, s e széles problémakör sok fejezetében a hogyan kérdésének érthető és gyakorlatias megvilágítását.

Gulyás Ottó halálának tizedik évfordulóján 1998. szeptember 10-én a munkatársak és a tanítványok egy csoportja kezdeményezésére Emlékülést rendezett a két szakterület. Az időpont e szokatlan megválasztása - hiszen általában a születés dátumához kapcsoljuk a megemlékezést - szintén arra próbálta ráirányítani a figyelmet, hogy Gulyás Ottó munkássága jelen van napjaink légkörtudományi problémáinak és alkalmazásainak széles körében.

A jelen kötet két részre tagozódik: az előadások első felét a személyes megemlékezések és a szakmai életút egyes mozzanatainak felelevenítése képviseli. A második, inkább szakmai jellegű, részben három olyan terület mai állásáról kapunk áttekintést, amelyek módszertani, vagy tartalmi vonatkozásban Gulyás Ottót is foglalkoztatták. A kötetet Gulyás Ottó fellelhető publikációinak - félő, hogy nem hiánytalan - listája zárja.

A kötet szerkesztőiként köszönetünket fejezzük ki a rendezést és az írásos megjelenést elsősorban támogató Országos Meteorológiai Szolgálatnak, valamint az ELTE Meteorológiai Tanszékének, az ELTE Valószínűségelméleti és Matematikai Statisztika Tanszékének, az MTA Meteorológiai Tudományos Bizottságának és a Magyar Meteorológiai Társaságnak a rendezésben való közreműködésért.

Köszönjük a család tagjainak megjelenését és érdeklődését. Reméljük, hogy az Emlék-ülés az Ő számukra is érzékelhetővé tette közösségeink nagyrabecsülését Gulyás Ottó iránt, amint ez az előadóknak és a rendezőknek szándékában állt.

Budapest, 1998. december 10.

*Matyasovszky István* és *Mika János*
szerkesztők

# Megnyitó

Götz Gusztáv

Köszöntöm a Gulyás Ottó kollégánk halálának tizedik évfordulója alkalmából rendezett emlékülésen megjelenteket – és külön szeretettel üdvözlöm Ottó feleségét: Lilit, fiát: Gyurit és a kisebbik unokát: Balázst, akik voltak szívesek jelenlétükkel ülésünket megtisztelni.

Abban az időben, amikor Ottó megkezdte szakmai pályafutásának meteorológiai szakaszát, tudományágunkban két, egymással szöges ellentétben álló nézet uralkodott. Az egyik álláspont szerint a légkör eseményei szigorúan determináltak: a megértésük és előrejelzésük felé vezető, egyedül üdvözítő utat a légköri folyamatokat kormányzó hidro-termodinamikai egyenletekre épülő modellekkel történő numerikus kísérletek végzése jelenti. A másik tábor – nem utolsósorban Norbert Wienernek, a Massachusetts Institute of Technology neves matematikus professzorának egy 1956-ban elhangzott előadására alapozva – azt vallotta, hogy a légkör folyamatai alapvetően véletlen jellegűek, azokat sztochasztikus-statisztikus modellekkel kell közelíteni, és azoknak az alkalmasan megválasztott regressziós függvényeknek, amelyekben a keresett jövőbeli értékek a jelen és a múlt értékek lineáris kombinációiként állnak elő, éppoly jól kell működniük, mint bármilyen más elérhető eljárásnak, beleértve a numerikus prognosztikában és a szinoptikus gyakorlatban alkalmazott módszereket is. Ottó ennek az utóbbi nézetnek a meggyőződéses híve volt; kutatásait a valószínűségszámítás és a matematikai statisztika területén végezte. Természetesen voltak közöttünk szenvedélyes viták – csakúgy, mint a világon mindenütt a két tábor képviselői között – a megismerés követendő módszereit illetően, ezeket a vitákat azonban mindig igyekeztünk a kölcsönös meggyőzés udvarias keretei közé szorítani. Kiálltunk eltérő nézeteinkkel a nyilvánosság elé is: a *Véletlenszerűek-e a légköri folyamatok?* címmel 1987 januárjában rendezett kerekasztal beszélgetés csak egyike volt a lehetséges közelítési módozatokat felvázoló szemináriumoknak. Komoly vállalkozásunknak ígérkezett az 1988. évi akadémiai Meteorológiai Tudományos Napok megszervezése, amely a *Matematikai módszerek a meteorológiában* témakörben volt hivatott színvonalas előadásokkal áttekintést adni a hazánkban folyó dinamikai és statisztikai kutatásokról. Ottó nagy lelkesedéssel vágott bele az előkészítő munkába – és a sors kegyetlen játéka volt, hogy a rendezvényt már nem érhette meg.

Az elmúlt tíz esztendő gyökeres változásokat hozott. Egyre szilárdabb matematikai alapokra helyeződő új diszciplína született, a *véletlenszerű* viselkedést tanúsító *determinisztikus* rendszerek elmélete, közismertebb nevén: a káoszelmélet. Ennek az elméletnek a meteorológiai alkalmazása keretében a korábban említett két nézet már nem ellentétes egymással: a légkör folyamatairól a különböző (időjárási és éghajlati) időskálákon csak a dinamikai és a statisztikai módszerek *együttesével* rajzolható teljes kép.

Ha Gulyás Ottó körünkben bejárt életútját röviden kellene jellemeznem, azt mondanám: alkotott, oktatott és tanítványokat nevelt. Ezt a szűkszavú jellemzést hivatottak bővebben a hallgatóság elé tárni a mai emlékülés előadásai.

# Gulyás Ottó találkozása a meteorológiával

Czelnai Rudolf

A megtisztelő felkérésnek, mely szerint ezen az emlékülésen a nyitóelőadás feladata éppen nekem jutott, olymódon szeretnék eleget tenni, hogy elmondom, hogyan is történt Gulyás Ottó találkozása a meteorológiával. Ezt örömmel teszem meg, de kötelességemnek is érzem, mert úgy esett, hogy az itt elmondandó nagy találkozás egy kisebb személyes találkozással kezdődött, melyben először csak mi ketten vettünk részt Gulyás Ottóval.

Az általam elmondandó történetnek van egy rövid, meg egy hosszú változata. A rövid változat onnan indul, hogy a Távközlési Kutatóintézetben, Csibi Sándor főosztályán a hetvenes évek eleje táján statisztikai alakfelismeréssel és tanulófelismerő algoritmusokkal kapcsolatos szemináriumokat szerveztek, és ezekre engem is rendszeresen meghívtak. Ezeknek a szemináriumoknak a légköre, mely teljesen eltért attól, ahogyan a Meteorológiai Intézetben pl. a referátumköri ülések akkoriban folytak, nekem nagyon tetszett. Érdemi vitákat lehetett hallani, sokkal pergőbb volt az ütem és aktívabbak voltak a résztvevők. Többször megfordult a fejemben, hogy ezt a szellemet importálnunk kellene.

Az előadók között többször szerepelt Gulyás Ottó, aki az Információközlés és Feldolgozás Osztály tudományos osztályvezetői posztját töltötte be. Előadásai számomra különösen érdekesek voltak, mert olyan matematikai statisztikai módszerekről beszélt, melyekről Tokióban - pár évvel korábban - azt hallottam, hogy kezdenek kulcsfontosságúvá válni a meteorológia területén. Ottóval többször beszélgettünk is ezekről a kérdésekről. A közös érdeklődésen túlmenően egyre növekvő szimpátia is fűzött hozzá. Megnyerő, szeretetre méltó egyéniségét egyre jobban értékeltem. Később ugyanezt érezték iránta más kollégáim is, akik közül többen jelen vannak a mai összejövetelen.

Talán nem is ért egészen váratlanul, mikor egy alkalommal Ottó hirtelen azt kérdezte: mit szólnék ahhoz, ha át akarna jönni hozzánk ? Be kell vallanom, hogy erre azonnal pozitív választ adtam, mégpedig anélkül, hogy - *horribile dictu* - az "üzemi háromszöggel", vagy a személyzeti vezetővel konzultáltam volna.

Az utóbbi ki is fejezte nemtetszését s figyelmeztetett: TKI főnökei ezt rossz néven vehetik! Ottó azonban biztosított, hogy ő már elmondta a dolgot közvetlen főnökének, Csibi Sándornak. Csibi Sándor pedig - akivel közben én is beszéltem - bölcsen és megértéssel fogadta ezt a fejleményt, mert pontosan tudta, hogy milyen feladatok keltették fel Ottó érdeklődését a meteorológia iránt. Rögtön rá is tértünk tehát a kikérés konkrét teendőire. Ezzel a rövid történetnek azt a részét, melyről másnak nem nagyon volt tudomása, el is mondtam. Ami ezután következett, arról Ambrózy Pál kollégám beszél majd.

A hosszabb történet, melyet ígértem, lényegében az előbbinek a hátterét adja meg és arról szól, hogy mi is volt a matematikai statisztikai stúdiumok helyzete a hazai meteorológia köreiben a hetvenes évek elején. Ezt azért kell vázolnom, mert csak így tudom érzékeltetni, hogy mi motivált, amikor Gulyás Ottó ajánlatát olyan gyorsan és lelkesen elfogadtam.

Azzal kell kezdenem, hogy a hetvenes évekig a hazai meteorológiai folklór a matematikai statisztikai és valószínüségelméleti alapismeretek tekintetében elemi szintű volt. Ez tükröződött az egyetemi oktatásban is. A leckekönyvemben bárki megnézheti, hogy 1950-ben, amikor az ELTE keretén belül a meteorológusképzés megindult, az oktatott tárgyak közül teljesen hiányoztak a valószínűségszámítás és a matematikai statisztika alapjaival foglalkozó stúdiumok. Ez nagyon feltűnő ellentmondásban volt azzal, hogy egyébként erős matematikai alapokat kaptunk és az analízist, a differenciálgeometriát és az algebrát olyan országos hírű professzorok tanították nekünk, mint Császár Ákos, Hajós György és Fuchs László. A harmadik évben az OMI egyik tudományos osztályvezetője leadott nekünk egy nagyon rövid anyagot valószínűségszámítás címszó alatt az éghajlati adatsorok homogenizálásáról és más effélekről. De ez semmit sem változtatott azon, hogy a matematikai statisztikai alapok módszeres felépítése hiányzott a tananyagból és méginkább hiányzott az a szemlélet melyet ezzel együtt tanárainknak a fejünkbe kellett volna csöpögtetniük jó időben, amikor még viszonylag fogékony állapotban voltunk.

Azt gondolom, hogy már az is eleve szemléleti probléma volt, hogy a matematikai statisztikát a tananyag összeállításakor mellőzték. Nyilvánvaló, hogy ezt a tárgyat nem tartották fontosnak. Talán az ötvenes évek "technológiai optimizmusa" okozta, hogy a nagyvilágtól elzárt, háborúban lerombolt és szegény Magyarország meteorológus köreiben - a távolból jött szakmai híreket félreértve- sokan hittek olyasmikben, hogy a meteorológia pár évtizeden belül "egzakt tudománnyá" fejlődik, amin olyasmiket értettek, hogy például a légköri dinamika törvényeinek pontos ismeretében s az új technika csodálatos eszközeinek birtokában nem lesz majd akadálya a 100%-os prognózisok készítésének. E szemléletben a véletleneknek, bizonytalanságnak, hibabecsléseknek és más efféle gyenge szívűségnek nem tulajdonítottak fontosságot.

Ezzel kapcsolatban fel szeretnék idézni egy érdekes élményt melyből kiderül, hogy nem egyedül a hazai meteorológus körökben uralkodtak ilyen illúziók, hanem a szomszédaink is hasonló cipőkben jártak. Az történt, hogy a hetvenes évek elején Belgrádban tartották a CBS alakuló ülését. Az ülés résztvevőit egy napon elvitték szakmai látogatásra az ottani új meteorológiai számítóközpontba. Ott egy lelkes, viszonylag fiatal kutató bemutatta saját előrejelző modelljét, s mondókáját a következő meglepő kijelentéssel fejezte be "Sikerült tehát olyan modellt kifejlesztenem, mely annyira stabil, hogy akár több hónapos előrejelzést is lehet vele készíteni".

Hüledezve néztünk egymásra, s akkor a mellettem álló cseh kolléga odasúgta: "Mindenhol vannak ilyen fickók, akik nincsenek tisztában azzal, hogy mi a különbség a modell numerikus stabilitása (vagyis azon tulajdonsága, hogy nem száll el) és a beválása között"... "Ez azért van " - tette hozzá - "mert nem tanultak matematikai statisztikát és nem állnak két lábon a szakmában".

De az itt vázolt torz felfogás nemcsak az elméleti meteorológia területén uralkodott, hanem valami módon kihatott az alkalmazott meteorológia területeire is. Pl. az egyetemen az agrometeorológia tanításakor senki sem beszélt arról, hogy azok a mezőgazdász kutatók, akikkel majd esetleg együtt kell dolgoznunk, a saját kísérleteikben bonyolult matematikai statisztikai teszteket alkalmaznak.

Ez persze bizarr helyzeteket eredményezett, amire elmondok egy példát. 1955 elején a martonvásári obszervatóriumba helyeztek. Alig voltam ott pár napja, amikor a nekünk otthont adó Mezőgazdasági Kutatóintézet igazgatója azzal a kéréssel fordult a meteorológiai obszervatórium stábjához, hogy szervezzünk egy matematikai statisztikai tanfolyamot az intézet mezőgazdasági kutatói számára. "Ti matematikusok vagytok, vagy mifélék" - mondta Rajki Sándor - "nektek ez nem okozhat gondot". Ezzel jól beletalált, mert rövid tájékozódás után már tudtuk: intézetének kutatói között többen is voltak, akik a matematikai statisztika módszereit nagyon magas szinten ismerték és művelték. Nekünk pedig ezekről a módszerekről halvány fogalmunk se volt.

Nem kellett nagy fantázia annak elképzeléséhez, hogy mekkora tekintélyvesztés lenne abból az egész meteorológiai szakma számára, ha kiderülne a valóság. Ámde főnöki döntés folytán mégis el kellett vállalnom a tanfolyam megtartását. Titkon kölcsön vettem az ottani könyvtárból néhány statisztikai könyvet, s mikor ezeket átnéztem, akkor láttam csak igazi rémülettel, hogy mibe keveredtem. Hónapokig jóformán aludni sem volt időm. Más kollégáimnak is voltak hasonló élményei. Mégis évek múltak el anélkül, hogy a kérdést rendeztük volna. Csak elszigetelt egyéni kezdeményezések történtek. Egyik-másik kollégám nagy erőfeszítéseket tett, hogy matematikai statisztikai felkészültségének hiányait pótolja. De az itt vázolt problémát ezek a hőstettek nem oldották meg.

Azt, hogy az akkori hazai meteorológus szakmai látókör a matematikai statisztika szögének fedetlenül hagyása folytán mennyire szűk volt, akkor vált megdöbbentően világossá számomra, amikor 1964-ben, ENSZ ösztöndíjasként, Leningrádban és Tokióban jártam. Mindkét helyen azt tapasztaltam, hogy a matematikai statisztikát valószínűségszámítást és pl. a sztochasztikus folyamatok elméletét centrális fontosságúnak tartják. A leningrádi GGO-ban történetesen Gandin és Kagan mellett dolgoztam, akik a legtermészetesebb hangon mondtak pl. olyanokat, hogy a meteorológiai előrejelzés kulcskérdése az adat-asszimiláció és azon belül a hibákkal terhelt hiányos kiindulási adatok objektív analízise. Sőt, már akkor beszéltek arról, hogy az aszinkron adatok felhasználásához ki kell dolgozni az ún. "négy dimenziós" adat-asszimiláció módszereit.

Tokióban Eichi Suzuki már 1964-ben (!) "ensemble prognózisokról" beszélt, igaz, hogy a kiindulási adatok perturbálását Monte Carlo módszerrel képzelte el, ami később nem bizonyult célravezetőnek. Japán kutatóktól hallottam először alakfelismerő algoritmusokról. Azt remélték, hogy a műholdas felhőanalízisek automatizálását ezen az úton sikerül megoldani. Ugyancsak foglalkoztatta őket az automatikus adatkontrol problémája. Náluk tudtam meg, hogy a turbulencia elmélet és a sztochasztikus folyamatok elmélete szorosan összefügg.

Mindezek a témák akkoriban majdnem teljesen kívül estek a hazai szakmai közösség látókörén. A probléma tehát nemcsak az volt ( és elsősorban nem az volt), hogy hiányoztak bizonyos módszertani ismeretek, melyek bármiféle hipotézis tudományos vizsgálatához elengedhetetlenül szükségesek, hanem az, hogy szóba sem jöttek azok a problémák, melyek akkoriban a meteorológus szakma aktuális feladataihoz kapcsolódtak. Nem tudom, hogy kell-e mondanom, hogy ez egy nagyon súlyos megállapítás, ami mindennél jobban érzékelteti, hogy Gulyás Ottóra mennyire szükségünk volt. Amikor átjött hozzánk, attól kezdve ezt az űrt sikerült betölteni és számos fiatal szakemberünk kapott tőle olyan indítást, mellyel a világ vezető kutatóhelyein is megbecsülést tudtak szerezni maguknak, s az egész magyar meteorológus közösségnek. Sikereik gyökerei sok esetben Gulyás Ottóhoz vezetnek.

Mi őbenne nem csalódtunk, de azt hiszem, hogy ő egy kicsit csalódott bennünk. Mert eredetileg az volt az álma, azért jött át hozzánk, hogy statisztikai alakfelismerés meteorológiai alkalmazásai területén végezzen kutatómunkát. Ezt az álmát nem nagyon tudta megvalósítani, mert mi nem hagytuk. Elhalmoztuk mindenféle más feladattal. De mégsem vesztegette az idejét hiába, mert teremtett egy iskolát az OMSZ keretein belül. Azt a szellemet, amit magával hozott számos követője viszi tovább és aligha lehet elképzelni ennél nagyobb dolgot.

Vagy talán mégis lehet! Arra gondolok, hogy Gulyás Ottó nemcsak magas fokú, alapos és megbízható szakmai felkészültségével, de emberi magatartásával is maradandó emlékű példát adott mindnyájunknak.

# Gulyás Ottó tevékenysége a Központi Meteorológiai Intézetben

Ambrózy Pál

Az előbb elhangzott előadásból megtudhattuk, hogyan került Gulyás Ottó kapcsolatba a meteorológiával, és melyek voltak azok a területek, ahol először munkálkodott meteorológusokkal együtt. A jelen előadás célja Gulyás Ottó azon tevékenységének ismertetése, amely már a Központi Meteorológiai Intézetben (KMI) zajlott le, a 70-es évek közepétől haláláig.

Szervezetileg ez úgy alakult, hogy 1973 végén létrejött a KMI-n belül a Módszertani Csoport, melynek 1974-ben Gulyás Ottó lett a vezetője. Munkatársai főleg fiatal matematikusok voltak. Rajtuk kívül dolgozott a csoportban 1-2 meteorológus is, de ők inkább csak szervezetileg tartoztak ide, nem pedig a kutatási témájuk miatt.

Az első nagy feladat, melyben rajta kívül matematikusként Faragó Tibor, meteorológusként Kaba Magdolna dolgozott, egy analógia-keresésre épülő távelőrejelzési módszer kidolgozása volt. Ennek lényege így foglalható össze: Valamilyen rendelkezésre álló időjárási típusrendszer - ilyen a hazánkban kifejlesztett és gyakorta alkalmazott Péczely-féle típusok rendszere, továbbá a német Hess és Brezowsky-féle tipizálás - szerint megállapítjuk a legutolsó 30 nap időjárási kódjait, majd az ezen időszak jellemző napi hőmérsékleti és csapadék adatait. Az így kapott sorokat rendre összehasonlítjuk egy hosszabb történeti idősor (esetünkben 60 év) azonos időszakainak (hónapjainak) megfelelő soraival alkalmasan választott analógia-index segítségével. Következő lépésként a leghasonlóbb 30 napos intervallumokat követő azonos hosszúságú (30 napos) időszak makroszinoptikus helyzeteinek sorozatait, ill. a statisztikailag hozzájuk rendelhető hőmérsékleti, csapadék, stb. sorokat határozzuk meg, s ez elvezet a 30 napos előrejelzés elkészítéséhez.

E módszerrel kapcsolatban számos matematikai vizsgálat folyt le Gulyás Ottó irányításával vagy részvételével, többek között a "legközelebbi társ szerinti döntés" alkalmazására, a regressziós típusú extrapolációra, továbbá a valószínűségi sűrűségfüggvény becslésére és illesztésére vonatkozóan.

Ebben az időben, tehát a 70-es évek második felében került a megvalósíthatóság közelébe az OMSZ azon terve, hogy az OMFB segítségével akkori viszonyok között nagy kapacitású számítógép birtokába jusson. A gazdasági tervezésekkel párhuzamosan intenzíven megindult a matematikai-statisztikai programrendszerek kidolgozása. Ebben Gulyás szintén fontos szerepet játszott, az elkészült tanulmányok, programok több száz oldalas jelentésekben kerültek napvilágra.

Közben tovább folytak az analógia-kereső eljárások terén elkezdett kutatások. Itt mindig nagy problémát jelentettek az analógiakeresés alapját képező időjárási helyzetek bonyolultsága, ezek egyetlen számmal (kóddal) való jellemzése, miközben ügyelni kellett arra, hogy az azonos csoportba kerülő esetek határozottan elkülönüljenek más csoportok jellemző tulajdonságaitól. A csoportosításnak (clusterezésnek) a matematikai statisztikában számos eljárása ismeretes, ezeknek összegyűjtése és meteorológiai feladatok megoldására történő alkalmazása szintén Gulyás nevéhez fűződik. Ennek nyomán történt kísérlet a 80-as évek elején teljesen statisztikai, tehát nem vizuális kiválasztáson alapuló clusterezési eljárásra az Atlanti-európai térség időjárási típusainak osztályozása céljából.

Gulyás Ottó munkássága nagymértékben hozzájárult a fentebb felsorolt eljárások bevezetésével a kezdetben egy havi, majd később 3 ill. 6 havi éghajlati előrejelzések matematikai megalapozásához. A Bartholy Judittal, Kaba Magdolnával és Légrádi Gáborral közösen írt "Meteorológiai mezők számítógépes analízise" c. pályázati tanulmány akadémiai kutatási jutalmat kapott 1978-ban.

Úttörő munkának számított az 1978-ban indult ún. Konzerv program keretében a termésmennyiség és érés előrejelzésének matematikai modellezése. Paradicsomra, uborkára, zöldborsóra és más konzervipari terményre történtek ilyen számítások. Nem a matematikai megközelítés hibája, hanem a növényfejlődés bonyolultsága miatt kellett a kapott összefüggéseket időről-időre korrigálni.

A Módszertani Csoport és Gulyás Ottó tevékenysége nem korlátozódott a KMI-re, ill. az éghajlati és agrometeorológiai kutatások matematikai oldalról történő támogatására. Ennek egyik példája a Tänczer Tiborral végzett csapadék-előrejelzési módszer, melyben egyaránt megtalálhatók dinamikai eljárások (advekciós előrejelzési technika, légköri dinamikai paraméterek használata), és statisztikai eljárások (tanuló algoritmusok alkalmazása) a prediktorok és a prediktandusz (12 órás csapadékösszeg) közötti kapcsolat felderítésére és optimalizálására.

Másik kirándulása 1984-ben a többsávos digitális műholdképek számítógépes clusterezése volt. Matematikai szempontból az ilyen műholdképek hatalmas tömegű, áttekinthetetlen adathalmazt jelentenek. A feldolgozás folyamatának egyik fontos láncszeme a clusterezés, melynek során az azonos jellegű képpontok közös csoportba kerülnek, és ezáltal lehet a felvétel képi struktúráját, lényegét megragadni. Szerzőtársaival (Korándi Márta és Ketskeméti László) gradiens operátoros clusterező algoritmust dolgozott ki. Tesztelését LANDSAT, METEOR, GOES műholdképeken végezték el. A feladat végrehajtása az akkor rendelkezésre álló IBM-370 számítógépen igen komoly munka volt, jelentős gépidőt kötött le.

Gulyás nevével találkozhatunk a jégeső-elhárítást megalapozó vizsgálatokban is. Az egyes radar-paraméterek matematikai-statisztikai elemzése útján kísérlet történt a jégkihullási valószínűség előrejelzésére. Egy Moldáviában kifejlesztett módszer adaptálásával és továbbfejlesztésével sikerült elérni, hogy a jeget adó felhőket 80-90 %-os biztonsággal megkülönböztethessék a jeget nem adóktól.

E rövid előadás keretében nem lehet felsorolni az összes általa irányított, vagy részvételével végzett kutatási témát. Ezért elnézést kérek mindazon egykori munkatársaitól, akik nevét nem említettem.

Közel másfél évtizedes OMSZ-beli pályafutása során tanulmányok tucatjai kerültek ki kezéből, nem szerette, ha valami íróasztalának fiókjában reked. Csaknem mindegyik munkáját társszerzőkkel írta. Ennek magyarázata egyszerű: igyekezett elméleti tudását a gyakorlatban hasznosítani, s ehhez mindig megkereste a feladat meteorológiai oldalát ismerő szakembereket. Többször kijelentette: ő matematikus, és nem akar meteorológus lenni, csak annyit akar tudni a feladat meteorológiai oldalából, amennyi a megfelelő matematikai módszer megválasztásához szükséges.

Publikációs tevékenységének csúcsát a halála évében megjelent testes egyetemi tankönyve jelenti, melynek társszerzője Dévényi Dezső volt. Ebben számos korábbi tanulmányának eszenciája is megtalálható.

Termékeny tudományos munkássága mellett nagy hozzáértéssel foglalkozott a fiatal szakemberek képzésével. Sok éven keresztül tanított matematikai statisztikát meteorológus hallgatók számára, sokan írtak vezetése alatt szakdolgozatot, doktori disszertációt, elismert aspiránsvezető volt. Iskola-teremtő munkásságának egyik gyümölcse volt az 1987-ben Síkfőkúton rendezett szeminárium, ahol közel 40 egykori tanítványa tartott előadást, felhasználva a tőle tanult matematikai ismereteket. Ő indította meg és egy évtizeden keresztül szervezte a KMI matematikai szemináriumát, kiváló lehetőséget nyújtva fiatal kutatók ismeretszerzésére és továbbképzésére.

Kezdeményezője, majd az OMSZ részéről koordinátora volt az OMSZ-KSH együttműködésnek, melynek keretében az OMSZ jelentős mennyiségű gépidőhöz jutott a KSH IBM-360, majd 370 típusú számítógépén, ami annak idején nagy szó volt, hiszen ilyen gép nem volt minden igénylő számára elérhető. A kapcsolat egészen addig fennállt, míg az OMSZ-KMI számítóközpontjában a TPA gépek, majd a Tatabánya téren a KEI mellett a BASF számítógép telepítésére sor nem került.

Szorosan vett szakmai munkáján túl, sok energiát szentelt tudománypolitikai kérdéseknek is. Nagy aktivitással és elhivatottsággal vett részt az OMSZ Tudománypolitikai Bizottságának (TUPIB) munkájában. Ma tíz évvel ezelőtt bekövetkezett halálával nem csak matematikusként, hanem közéleti személyiségként is nagy űrt hagyott maga után.

# Gulyás Ottó munkássága matematikus szemmel

Szeidl László

Egy korán megtört, de szakmailag mégis rendkívül gazdag életpályáról szeretnék beszélni. Jelen megemlékezés keretein belül természetesen nem lehet részletesen kitérni Gulyás Ottó teljes szakmai tevékenységére, hanem csak annak főbb vonásaira. Célom nem egy leltár készítése, hanem egy kiváló szakember bemutatása.

Mindenekelőtt egy személyes megjegyzést szeretnék tenni. Ottóval 1972-ben ismerkedtem meg közelebbről, az MTA Számítástechnikai és Automatizálási Kutató Intézete és a Távközlési Kutató Intézet közös szakmai szemináriumán, s kapcsolatunk egészen haláláig megőrződött. A számos közös szeminárium, a sok-sok kötetlen beszélgetés alkalmat adott mélyebb szakmai és emberi ismeretségre. A szemináriumoknak nem csak új eredmények ismertetése volt a célja, hanem adott területeknek, pl. idősorok elemzésének a feldolgozása, alaposabb megismerése és közkincsé tétele is (könyvek formájában). Ezekben a munkákban Ottó mindig lelkesen és aktívan vett részt.

Bár kutatási területeink, szakmai érdeklődési körünk és szemléletünk nem esett távol egymástól, a sors úgy hozta, hogy közös kutatási feladaton nem dolgoztunk együtt. Ebben, sajnos, nagy szerepe van korai halálának is, mivel nem sokkal előtte kezdtem el foglalkozni meteorológiai idősorok statisztikai modellezési kérdéseivel, amely szakmailag még közelebb hozott minket egymáshoz.

Matematikus szemmel érdekes elgondolkodni azon, hogy Gulyás Ottó szakmai útja hogyan vezetett a tisztán elméleti-matematikai kérdések vizsgálatától gyakorlati feladatok matematikai-számítástechnikai modelljeinek kidolgozásához, megoldásához és hogyan lett ő a statisztikus meteorológia magas szintű szakmai képviselője, oktatója és elkötelezett, lelkes propagálója.

A '60-as évek végén, '70-es évek elején a Távközlési Kutató Intézetben dolgozott egy igen erős szakmai felkészültségű matematikusokból, mérnökökből, meteorológusokból és orvosokból álló kollektíva. Ennek munkájában Gulyás Ottón kívül Ambrózy Pál, Csibi Sándor, Faragó Tibor, Götz Gusztáv, Győrfi László és mások vettek részt. Ennek a kollektívának az alapvető célja az volt, hogy kifejlesszen kisszámítógépekre számítástechnikai diagnosztikai berendezéseket. Igen tanulságos és a feladat komplex jellegét jól mutatja a Távközlési Kutató Intézet 3 vaskos tanulmánya, amely 3 egymást követő évben (1969, 1970, 1971) készült el, s amely számot ad az adott időszakban elvégzett munkákról. Ezekben a munkákban jelentős szerepet vállalt magára Gulyás Ottó nem csak a matematikai kutatásokban, hanem a gyakorlati modellek felállításában, az algoritmusok kialakításában és azok számítástechnikai megvalósításában is. Lényegében az ezirányú eredményeinek összefoglalásaként született meg 1973-ban kandidátusi értekezése is.

Ennek az igen sokrétű feladatnak az egyik kulcskérdése olyan tanuló-felismerő algoritmusok kidolgozása volt, amely elvégezte a vizsgált objektumok automatikus osztályozását, vagyis az objektumok szeparálását, kategorizálását és osztályozását. Természetesen ezeknek az algoritmusoknak a kidolgozása szoros összefüggést mutat az egyes gyakorlati feladatok matematikai-számítástechnikai modelljének kidolgozásával, lényegkiemelő módszerek kifejlesztésével és megvalósításával.

Mielőtt a problémakör néhány fontos matematikai vonatkozására kitérnék, itt jegyzem meg, hogy e munka keretében Gulyásék által megoldott első gyakorlati feladat a konvektiv aktivitás (*zivatar*) előrejelzésével függött össze. A meteorológiai modellt Szalay és Götz dolgozták ki (*Időjárás*, 1971), majd közösen alakították ki a tanulóalgoritmusok alkalmazására a megfelelő változatot és végezték el az első számítógépes kísérleteket. Úgy gondolom, hogy ez a közös munka döntő befolyással volt Gulyás Ottó későbbi pályafutására is.

Mint korábban megjegyeztem, Gulyás Ottó 1973-ban készítette el kandidátusi disszertációját, melynek címe: "Az alakfelismerés néhány matematikai kérdése és alkalmazása" volt. Az értekezés jól visszatükrözi a vizsgált probléma komplex jellegét, ugyanakkor rávilágít a szerző alapos matematikai felkészültségére, igen széleskörű matematikai intelligenciájára, kreativitására és a gyakorlati feladatokkal szembeni fogékonyságára. Egyébként az értekezés nem csak az addigi munkáinak egyfajta összefoglalását jelentette, hanem szakmai útjának egyik fontos állomása is volt, hiszen nem sokkal később (1974) a Központi Meteorológiai Intézetbe való kerülésével feladat- és érdeklődési köre nagymértékben megváltozott.

Az elkövetkezőkben röviden vázolni szeretném azokat a matematikai megközelítéseket és problémákat, amelyek e többéves kutatómunka során felmerültek és amelyek sikeres megoldást nyertek.

Az alakfelismerés feladata olyan módszerek kidolgozása, melyekkel a vizsgált objektumokról eldönthetjük, hogy meghatározott osztályok közül melyikbe tartoznak. Emellett fontos szempont az is, hogy a kidolgozott eljárás számítástechnikailag megvalósítható legyen. Ez az osztályozás általában statisztikus jellegű, ezért az alakfelismerés a feladat célkitűzésében, módszereiben a statisztikus döntéselmélethez áll közel, azonban több olyan sajátosságot mutat, ami miatt más megközelítést is igényel. Ilyen sajátosságot jelent pl. az, hogy

1) általában hiányoznak az a priori ismeretek;

2) a megfigyelés általában nem az osztályozandó objektum, hanem annak egy véges dimenziós Euklideszi térbe történő leképezése. Ez utóbbi a gyakorlati alkalmazások során éppen a lényegkiemelést jelenti, amely a feladatok igen fontos részét képezik nem csak matematikai, számítástechnikai szempontból, hanem a különböző kutatási területek szempontjából is, hiszen a különböző területek regisztrátumai teljesen eltérő vonásokat és meghatározó tulajdonságokat mutathatnak (pl. meteorológiai, orvosi alkalmazások esetén).

3) A megfigyelések és hozzátartozó paramétereik sorozata, vagyis a tananyag alapján a megfelelő tanuló algoritmusnak olyan döntésfüggvény-sorozatot kell generálnia, amelyhez tartozó hibavalószínűség bizonyos értelemben optimális.

Gulyás Ottó fontos és erősen általános elméleti matematikai eredményeket ért el mind a lényegkiemelés, mind pedig a tanuló algoritmusok vonatkozásában a reprodukáló magú Hilbert terek módszerének felhasználásával. Vizsgálataiban az osztályozandó objektum egy sztochasztikus folyamat (pl. meteorológiai idősorok, vagy EKG regisztrátumok), s ezekkel kapcsolatban többféle lényegkiemelő megközelítést vizsgált, úgy mint a mintavételi sorfejtés, a Karhunen-Loéve sorfejtés, a regressziós típusú sorfejtés és ért el általános eredményeket - a jól ismert standard tételeket magába foglalva.

Munkájának egyik lényeges része volt a tanuló algoritmusok (elsősorban potenciálfüggvényes algoritmusok) tárgyalása és reprodukáló magú Hilbert terek elméletének felhasználásával történő általánosítása. Ez az általánosítás sok esetben lehetővé teszi a potenciálfüggvény konkrét megválasztását. Több algoritmus konvergenciáját bizonyította be és adott becslést a konvergencia sebességére.

Az elért elméleti-matematikai eredmények igazi felhasználást is nyertek a már említett konvektiv aktivitás előrejelzésével kapcsolatban és az EKG regisztrátumok vizsgálatánál. Talán nem árt hangsúlyoznom, hogy ezeknek az eredményeknek az elérésében a kutató kollektíva többi matematikus, mérnök, meteorológus és orvos tagjának is óriási szerepe volt, azonban ez nyilvánvalóan nem csökkenti Gulyás Ottó érdemeit.

Gulyás Ottó szakmai tevékenysége lényeges változáson ment keresztül a Központi Meteorológiai Intézetbe való kerülésével. Bár matematikai szempontból érdeklődési köre nem változott, kutatómunkája középpontjába - magától értetődően - elsősorban meteorológiai problémák kerültek.

Nagy energiákat fektetett abba, hogy minél szélesebb körben nyerjenek alkalmazást a meteorológiai kutatások során a korszerű valószínűségelméleti, statisztikai módszerek. Ezt nem csak az elméleti háttér leírásával, előadások tartásával, hanem programok és alkalmazási példák sorával is megpróbálta elősegíteni.

Matematikus szemmel érdekes felfedezni azokat a korábbi munkáiban is megmutatkozó matematikai megközelítési módokat, módszereket és egyáltalán azt a szemléletet, amelyek gyümölcsözően hatottak a különböző - meteorológusokkal közösen folytatott - kutatómunkákra. Ennek bemutatására szeretnék megemlíteni néhány publikációt. Például: "*Az analógia elvén alapuló prognosztikai módszerek matematikai modellje*" (*Időjárás*, 1975, társszerzők: Faragó Tibor és Kaba Magdolna), "*Az analógia fogalma és felhasználása típusok képzésére I. és II.*" dolgozat (*Időjárás*, 1977. 1. és 6. szám, társszerzők: Bartholy Judit és Légrády Gábor).

Ezekben a cikkekben a problémakör olyan elvi és matematikai kérdéseiről van szó, mint a meteorológiai típusképzés matematikai modelljének elemzése, a már meglévő meteorológiai típusok összehasonlító statisztikai elemzése és az automatikus osztályképző és osztályba soroló számítógépes módszer kidolgozása és alkalmazása.

Hasonlóan jó példákként említhetők a következő publikációk is: "*A csapadék rövidtávú előrejelzése tanuló algoritmusok segítségével*" (*Időjárás*, 1979, 2. szám, társszerzők: Bak Judit és Tänczer Tibor) c. cikk és az "*A method of analysing types using analogy indeces*" (*Acta Climatologica*, 1980, társszerző: Bartholy Judit), valamint a "*Távsávos digitális műholdképek számítógépes clusterezése*" (*Időjárás*, 1984, 2. szám, társszerzők: Ketskeméty László és Korándi Márta) dolgozatok.

Természetesen lehetne találni még további publikációkat, de talán ezek a legjellemzőbbek.

Szeretnék szólni Gulyás Ottó kiemelkedő oktatói, utánpótlás-nevelői tevékenységéről is. Tudományos munkája a Központi Meteorológiai Intézetbe való kerülése után kiegészült igen komoly és rendszeres oktatói munkával is különösen azután, hogy az ELTE Természettudományi Karán 1976-ban újra megindult az önálló meteorológusképzés. Több mint egy évtizeden keresztül vett részt a Valószínűségelméleti és Statisztikai Tanszék oktatói munkájában. Ő tartotta a meteorológus hallgatóknak a "Valószínűségelmélet és matematikai statisztika" c. tárgyat. Óráira mindig körültekintően és akkurátusan készült fel, a tanítás mindig szívügye volt.

Nemcsak az oktatást magát tartotta fontosnak, hanem az azt nagymértékben támogató jegyzetek, könyvek és egyéb oktatási segédanyagok készítését is. 1978-ban jelent meg a "*Valószínűségszámítás és matematikai statisztika*" c. egyetemi jegyzete a Tankönyvkiadónál és 10 évvel később a Dévényi Dezsővel közösen írt "*Matematikai statisztikai módszerek a meteorológiában*" c. könyve, ugyancsak a Tankönyvkiadónál.

Az egyetemi oktatásban is jól alkalmazható *"Idősorok analízise"* (Műszaki Könyvkiadó, 1986, szerkesztette: Tusnády Gábor és Ziermann Margit) és a *"Többváltozós statisztikai analízis"* (Műszaki Könyvkiadó, 1986, szerkesztette Móri F. Tamás és Székely J. Gábor) könyvekbe egyaránt írt egy-egy fejezetet. Számos tanítványa van. Sokan írtak szakmai irányítása mellett egyetemi szakdolgozatot, doktori disszertációt és több aspiránsa szerzett kandidátusi címet. Kutatás iránti érzékenységet és igényességet felkeltő attitűdje a mai napig kihat.

Gulyás Ottó tudásátadó tevékenysége nem szorítkozott kizárólag az egyetemi szférára. Szemináriumok szervezésével, előadások tartásával, szemináriumi anyagok készítésével állandóan stimulálóan hatott szakmai környezetére.

Sokat foglalkozott a szakterület tudománypolitikai kérdéseivel és talán sikerült új szint, új szemléletet is vinni abba. Több dolgozatában és előadásában, mint pl. *"Matematikusok a meteorológiában", Légkör*, 1985. 2. szám. *"Megjegyzések a matematikai statisztika és meteorológia kapcsolatáról"* c. előadásában igyekezett tisztázni a matematikusok helyét és szerepét a meteorológiai kutatásokban, illetve fejtette ki elképzeléseit arról, hogy milyen kapcsolat kialakítása szolgálná az eredményes előrelépést.

A meteorológia társadalmi megítélése, népszerűsítése is komolyan foglalkoztatta. A *"Néhány szó a meteorológia népszerűsítéséről", Légkör*, 1987. 2. szám c. cikkében kiemeli, hogy mennyire lényeges "a társadalom tájékoztatása az információk helyes értelmezéséről, az információk egyértelmű, pontos közlése, a tudományos korrektség a hibák elismerésében. Ennek érdekében nagyon fontos teendő a társadalom meteorológiai kultúrájának emelése."

E cikk egyik félmondatából az is kiderült, hogy a matematikától elindulva másfél évtized alatt milyen közel került a meteorológiához, hiszen már nem kívülállóként mondta: "...a napi munkáink színvonalas elvégzésén kívül mi is tehetünk többet a meteorológia társadalmi megítélésének kedvezőbb alakulásáért..."

Gulyás Ottó személyében egy nagy szakmai tudású, széles műveltségű és látókörű matematikust, mondhatom talán azt is, hogy statisztikus meteorológust, remek gyakorlati érzékkel bíró modellezőt és kiváló kollégát vesztettünk el. Emlékét megőrizzük!

# Éghajlati idősorok statisztikai elemzése

Szentimrey Tamás

*Bevezetés*

1981-ben az egyetem matematikus szakának befejezése után az OMSZ statisztikai Módszertani Csoportjánál helyezkedtem el. A csoport vezetője Gulyás Ottó volt, nem csupán hivatalos értelemben, hanem az elhivatottságot tekintve is. Ennek a viszonylag csekély létszámú csoportnak az alapvető feladatköre a matematikai statisztikai módszerek elfogadtatása, terjesztése volt a meteorológia területén, és tevékenységének meghatározó személyisége Gulyás Ottó volt. Természetesen sok személyes emlékem fűződik hozzá, hiszen éveken keresztül együtt dolgoztunk, azonban egyrészt érzésem szerint ezek felidézéséhez tíz év talán még nem elég hosszú idő, másrészt fontosabbnak tartom, hogy a magam nyilvánvalóan szubjektív szemszögéből megpróbáljam a tevékenységét bemutatni, értékelni. Ha egy mondatban kellene összefoglalnom, akkor azt mondanám, hogy Gulyás Ottó missziónak tekintette a matematikai statisztika tudományának megismertetését valamint alkalmazási feltételeinek megteremtését a hazai meteorológus tudományos közéletben. Miben is nyilvánult meg ez? Véleményem szerint három területet tartott alapvető jelentőségűnek. Mindenekelőtt az oktatást, beleértve a valószínűségszámítás és matematikai statisztika elméletének megismertetését, valamint a sztochasztikus modellezés, gondolkodás elfogadtatását is. Az ő idejében kezdődött meg a meteorológus hallgatók ez irányú képzése, továbbá színvonalas szeminárium sorozatokat szervezett. Hasonlóan fontosnak tartotta a létező matematikai statisztikai módszerek, eljárások lehetőleg minél szélesebb körű és nem utolsósorban korrekt alkalmazását. Ezen a téren ő maga járt elöl jó példával, hiszen számos meteorológiai kutatási témában vett részt, hasznosította tudását. És végül de nem utolsósorban, igen nagy súlyt helyezett a matematikai statisztikai módszerek kutatására, fejlesztésére a meteorológia tudományának sajátos igényei szerint. Úgy gondolom, hogy célkitűzéseiből igen sok minden megvalósult, illetve remélhetőleg megvalósulóban van, aminek következtében a hazai statisztikus meteorológia nemzetközi mércével mérve is megállja a helyét. Azt hiszem, hogy ebben Gulyás Ottóé a fő érdem.

Előadásom tárgya az éghajlati idősorok elemzésével kapcsolatos, azonban nem kívánok teljes általánosságában foglalkozni a kérdéskörrel, hanem inkább egy rövid történeti áttekintés bemutatására törekszem. A múltat és a jelent szeretném összehasonlítani, vagyis, hogy a témával kapcsolatosan milyen problémákkal foglalkoztunk még Gulyás Ottóval együtt, illetve milyenekkel foglalkozunk jelenleg. Az évek folyamán új éghajlati kérdések is felmerültek, hiszen míg korábban az egyik központi kérdés az éghajlatváltozás detektálása volt, az adatsorok alapján, mára az éghajlati adatsorok megbízhatósága is kérdésessé vált, és ezzel kialakult egy új terület, a homogenitásvizsgálat. A hangsúlyt azonban nem az eltérésekre szeretném helyezni, hanem a hasonlóságra, nevezetesen arra a törekvésre, hogy az éghajlati problémák megoldására adekvát statisztikai módszereket fejlesszünk, illetve alkalmazzunk.

## *1. Éghajlati idősorok trendvizsgálati modelljei*

A statisztikai trendvizsgálat diszkrét esetének általános modellje, hogy a vizsgált idősor

$$Y(t) = m(t) + e(t) \qquad t=1,2,\ldots,n \qquad (1)$$

alakban áll elő, ahol $m(t)$ az $Y(t)$ időbeli változásának alaptendenciáját jellemző trendfüggvény - várható érték függvény -, az $e(t)$ "zajrész" elemei pedig azonos eloszlásúak, 0-várható értékűek és teljesen függetlenek. A továbbiakban, a különféle módszerek ismertetésénél általában feltételezzük a zaj normalitását is, azaz, hogy $e(t)$ ( $t=1,2,\ldots,n$ ) normális fehér zaj.

A statisztikai trendvizsgálat célja a trendfüggvény és a zaj szétválasztása, pontosabban minél több, jól értelmezhető információ szerzése a trendfüggvényről egy adott valószínűségi szinten. Lehetőségeink korlátozottak, hiszen az adott idősor egyetlen realizációja alapján a trendfüggvényt nem ismerhetjük meg, legfeljebb bizonyos jellemzői becsülhetők, valószínűsíthetők. Lényeges, hogy a trendfüggvényre vonatkozó állításaink megbízhatóságának mértéke éppen az adott szignifikancia-szint, így a hipotézisvizsgálati eljárásoknak - ebben a vonatkozásban is - igen nagy szerepe van.

Az éghajlati idősorok trendvizsgálatánál korábban, így tíz évvel ezelőtt is, abból a modellből indultunk ki, hogy

I. $m(t) = C(t)$

$C(t)$: a feltételezett éghajlati változás jele,

azaz a trendfüggvény reprezentálja az esetleges éghajlatváltozást.

Gulyás Ottóval együtt olyan statisztikai módszerek fejlesztésére, illetve alkalmazására törekedtünk, melyek lehetővé teszik az éghajlati változás korrekt jellemzését, az adott trendvizsgálati modell alapján. Ebből a szempontból az információk értelmezhetőségét, valamint a hipotézisvizsgálati problémákat tartottuk a legfontosabb kérdéseknek. (Lásd 2., 3. fejezetek)

Az utóbbi években egyre több kétség merült fel az éghajlati idősorok megbízhatóságával kapcsolatban, nevezetesen, hogy bizonyos lokális beavatkozások következtében nem tekinthetőek homogénnek. Ilyen hatások lehetnek a megfigyelési rendszerben bekövetkező változások, illetve egyéb lokális antropogén tényezők, melyek okozhatnak bizonyos változást az adott éghajlati elem változékonyságában, ami helytelen következtetésekhez vezethet az éghajlat vizsgálata szempontjából. Az ennek megfelelően módosított trendvizsgálati modell, hogy

II. $m(t) = C(t) + IH(t)$

$IH(t)$: az esetleges inhomogenitás jele,

azaz a trendfüggvény az éghajlati változás jelén kívül, még esetleg bizonyos inhomogenitásokat is tartalmaz. Ez esetben a vizsgálatok célja, minél több szignifikáns és jól értelmezhető információ szerzése $C(t)$ - ről és $IH(t)$ - ről, továbbá az esetleges inhomogenitások kiszűrése. ( Lásd 5. fejezet ).

## 2. Lineáris analitikus trendvizsgálat és általánosítása

A lineáris analitikus modellnél feltételezzük, hogy

$$m(t) \equiv \sum_{i=1}^{L} c_i \cdot f_i(t) \qquad (\, L \ll n\, ;\, t=1,2,\ldots,n\, ) \tag{2}$$

ahol $f_i(t)$ egy adott függvényrendszer. Az ismeretlen $c_i$ együtthatók a legkisebb négyzetek módszere szerint becsülhetők, az $Y(t)$ idősor alapján. Természetesen a (2) azonosság feltételezése csupán egy a priori leegyszerűsítése a tapasztalatilag esetleg elfogadható (1) trendvizsgálati modellnek, abból a célból, hogy matematikailag kezelhetővé tegyük a feladatunkat. Tehát óhatatlanul felvetődik a kérdés, hogy általános esetben hogyan értelmezhetők a lineáris analitikus trendvizsgálat eredményei.

### 2.1. A lineáris analitikus modell általános értelmezése

Az általános értelmezés szerint a $c_i$ együtthatókat az alábbi minimum-feladattal definiáljuk:

$$\sum_{t=1}^{n} \Big( m(t) - \sum_{i=1}^{L} c_i \cdot f_i(t) \Big)^2 = \min ,$$

azaz a $c_i$ együtthatók a trendfüggvény - adott függvényrendszerre vonatkozó - vetületének együtthatói. Az együtthatók becslése továbbra is a legkisebb négyzetek módszere szerint történik, és a becslések tulajdonságai hasonlóak a pontos illeszkedés - a (2) azonosság teljesülése - esetére vonatkozó tulajdonságokhoz (Szentimrey,1989 ).

### 2.2. Hipotézisvizsgálati kérdések az általános értelmezés alapján

i, Null - hipotézis: $c_i = 0$
Az általános esetben is: t-próbák ( T próbastatisztikák ).

ii, Null - hipotézis: $\forall\, c_i = 0$
Az általános esetben is: F-próba.

iii, Null - hipotézis: elfogadható a (2) lineáris modell.
A vizsgált $Y(t)$ idősorhoz hozzárendelhető olyan $Z(t)$ idősor, melyre teljesülnek az alábbiak:
- a $Z(t)$ idősor elemei teljesen függetlenek, normális eloszlásúak, azonos szórásúak,
- a $Z(t)$ idősor elemeinek várható értéke 0, akkor és csak akkor, ha igaz a (2) lineáris modell.

### 3. *Trendvizsgálat ortogonális sorfejtés alapján*

Tegyük fel, hogy az $\{ f_i(t) \; ; i = 1,2,\ldots,n \}$ függvényrendszer teljes és ortogonális az $\{ 1,2,\ldots,n \}$ halmazon. Legyen a trendfüggvény diszkrét sorfejtése:

$$m(t) = \sum_{i=1}^{n} c_i \cdot f_i(t) \qquad t = 1,2,\ldots,n$$

A trendfüggvény megismerése szempontjából az ismeretlen $c_i$ együtthatók érdekelnek bennünket, melyek becsléseit ($c^*_i$) az alábbi sorfejtés alapján nyerhetjük:

$$Y(t) = \sum_{i=1}^{n} c^*_i \cdot f_i(t) \qquad t = 1,2,\ldots,n$$

#### 3.1. *Hipotézisvizsgálat*

A vizsgálat tulajdonképpeni célja a zérustól különbözö $c_i$ együtthatók meghatározása, adott szignifikancia-szinten. A válogatás során két fajta hibát követhetünk el: elsöfajú hibát, amikor létezik olyan zérus $c_i$ együttható, amiröl úgy döntünk, hogy zérustól különbözik, illetve másodfajú hibát amikor is létezik olyan zérustól különbözö $c_i$ együttható, amiröl úgy döntünk, hogy zérus. Statisztikai szempontból egy ilyen döntési eljárás akkor tekinthetö egzaktnak, ha az adott modellen belül meg tudjuk adni az elsöfajú hiba valószínüségét ($q$), illetve ami ezzel ekvivalens, a szignifikancia-szintet.

Szignifikáns döntési eljárás

Ha a döntés: $c_{i,1} \neq 0, \ldots\ldots, c_{i,K} \neq 0$, akkor

$$P( \; \exists c \in \{c_{i,1}, \ldots\ldots, c_{i,K} \}: c = 0 \; ) = q,$$

ahol $1 - q$ az adott szignifikancia-szint. A döntési eljárás végrehajtása a $c^*_i$ becslések elemzése alapján történik, az (1) trendvizsgálati modellböl következö statisztikai tulajdonságaik figyelembevételével.

#### 3.2. *Alkalmazások*

Periodicitás-vizsgálat ( diszkrét Fourier-féle sorfejtés )

A "Trendvizsgálat, ortogonális sorfejtés alapján" c. témakörrel pontosan a periodicitás-vizsgálat kapcsán kezdtünk el foglalkozni. (Gulyás és Szentimrey, 1986) , ( Hamed, Gulyás és Ketskeméty, 1986 ), (Hamed, Szentimrey és Gulyás, 1988 )

Eredeti célunk a kérdéskörrel kapcsolatos félreértések tisztázása volt.
A diszkrét Fourier-féle sorfejtéssel történő vizsgálatnál a trendfüggvényre jellemző esetleges periodicitásokat kívánjuk detektálni. Ennél a sorfejtésnél az

$1, \{ \cos(2\pi it/n), \sin(2\pi it/n) \}$ ,  $i = 1,2,\ldots,[n/2]$

trigonometrikus ortogonális rendszert alkalmazzuk, tehát lényegében a

$T_i = n/i$  $i = 1,2,\ldots,[n/2]$

periódusokhoz tartozó periodikus összetevőket vizsgáljuk. A hipotézisvizsgálat az úgynevezett periodogram elemzése alapján történik, ugyanis a sorfejtésnél egy vizsgált periódust két összetevő is reprezentál. A diszkrét Fourier-féle sorfejtés széles körben elterjedt módszer, azonban a "szignifikáns" periódusok fizikai értelmezése csaknem kivétel nélkül komoly problémát jelent.

Ablaktechnika ( diszkrét „Haar-féle" ortogonális sorfejtés )

Egy másik alkalmazási lehetőség a Haar-féle ortogonális rendszer diszkrét változatának kidolgozásán alapszik ( Szentimrey, Faragó and Szalai, 1992 ).
A vizsgálat célja, hogy a trendfüggvény esetleges változásáról bizonyos intervallumokon belül szerezzünk információt, pontosabban azt akarjuk eldönteni, hogy elfogadható-e az $m^{(1)} = m^{(2)}$ null-hipotézis egy adott intervallum-rendszer - ablakrendszer - elemein belül, ahol $m^{(1)}$ , $m^{(2)}$ a trendfüggvény megfelelő részátlagai. A probléma visszavezethető az ortogonális sorfejtés modelljére, ahol a Haar-féle ortogonális rendszer diszkrét megfelelőjét alkalmazzuk. Ez esetben az ortogonális rendszer elemei "lépcsősfüggvények", egy-egy ugrással, tehát végül is egy olyan ortogonális sorfejtésről van szó, ahol az együtthatók illetve az eredmények "fizikailag" is értelmezhetők. A vizsgálati eljárás lényegében úgy is felfogható, mint sok-sok független kétmintás t-próba egyszerre történő végrehajtása.

### *4. Éghajlati idősorok összehasonlítása lineáris regresszió alapján?*

Az éghajlati idősorok statisztikai elemzésének egy igen fontos területe az idősorok összehasonlítása. Az összehasonlítás gyakran a lineáris regressziós modell alapján történik - a múltban általában mi is ezt alkalmaztuk -, azonban egy ezzel kapcsolatos ellentmondásra most ez úton szeretnénk felhívni a figyelmet.
Legyenek $Y(t), X_1(t),\ldots, X_N(t)$ ( $t=1,2,\ldots,n$ ) éghajlati idősorok, és tételezzük fel , hogy az $Y(t)$ idősornak az $X_1(t),\ldots, X_N(t)$ idősorokra vonatkozó lineáris regressziója:

$$Y_r(t) = c_0 + \sum_{i=1}^{N} c_i \cdot X_i(t) \qquad t=1,2,\ldots,n$$

ahol $c_0, c_i$( $i=1,2,\ldots,N$ ) konstansok, azaz a lineáris regressziós kapcsolat időben állandó.

Felmerül a kérdés, hogy alkalmazható-e egyáltalán ez a modell az éghajlati idősorok vizsgálatánál? A következő - könnyen belátható - tétel azt mutatja meg, hogy ez a látszólag természetes matematikai feltételezés milyen klimatológiai ellentmondáshoz vezet.

*Tétel*

Tegyük fel, hogy (tfh) az $Y_1(t)$, $Y_2(t)$ ( $t = 1,2,\ldots,n$ ) idősorok között a lineáris regressziós kapcsolat „állandó", azaz

$$Y_1(t) = a_1 + b_1 \cdot Y_2(t) + e_1(t)$$

$$Y_2(t) = a_2 + b_2 \cdot Y_1(t) + e_2(t),$$

ahol $a_1$ , $b_1$ , $a_2$ , $b_2$ konstansok, és a lineáris regresszió tulajdonságai alapján pedig a regressziós hibák várható értékeire teljesül, hogy

$$\mathrm{E}(e_1(t)) = \mathrm{E}(e_2(t)) = 0 \qquad t = 1,2,\ldots,n.$$

Ekkor igazak az alábbiak:

i, $\mathrm{corr}( Y_1(t), Y_2(t) ) = \text{konstans} = r,$ $\qquad$ ( $t = 1,2,\ldots,n$ )

azaz a korrelációs kapcsolat szintén állandó.

ii, Ha $r^2 = 1$ , akkor kölcsönösen egymás lineáris függvényei.

iii, Ha $r^2 < 1$ , akkor az adott idősorok várható értékei időben nem változnak:

$$\mathrm{E}(Y_1(t)) = \text{konstans} , \quad \mathrm{E}(Y_2(t)) = \text{konstans}.$$

*Következmény*

Tekintettel arra, hogy a ii, eset a gyakorlatban szinte soha nem fordul elő, így tehát azt mondhatjuk, hogy a lineáris regressziós modell már eleve kizárja az éghajlatváltozást, legalábbis várható értékben. Ugyanakkor az éghajlati idősorok vizsgálatának egyik legfontosabb célja pontosan az esetleges változás detektálása. Következésképpen nem javasoljuk az ilyen jellegű vizsgálatoknál a lineáris regressziós modell alkalmazását, hiszen a használt modell implicite a "változatlanság" eredményét is tartalmazza.

## *5. Homogenitásvizsgálat*

Az éghajlati idősorok homogenitásvizsgálatának ( lásd 1. fejezet: II. modell ) alapvető módszere az idősorok összehasonlítása, és az ilyen összehasonlításon alapuló vizsgálati módszereket szokás relatív vizsgálatoknak is nevezni. A továbbiakban az általunk kidolgozott statisztikai eljárás leglényegesebb matematikai alapelveit kívánjuk vázolni.
( MASH : Multiple Analysis of Series for Homogenization ; Szentimrey , 1997 )

### 5.1. *Additív modell*

Vizsgált idősor

$$Y(t) = C(t) + IH(t) + e(t) \qquad t=1,2,...,n$$

Referencia idősorok

$$X_i(t) = C(t) + e_i(t) \qquad ( i=1,2,..,N ;\ t=1,2,...,n )$$

$C(t)$ : a feltételezett közös éghajlati változás jele
$IH(t)$ : az esetleges inhomogenitás ( lépcsősfüggvény )
$e(t)$, $e_i(t)$ ( $i$=1,2,..,$N$ ): fehér zajrészek, együttes eloszlásuk normális

A vizsgálat célja:
Az $IH(t)$ inhomogenitás becslése ( töréspontok detektálása, ugrások becslése ) ,
és az $Y(t)$ korrigálása

### 5.2. *Az idősorok összehasonlítása*

Differencia idősorok

$$Z_w(t) = Y(t) - \sum_{i=1}^{N} w_i \cdot X_i(t) = IH(t) + d_w(t) \qquad t=1,2,...,n$$

ahol $\sum w_i = 1$ és $w_i \geq 0$ ( $i$=1,2,..,$N$ ), azaz a referencia idősorokhoz tartozó súlytényezők nem negatívok és az összegük egy. A differencia idősorok már nem tartalmazzák az éghajlati változás jelét, tehát alkalmasak az esetleges inhomogenitás vizsgálatára.

Optimális differencia idősor

A vizsgálat hatékonysága a "jel, zaj arányon" is múlik, ezért célszerű a szórásnégyzet minimalizálása. Optimálisnak tekintjük azt a differencia idősort, amelynek a szórásnégyzete minimális, azaz

$$\text{Var}\,(Z_w) = \text{Var}\,(d_w) = \min$$

Optimális differencia idősor rendszer

Általában nem tételezhető fel a referencia idősorok homogenitása, ezért több differencia idősort vizsgálunk, amelyekre teljesül, hogy nem tartalmaznak közös referencia idősort, és a szórásnégyzetük „minimális". Az első feltétel biztosítja a detektált töréspontok elkülöníthetőségét a vizsgált idősor számára.

### 5.3. *A töréspont detektálás matematikai formalizálása*

Legyen a vizsgált differencia idősor

$$Z_w(t) = IH(t) + d_w(t) \qquad ( t=1,2,...,n )$$

és tfh a valódi töréspontok halmaza : $\{ 1 \leq P_1 < P_2 <...< P_L < n \}$

azaz $IH( P ) \neq IH( P+1 )$.

Az töréspontok detektálásánál kétféle hibát követhetünk el. Egyrészt homogén szakaszoknál is töréspontokat vélhetünk felfedezni, másrészt esetleg létező töréspontokat nem tudunk kimutatni. Mivel általában a homogenitás null-hipotézise a természetes kiindulópontunk, ezért kézenfekvő a statisztikai konvencióknak megfelelően az első hibalehetőséget elsőfajú hibaként, míg a második hibalehetőséget másodfajú hibaként definiálni. A vizsgálatunk akkor tekinthető statisztikai szempontból egzaktnak, ha az adott modellen belül meg tudjuk adni az elsőfajú hiba valószínűségét, illetve ami ezzel ekvivalens, a szignifikancia-szintet.

Tfh a töréspontok becslései : $\{ 1 \leq P^*_1 <...< P^*_{L^*} < n \}$ahol $L^*$ töréspontok számának becslése.

(i) Elsőfajú hiba (szignifikancia)

Létezik olyan $P^*_l : ( P^*_{l-1} , P^*_{l+1} ) \cap \{ P_1 < P_2 <...< P_L \} = \varnothing$.

Feleslegesen detektálunk töréspontot. Törekedni kell a hiba valószínűségének ($q$) megadására ( szignifikancia-szint: $(1-q)$) .

(ii) Másodfajú hiba (hatékonyság)

Valódi töréspontokat elhanyagolunk. Lehetőleg minél többet!

### 5.4. *Szignifikáns eljárás a töréspontok detektálására*

Alapstatisztikák

$$ST( i, p, j) = \frac{(p-i+1)(j-p) \; ( Z_{i,p} - Z_{p+1,j} )^2}{(j-i+1) \; S^2( Z )} \qquad ( 1 \leq i \leq p < j \leq n )$$

ahol i, j és p tetszőleges egészek, mélyekre teljesül a zárójelek közötti reláció :

$Z_{i,p}$ : az $[ i , p ]$ intervallumhoz tartozó átlag

$Z_{p+1,j}$ : a $[p+1 , j ]$ intervallumhoz tartozó átlag

$S^2( Z )$ : egy 'jó' becslése a $\mathrm{Var}(Z_w)$ szórásnégyzetnek.

Az intervallumok inhomogenitásának jellemzése

Az előbbi statisztikák alapján jellemezzük az intervallumok inhomogenitását:

$INH([k, l]) = \max(ST(i, p, j) \mid k \leq i \leq p < j \leq l)$

A kritikus érték ( $\alpha$ ) meghatározása

$P(INH([1, n]) > \alpha \mid Z_W(t) \text{ homogén }) = q$

ahol $1-q$ az adott szignifikancia-szint. Konkrét értékek Monte-Carlo módszerrel számolhatók.

Az eljárás alapvető tulajdonsága ( szignifikancia, hatékonyság )

$$\max_{l}(INH((P^*_{l-1}, P^*_{l}]) \leq \alpha < \min_{l}(INH((P^*_{l-1}, P^*_{l+1}]),$$

következésképpen, az adott szignifikancia - szinten:

- egyrészt valamennyi ( $P^*_{l-1}$, $P^*_{l}$ ) intervallumról elfogadható, hogy nem tartalmaz töréspontot,

- másrészt valamennyi ( $P^*_{l-1}$, $P^*_{l+1}$ ) intervallum tartalmaz töréspontot

## *Irodalom*


*Gulyás, O., Szentimrey, T.*, 1986: „An iterative method for the periodic analysis of long time series." Proceedings of the Third International Conference on Statistical Climatology, Vienna, 28-33.

*Hamed, A., F., Gulyás, O. és Ketskeméty, L.*, 1986: „Meteorológiai idősorok periodicitásának elemzése", Időjárás, 90, 14-23.

*Hamed, A., F., Szentimrey, T. és Gulyás O.*, 1988: „Meteorológiai idősorok periodicitásának elemzése II..", Időjárás, 92, 38-45.

*Szentimrey, T.*, 1989: "A lineáris analitikus trendvizsgálat néhány elvi, módszertani kérdése", Időjárás, 93, 151-156.

*Szentimrey, T., Farago, T. and Szalai, S.*, 1992: "Window technique for climate trend analysis", Climate Dynamics, 6:127-134.

*Szentimrey, T.*, 1997: "Statistical procedure for joint homogenization of climatic time series", Proceedings of the Seminar for Homogenization of Surface Climatological Data, Budapest, 47-62.

# Meteorológiai mezők objektív osztályozása

Bartholy Judit

*Bevezetés*

Nagy örömmel teszek eleget a felkérésnek, s tartom meg előadásomat ezen az ankéton, melyet Gulyás Ottó emlékének szentelünk. 1974. októberében, mint az ELTE TTK harmadéves meteorológus hallgatója kerestem fel Gulyás Ottót, s kértem meg nem írhatnám-e szakdolgozatomat az ő vezetése alatt. Ebben a szemeszterben az "Alakzatfelismerési technikák alkalmazásai a meteorológiában" címmel tartott nekünk speciális kollégiumot. Ettől kezdve - a 14 évvel később bekövetkező haláláig - vele dolgoztam. Ő volt a témavezetője szakdolgozatomnak, egyetemi doktori disszertációmnak és kandidátusi értekezésemnek, s mindhárom témája a matematikai statisztika és valószínűségszámítás meteorológiai alkalmazásai tág tárgykörbe sorolható. Meghatározóak voltak nekem ezek az első évek, s úgy hiszem tőle kaptam a legtöbb szakmai útravalót. A teljességre való törekedés nélkül, néhány olyan gondolatot, alapelvet idéznék fel, mely tőle származik, ma is időszerű, s szükséges továbbadni az utánunk jövő nemzedékeknek.

Missziót vállalt fel, s tudatosan teljesítette. Hitte, hogy a meteorológia tudományának szüksége van egy matematikusra, aki segít áttörni a falakat, s rajta keresztül beáramolhatnak az új valószínűségszámítási, matematikai statisztikai módszerek a szakterületre. Ahogy Czelnai Rudolf előadásából hallhattuk ez a hit reális szakmai igényeken alapult, de mégis sok ellenállásba ütközött. Gulyás Ottó szisztematikusan dolgozott, sok fronton egyszerre közelítette az általa elérni kívánt célt:

- megcélozta a fiatal nemzedéket, az egyetemi hallgatókat, a néhány éve végzetteket, bekapcsolódott az egyetemi *oktatásba*,
- nyílt légkörű, *matematikai szemináriumsorozatot* indított el, mely még halála után is működött néhány évig,
- nagyszámú *szakdolgozathoz, doktori értekezéshez, kandidátusi disszertációhoz* ajánlott témát, s vállalta vezetésüket,
- számos meteorológiai kutatási projektbe kapcsolódott be, s adott módszertani tanácsokat,
- megalakította a közvetlen igazgatói alárendeltségű Módszertani csoportot, mely lehetővé tette az intézeti struktúra egységei közötti szabad közlekedést, elhárította a széleskörű tanácsadás, együttműködések akadályait,
- több országos szakmai konferenciát szervezett, s ezáltal a vidéki egyetemek, szakmai fórumok felé is nyitott.

A jól megszervezett missziós munka mellett sikerének másik forrása iskolateremtő egyénisége volt. Vonzotta a fiatalokat. Mivel? Nehéz erre pontos választ adni. Aki felkereste őt, hogy részt szeretne venni kutatásokban, diákköri dolgozatot, vagy szakdolgozatot írna, biztos lehetett abban, hogy értelmes munkákkal lesz foglalkoztatva, számíthatott személyreszabott feladatokra, szigorú értékelésre egyszóval objektív megmérettetésre. Nemcsak mással volt szigorú, hanem magával is, ami a hitelesség egyik nélkülözhetetlen eleme.

Tanító volt. Nemcsak a szakmai tudást akarta továbbadni, hanem az évtizedek alatt önmaga számára összegyűjtött élettapasztalatokat is. A kutatások etikájára, a munka moráljára vonatkozó alapvető elveket, de ugyanúgy az órára/előadásra való felkészülés "konyhaszabályait", a hozzászólóknak, az opponenseknek való udvarias válaszadás trükkjeit, az akkor még nagyon idegennek tűnő tudományos "manager" szakma fontosságát. Mégis, mint ahogy gyakran előfordul szerte a világon nem mindenben, s főleg nem a befektetett energiák arányában kapott megbecsülést és elismerést Gulyás Ottó a meteorológiai szakterületen kifejtett munkájáért.

Gulyás Ottó iránti tiszteletből és a matematikai statisztikai módszerek meteorológiában való elterjesztése területén végzett tevékenységére emlékezve, ebben a retrospektív előadásban összefoglalom a meteorológia szakterületén használatos osztályozási módszereket, s néhány alkalmazást is bemutatok, értékelek.

*Nagytérségű osztályozások szubjektív és objektív módszerei*

Két típusát különíthetjük el a meteorológia területén jelentkező osztályozási feladatoknak: az egyik esetben egy vizsgált térrészen belül (regionalizáció), a másik esetben az időskálán történik a hasonló objektumok csoportosítása (mezők osztályozása). Mindkét esetben történhet ez az osztályozás szubjektív, vagy objektív módszerek felhasználásával.

Pontosabban megfogalmazva az osztályozások két alaptípusát:

*1. **Regionalizáció***

Az egész Föld, az északi, vagy déli félteke, vagy egy makrotérség (pl. Atlanti-Európai térség) valamilyen szempont szerint hasonló régióit határozzuk meg. Közismertek az ebbe a kategóriába tartozó Köppen-, vagy Trewartha-féle éghajlatosztályozások, ahol részben szubjektív módon (több évtizedes emberi megfigyeléseken alapuló), részben objektív eljárással (az évi középhőmérsékletek, ill. az évi csapadékösszegek küszöbértékei alapján) történt a régiók határvonalainak meghatározása.

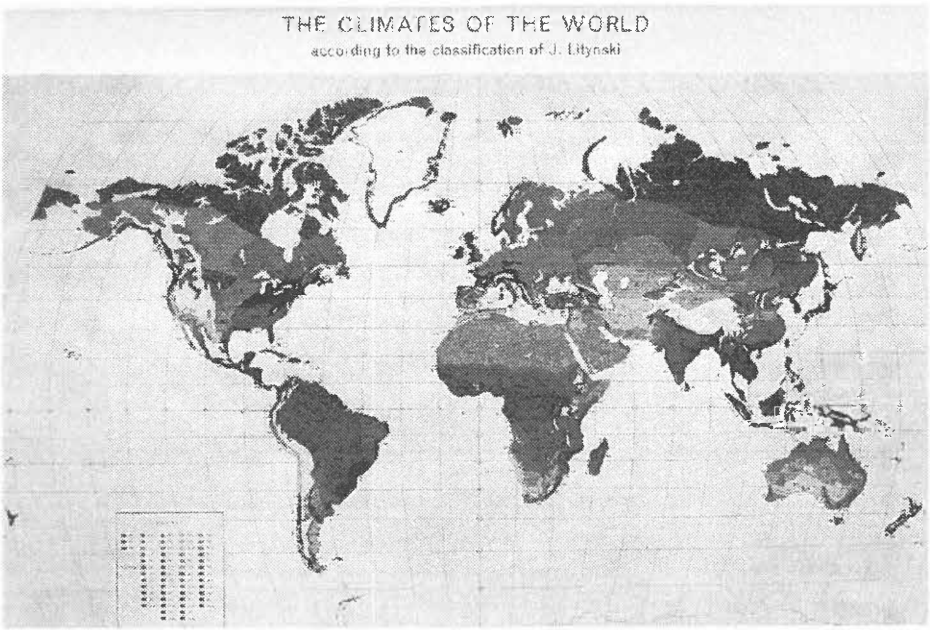
A regionalizáció teljesen objektív módon is történhet, mint ahogy Litynskinél (1983) láthatjuk, aki csupán számítógépes algoritmusok, s három mért, ill. számított paraméter (az évi középhőmérséklet, az évi csapadékösszeg, s a kontinentalitás mértékére definiált mennyiség) segítségével osztja 60 éghajlati régióra a teljes földfelszínt, melyeket az 1. ábrán mutatunk be.

*2. **Mezőosztályozások***:

A meteorológia területéről elsősorban a makrocirkulációs tipizálásokat soroljuk ebbe a csoportba, melyek célja nagy földrajzi térséget lefedő cirkulációs helyzetek osztályozása. Azaz, kerüljenek a hasonló cirkulációs helyzetek azonos osztályba, míg az egymástól különbözőek különböző osztályba. Itt is a szubjektív (pl. Hess-Brezowszky-féle makroszinoptikus típusok) ill. az objektív technikák terjedtek el, s az elmúlt évtizedekre mindinkább az utóbbi módszerek térhódítása a jellemző.

Amennyiben egy kicsit tágabban értelmezzük ezt a kategóriát és a kvázi-stacionáris makricirkulációs folyamatok közös karakterisztikáit kívánjuk megkeresni, kiemelni, csoportosítani, úgy a modern matematikai eszközök közül válogatva három út kínálkozik: a távkapcsolat analízis, az empirikus ortogonális függvények analízise, ill. a clusteranalízis.

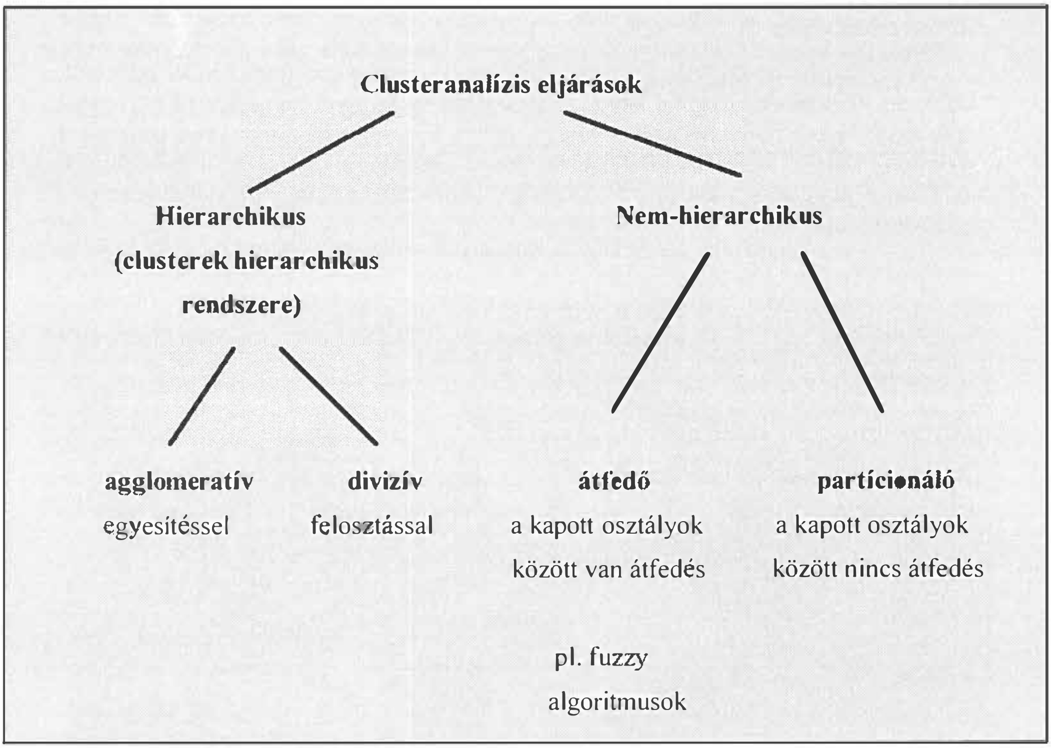
A távkapcsolatanalízis az egymástól nagy távolságra lévő akciócentrumok koherens változásait vizsgálja. ennél az eljárásnál nagyszámú korrelációs kapcsolatot tartalmazó térkép felhasználásával választjuk ki a legszignifikánsabb távkapcsolatokat (Namias, 1981). A módszer gyengéje, hogy esetenként több ezer térképből kell szubjektív válogatással kiszűrni a legerősebb kapcsolatokat, mely műveletben nagy a hiba lehetősége.




1. ábra:

Litynski (1983) a Föld 1300 állomásának évi középhőmérséklet, évi csapadékösszeg, s kontinentalitási index adatsorai alapján 64 klímát definiál (eloszlásfüggvényeik alapján), melyből csupán 60 fordul elő a valóságban

Az empirikus ortogonális függvényanalízis során az adott változó korrelációs mátrixának sajátértékegyenletét oldjuk meg. Mód nyílik a térség akciócentrumainak megjelenítésére a sajátértékek és sajátvektorok segítségével. Ennek az eljárásnak mindenképpen előnye az előző módszerrel szemben, hogy eredményként minden esetben csak néhány karakterisztikus alapmódot kapunk, valamint az algoritmus optimalitása is biztosított. E módszernél viszont - mivel matematikai struktúrákkal, s nem fizikai rendszerekkel operálunk az eljárás során - semmi biztosíték nincs arra nézve, hogy az eredményül kapott sajátvektorok, ill. alapmódok valódi fizikai tartalommal bírnak, s nem csupán fikciók. Rotációs eljárásokkal szoktak javítani a fenti problémán (Richman, 1985).



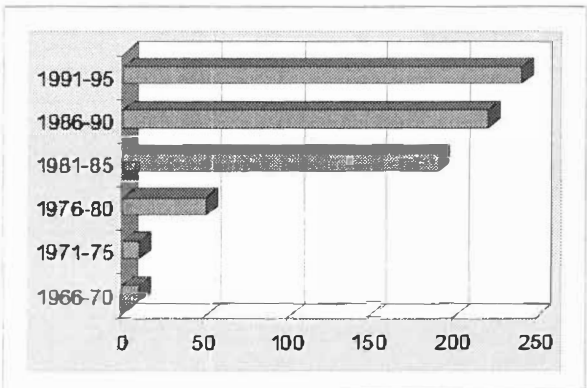

2. ábra:

Clusterező módszerek áttekintése

A clusteranalízis egyaránt alkalmas egy térség regionalizálására, vagy objektumok, mezők tipizálására, valamely hasonlósági kritérium alapján. A clustertechnikák módszertani áttekintő vázrajzát láthatjuk a 2. ábrán. A két fontos alaptípus a hierarchikus és a nem-hierarchikus eljárások csoportja. Az első esetben az algoritmusok alkalmazása a clusterosztályok egy hiarerchiáját feltételezi, s vagy a kiinduló csoportok összevonásával (agglomeráció), vagy azok felosztásával (divízió) jutunk a keresett osztályok rendszeréhez. A második, ún. nem-hierarchikus (gyakran dinamikusnak nevezett) módszerek esetén mindig egy konvergens iterációs technikával jutunk az általunk előre megadott számú clusterrendszerhez. Ezen osztályok átfedhetik egymást, vagy lehetnek diszjunktak, s e kritériumnak megfelelően nevezzük az eljárásokat átfedőnek, vagy partícionálónak. A legnagyobb problémát a clusterező eljárásoknál, mindig a távolságfogalom definiálása, a metrika megválasztása, ill. az iteratív eljárások leállítási szabályának megadása jelenti. Egy sikeresen kivitelezett clusterezést mindig több módszer kipróbálása, az adatok feltáró analízise előzi meg.

Az elmúlt két évtizedben 43 különböző clusterező algoritmust alkalmaztak a meteorológia szakterületén a kutatók. A 3. ábrán az 1965-95-ös időszakra vizsgáltuk a clustertechnikákat alkalmazó vezető folyóiratokban megjelent angolnyelvű cikkek számának alakulását. Az ábra tanúsága szerint az 1980-as évek elejére tehető a módszer elterjedése a meteorológiában, s azóta folyamatosan nő e területen publikált szakcikkek száma. Szokatlan ez a jelenség, hiszen más új eljárások megismerése és elterjedése általában sokkal hamarabb lecsengett a múltban. Erre részleges magyarázatot adhat, hogy a 90-es évek elején nagy lendülettel indult be a clusteranalízisnek egy új területen való alkalmazása, s ez az ensemble előrejelzések clusterezése.

A fentiekben felsorolt módszerek nagyszámú alkalmazásával találkozhatunk a hazai szakirodalomban is. Számos téma kezdeményezése még Gulyás Ottó nevéhez fűződik (Ambrózy-Bartholy-Gulyás, 1984; Bartholy, 1989).




3. ábra:

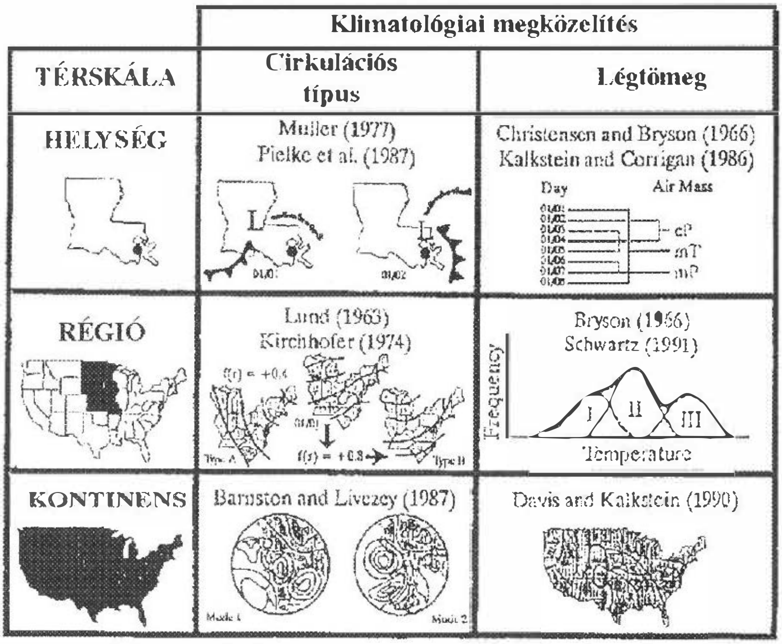
**Clusterező algoritmusokat alkalmazó kutatások száma**
**(1960-1995)**
**(Nemzetközi angolnyelvű folyóiratokban publikált cikkek alapján)**

Az osztályozási eljárások alkalmazásának leggyakoribb területei:

- a hosszútávú előrejelzések, az agrometeorológiai előrejelzések,
- a klímaszcenáriók készítéséhez alkalmazott leskálázási modellek (Bartholy-Duckstein, 1994, Bartholy-Matyasovszky-Bogárgi, 1995),
- az El Nino jelenség regionális hatásanalíziséhez (Bartholy-Pongrácz, 1998).

Ezen vizsgálatok eredményeire vonatkozó publikációk sora látott már napvilágot.

Kalkstein (1997) nevéhez fűződik egy újszerű rendszerező munka, melyben SSP (Spatial Synoptic Classification–Térbeli szinoptikus osztályozások) elnevezéssel kísérli meg felvázolni a lehetséges szinoptikus közelítésű osztályozásokat. E munkának egy kulcsdiagrammja a 4. ábrán kerül bemutatásra, melyen jól áttekinthetők a térskálák és a hozzá tartozó módszertani megközelítések.




Forrás: Kalkstein, 1997

4. ábra:

Lehetséges szinoptikus közelítésű osztályozások

Szerző nagy súlyt helyez a cirkulációs tipizálás mellett az adott térségre jellemző légtömegek osztályozására. Különösen figyelemreméltó az összesített eloszlásfüggvények maximumhelyek száma szerinti felbontása részgörbékre (feltételezve, hogy mindegyik egy-egy specifikus légtömeghez tartozó hőmérsékleti eloszlásfüggvény).

## *Irodalom*


*Ambrózy P. - Bartholy J. - Gulyás O.*, 1984: A system of seasonal macrocirculation pattems for the Atlantic-European region. Időjárás, Vol. 88. No. 3. pp. 121-133.

*Bartholy J.*, 1989: Determination of seasonal macrosynoptical types using cluster analysis and rotated EOF analysis. Acta Climatologica. Tomus XXI-XXIII. Fasc. 1-4. pp. 23-33.

*Bartholy J. - Duckstein L.* (1994): A subjective macrocirculation classification for the western region of the United States. *Annales Universitatis Scientiarum Budapestinentis de Rolando Eötvös Nominatae, Sectio Geophysica et Meteorologica*, Tomus X. pp. 7-22.

*Bartholy J. - Matyasovszky I. - Bogárdi I.* (1995): Effect of climate change on regional precipitation in Lake Balaton watershed. *Theoretical and Applied Climatology*, Springer Verlag Vol. 51., No. 4., pp. 237-250.

*Bartholy J., Pongrácz, R.*, (1998): Comparing signals of ENSO and NAO for selected regions of the northern hemisphere. *Annales Geophysicae*, Supplement II., Vol. 16. C 366.

*Litynski, J. K.*, 1983: The numerical classification of the world's climates. WMO World Climate Programme, PMC/63, Geneva, p. 51.

*Namias, J.*, 1981: Teleconnections of 700 mb height anomalies for the Northem hemisphere. Calcofi Atlas, No. 29.

*Richman, B. M.*, 1985: Rotation of principal components. (Rewiev paper), Illinois State Water Survey, Climate and Met. Sec. p. 73.

# Bayes - becslés a meteorológiai adatasszimilációban

Dévényi Dezső

### *Bevezetés*

A meteorológiai adatasszimiláció célja a különböző időpontokban beérkezett és különböző időpontokra, különböző térbeli pontokra, valamint különböző meteorológiai állapothatározókra vonatkozó mérési adatokból a numerikus előrejelző modellek számára a lehető legpontosabb és kiegyensúlyozott kezdeti feltételek biztosítása. A feladat megoldását számos tényező nehezíti. Ezek közül az egyik legjelentősebb az, hogy az adatok által biztosított térbeli felbontás jelentősen rosszabb a modellek által ma már elért és megkövetelt felbontásnál, azaz a mérések jelentősen alul- mintavételezik a meteorológiai folyamatokat. Bizonyos új és kevésbé új mérőeszközök, mint pl. a meteorológiai műholdak és az időjárási radarok megoldást kínálnak erre a problémára, ugyanakkor ezek az eszközök általában nem a numerikus prognosztikai rendszerekben szükséges állapothatározókat mérik, hanem azokkal bonyolult relációban álló, gyakran térben vagy időben integrált mennyiségeket mérnek. Mivel az adatok jelentős része ma már nem kapcsolódik a standard fő- és mellék mérési terminusokhoz, ezért olyan technikákra is szükség van, amelyek segítségével lehetséges egy adott idő-intervallum mérési adatainak felhasználása a kiindulási feltételek meghatározásában. Az eddig vizsgált különböző eljárások felölelik a 3- és 4-dimenziós variációs analízis módszereket, ill. a Kálmán-filter alkalmazását.

Az elmúlt három évtizedben a meteorológiai adatasszimilációban, ill. objektív analízisben az optimális interpoláció módszere (Gandin, 1963) vált uralkodóvá. Az optimális interpoláció azonban nemcsak hatásos analízis módszernek bizonyult, hanem kikényszerítette és megnyitotta az utat az előrejelzési és mérési hibák statisztikai szerkezetének tanulmányozásához, az optimális állomáshálózatok tervezéséhez és közvetett módon, a meteorológiai folyamatok érzékenységének és előrejelezhetőségének vizsgálatához. Kezdetben különösen motiválta ezeket az irányzatokat az az álláspont, hogy a numerikus előrejelzések javításához elsősorban a kezdeti feltételeket kell javítani, mivel azok annyira rosszak, hogy a legkisebb javítás is érezhető javulást eredményez az egyre finomabb felbontású modellekkel generált előrejelzések minőségében. Ma már sokkal kiegyensúlyozottabb álláspont uralkodik, mivel az elmúlt 5 év egyik fontos irányváltása volt a modell dinamika helyett a modellekben a finomabb skálájú folyamatok parametrizációjának javítására történő erőkoncentráció. Ez a megkezdett munka várhatóan további lendülettel folytatódik, mivel a talajfolyamatok, a konvekció, a légköri turbulencia, stb. jobb leírására irányuló törekvések ráirányították a figyelmet az említett területeken meglévő tudatlanságunkra. Ugyanakkor ez a fejlődés megköveteli olyan változók mérését és az adatasszimilációba való bevonását, amelyekre eddig nem került sor (pl. talajnedvesség, vegetációs adatok, stb.).

Mivel az optimális interpoláció módszerében együtt kapnak helyet a modell-előrejelzésekből származó háttér információk (szokás ezeket első közelítésnek is nevezni) és a mérésekből származó adatok, ezért a modellhibák és a kapcsolódó érzékenységi és előrejelezhetőségi problémák ma már nem különülnek el az adatok hibáinak és az adatoknak az analízisre gyakorolt hatásának vizsgálatától. A legáltalánosabb matematikai keretet itt parciális sztochasztikus differenciál egyenletek rendszere jelenti (ami a légkördinamikát és -fizikát írja le), amelyet a kvázi-folyamatosan (vagy csak diszkréten) beérkező mérési adatok sztochasztikus folyamata vezérel. A feladat természetesen az optimális vezérlés meghatározása.

A Bayes-becslés módszerei, ill. annak az inverz feladatok megoldására kidolgozott változatai egységes fogalmi rendszert és metodikát kínálnak a meteorológiai adatasszimiláció fentebb bemutatott problémájának megoldására ás az optimális interpoláció módszerén való túllépésre. A kérdésnek ma már jelentős irodalma van (a legutóbbi idők fejleményeit illetően lásd WMO, 1997) és néhány gyakorlati eredmény is elérhető. Ugyanakkor még mindig nem eléggé és széles körben ismertek azok az általános elvek, amelyek ezt a megközelítést jellemzik. Ezért a jelen dolgozat első fejezetében összefoglaljuk azokat az alapvető gondolatokat és eszközöket, amelyek a Bayes-féle módszert jellemzik. A második fejezetben formulázzuk meg a Bayes-féle becslést használó általános adatasszimilációs módszert. A harmadik fejezetben mutatunk be egy működő 3-dimenziós variációs analízis módszert, amit a NOAA Forecast Systems Laboratory-ban (FSL) fejlesztettünk ki. Végül összefoglaljuk eredményeinket és utalunk a további fejlődés irányaira.

Ezt az írást Dr. Gulyás Ottó emlékének ajánlom. Visszapillantva a múltra talán érdemes ott kezdeni, hogy a 70-es évek elején szellemi pezsdülés indult meg az Országos Meteorológiai Szolgálatnál (OMSz). Fiatal szakemberek érkeztek a Szolgálathoz és, talán a Szolgálat történetében először, matematikusokkal is kialakult szoros szakmai együttműködés. Az együttműködő matematikusok között volt Gulyás Ottó, aki akkoriban a Távközlési Kutató Intézetben (TÁKI) volt osztályvezető és nagyon aktívan dolgozott a matematikai statisztika alkalmazásai terén. Később átjött az OMSz-ba és körülötte egy iskola alakult ki. Mint fiatal kezdő szakember én is hamar kapcsolatba kerültem Ottóval, aki könnyen megtalálta a hangot a fiatalokkal és akik szívesen dolgoztak vele. Bár soha nem dolgoztam közös témán Ottóval, azért ő bevont az OMSzos továbbképzésekbe, az egyetemi oktatásba és nagy vállalkozásként elkészült egy közös tankönyv (Dévényi és Gulyás, 1988), amit abban az időben az egész meteorológiai tankönyv irodalmat tekintve is úttörő munkának lehetett tekinteni. Én nagyon élveztem a könyvön együtt végzett munkát, pedig 2 teljes éven keresztül elég kemény elfoglaltságot jelentett és a munka nem volt mentes a vitáktól sem. Ezek a viták azonban egyáltalán nem jelentettek törést a kapcsolatunkban, sőt, inkább fokoztak a kölcsönös megbecsülést. Sokat terveztük a könyv angol nyelvű kiadását és úgyszintén egy második kötetet haladóbb témákra és további gyakorlati alkalmazásokra. Sajnos ezek a szép tervek nem valósulhattak meg Ottó korai halála miatt. Kíváncsi vagyok, hogy Ottó hogyan értékelné ezt a mostani dolgozatot. Egyes részeknél bizonyos vagyok az egyetértésében, mert azokat a dolgokat tőle tanultam.

## *1. A Bayes-becslés*

Ebben a fejezetben egy rövid bevezetést kívánunk adni a Bayes-becslés gondolatvilágába. Nem szándékozunk technikai részletekbe belemenni. Azokat illetően lásd pl. Dévényi és Gulyás (1988) vagy Box és Tiao (1973), avagy Sivia (1996). Egy egyszerű példán keresztül illusztrálni fogjuk a bemutatott elveket.

Érdemes felidézni, hogy maga a Bayes-tétel, ami a Bayes-becslés elmélete mögött áll, egyike a valószínűségszámítás egyszerűbb tételeinek és lényegében minden bevezető valószínűségszámítási tankönyvben megtalálható. A szerzők többsége nem szokta elfelejteni azt a standard kifejezést sem leírni, hogy a Bayes-tétel lényegében egy tautológia. Ugyanakkor a Bayes-féle megközelítés általában heves viták tárgya és a Bayes-becslési módszerek alkalmazása igen hullámzó népszerűséget mutat az egyes időszakokban. Az uralkodó divattól a szinte teljes elfeledésig mindent felölel a népszerűségi skála. A "népszerűségi mutató" ilyen ingadozását nem túl könnyű megmagyarázni, de bizonyosan nagy szerepet játszik benne, hogy a Bayes-becslésben használt fogalmakat, elsősorban az un. a'priori ismeret vagy a'priori eloszlás fogalmát sokszor nem értelmezik tisztán.

A matematikai gondolkodás alapvető formája a deduktív következtetés, amikor axiomákból és definíciókból kiindulva egész elméleteket lehet felépíteni. Nem így működnek azonban a természettudományok, ahol leginkább egy állandó "iteratív javítási folyamatról" beszélhetünk: a gyűjtött adatok (gyakran statisztikai) vizsgálaton esnek át, majd az így szerzett ismeretek alapján újabb adatgyűjtésre kerül sor, vagy másképpen fogalmazva: a tudományos hipotézis és a kísérlet állandó ismétlésével és korrekciójával fejlődik a tudományos ismeret. A Bayes-féle elmélet ideális keretet nyújt ezen állandó iteráció formalizált leírásához. Viszont kemény és sokszor személyeskedő vitákra adnak lehetőséget az a'priori információ (tudás) körüli viták. Természettudományos, fizikusi körökben az utóbbi időben a Bayes-becslés jelentős hívőtáborra lelt és a fejlemények egy része a WEB-en is jól követhető (lásd pl. http://bayes.wustl.edu/, ahol a közelmúltban elhunyt E. T. Janes professzor Bayes-i szellemben írt tankönyvének elkészült része is elérhető). Egy kicsit részletesebben megvilágítani a Bayes elmélet hátterét, tegyük fel, hogy valamilyen $y = (y_1, y_2, \ldots, y_n)$ adathalmaz (megfigyelések) eloszlási sűrűsége az $f(y|x)$ feltételes sűrűség, és a feltételes sűrűségben szereplő $x = (x_1, x_2, \ldots, x_k)$ paraméter-együttesről tudjuk, hogy előzetes vagy a'priori sűrűsége $f(x)$. Tehát az $f(x)$ sűrűség azt fejezi ki, hogy mit tudunk az x-ről a megfigyelési adatok nélkül. A megfigyelési adatok birtokában a Bayes-tétel az alábbi formában írható fel:

$$f(x|y) = \frac{f(y|x)f(x)}{f(y)} \tag{1}$$

Végülis (1)-ből az $f(x|y)$ feltételes sűrűség azt mondja meg, hogy az adatok birtokában mennyiben változik az x paraméter, vagyis az új adatok fényében milyen új értékeket kapunk a paramétereinkre. Szokás ezt a feltételes sűrűséget a'posteriori feltételes sűrűségnek is nevezni. Az adatok birtokában (1)-ben az $f(y|x)$ sűrűséget lehet az x függvényének is tekinteni (y helyett) és ebben az esetben az x likelihood függvényének nevezzük az adott y-ra és az $L(x \mid y)$ jelölést alkalmazzuk. Ezekkel a módosításokkal a Bayes-tétel új alakja

$$f(x|y) = L(x|y)f(x) \tag{2}$$

azaz az a'posteriori sűrűség (eloszlás) arányos a likelihood es az a'priori sűrűség (eloszlás) szorzatával. A likelihood függvény igen fontos szerepet játszik: rajta keresztül módosítják az előzetes ismeretet az adatok. Egyszerű meggondolással látható, hogy a likelihood függvény csak egy szorzótényező erejéig van meghatározva; a likelihood függvénynek csak a relatív értéke fontos.

Amennyiben a Bayes-tétel segítségével sikerül meghatározni az a'posteriori eloszlást, különböző kérdések tehetők fel az így eredményezett új változó statisztikai tulajdonságaira (mi a várható értéke, szórása, a maximális valószínűséghez tartozó értéke, stb.) vonatkozóan. Már most óvjuk viszont az olvasót attól, hogy a Bayes-tételt egy minden problémára orvosságot adó csodaszernek tekintse. Az elméletileg egyszerű megfontolások sokszor eredményeznek a gyakorlatban kiszámíthatatlan eloszlásokat és ilyen esetekben további ügyeskedéseket kell alkalmazni. Viszont a Bayes-tétellel tárgyalt kérdésekkel nagyon gyakran találkozunk az adatasszimilációban, ahol egy háttér (előzetes) információ áll rendelkezésünkre és felvetődik az a kérdés, hogy új mérések bevonásával mi az optimális kombinációja a háttér (előzetes) értékeknek és a megfigyeléseknek (feltételezve, természetesen, hogy mindegyik hibával terhelt). Az alábbi egyszerű példában egyetlen térbeli pontra fogunk objektív analízist végezni úgy, hogy az adott pontban rendelkezésre áll a háttér és a mért érték (tehát, szemben a fenti esettel, ahol vektorváltozókkal dolgoztunk, itt egy egyváltozós esetet mutatunk be). A szereplő eloszlásokat normális eloszlásnak fogjuk venni és ezzel egy könnyen (analitikusan) számolható esethez jutunk. Az irodalomban elfogadott szokásnak megfelelően, $y_0$ jelöli az adott pontbeli megfigyelést, és $x_b$ az adott pontbeli háttér (background) információt. A feladat természetesen az, hogy az (1) Bayes-formulában szereplő tagokra a megfelelő sűrűségeket beírjuk.

Egy nagyon természetesen előzetes ismeretnek tekinthető az, hogy az x változónk közel van az $x_b$ előzetes értékhez és az eltérés normális eloszlást követ $x_b$ várható értékkel és $\sigma_b$ szórással:

$$f(x) = \mathcal{N}(x|x_b, \sigma_b) = \frac{1}{\sqrt{2\pi\sigma_b}} \exp\left[-\frac{1}{2}\frac{(x - x_b)^2}{\sigma_b}\right] \tag{3}$$

ahol $\mathcal{N}$ a normális eloszlásra utal és a többi jelölés általánosan ismert. Ha az y mérés közvetlenül az x-re vonatkozik és a mérés varianciája $\sigma_0$, akkor az x-re vonatkozó feltételes sűrűség

$$f(y_0|x) = \mathcal{N}(y_0|x, \sigma_0) = \frac{1}{\sqrt{2\pi\sigma_0}} \exp\left[-\frac{1}{2}\frac{(y_0 - x)^2}{\sigma_0}\right] \tag{4}$$

(4)-et x-szerint integrálva:

$$f(y_0) = \int p(y_0|x)p(x)dx \tag{5}$$

Most kihasználjuk, hogy normális eloszlásokról van szó:

$$p(y_0) = \mathcal{N}(y_0|x_b, \sigma_b + \sigma_0) \tag{6}$$

Most az összes származtatott sűrűséget írjuk be a Bayes-tétel (1) alakjába és végezzük el a számításokat. Ekkor

$$p(x|y_0) = N(x|x_{ered}, \sigma_{ered}) \tag{7}$$

ahol

$$\frac{x_{ered}}{\sigma_{ered}} = \frac{y_0}{\sigma_0} + \frac{x_b}{\sigma_b} \tag{8}$$

és

$$\frac{1}{\sigma_{ered}} = \frac{1}{\sigma_0} + \frac{1}{\sigma_b} \tag{9}$$

Érdemes megjegyezni, hogy ezt az eredményt már Gauss is ismerte.

Az előző eredményhez azonban úgy is eljuthatunk, hogy vesszük a Bayes-tétel logaritmusát:

$$-\ln[f(x|y_0)] = -\ln[f(y_0|x)] - \ln[f(x)] + \mathrm{const} \tag{10}$$

és beírjuk a szereplő sűrűségeket:

$$-\ln[\mathbf{f}(\mathbf{x}|\mathbf{y}_0)] = \frac{1}{2}\frac{(\mathbf{x}_b - \mathbf{x})^2}{\sigma_b} + \frac{1}{2}\frac{(\mathbf{y}_0 - \mathbf{x})^2}{\sigma_0} + \mathbf{const} \tag{11}$$

Ha most meghatározzuk ezen kifejezés minimumát (azaz maximalizáljuk a valószínűséget), akkor az előző eredményt ((8) és (9) összefüggések) kapjuk vissza.

A (11) jobboldalát tekinthetjük egy ún. költség- vagy veszteségfüggvénynek, amelynek minimalizálásával meghatározhatjuk az x optimális értéket. Ezt az értéket tekinthetjük aztán az objektív analízis eredményének. Funkcionálok szélsőértékeire (minimumara) vonatkozó feladatokkal a variaciószámítás foglalkozik és az első variációs objektív analízis módszer kidolgozása Sasaki (1955) nevéhez fűződik. Ő azonban nem kapcsolta a variációs funkcionálban szereplő paramétereket a feladat eloszlási paramétereihez, így módszere sok tekintetben formálisnak tekinthető. Azonban ez az első próbálkozás később nagyon hasznosnak bizonyult a Bayes-féle módszeren alapuló objektív analízis módszerek fejlesztésekor.

A bemutatott példában a megfigyelés hibáját normálás eloszlásúnak tekintettük, ugyanakkor az operatív meteorológiai gyakorlat azt mutatja, hogy a beérkező adatok bizonyos százaléka erősen hibás. Ezért a fenti módszer gyakorlati alkalmazása során érdemes azt feltételezni, hogy a megfigyelési hiba két eloszlás: egy normális eloszlás (a "jó" adatok hibaira) és egy egyenletes eloszlás (a "rossz" adatok hibaira) keveréke. Ebben az esetben a veszteségfüggvény viszont már nem lesz kvadratikus és így több lokális szélsőértékhelye lehetséges. Megfelelően robusztus minimalizációs módszert célszerű alkalmazni, hogy a háttér (background) értékhez legközelebbi lokális szélsőértékhez jussunk. Az előző példa abban is ideális volt, hogy a mérési és analízis pont egybeesett, ill. a mérés maga közvetlenül az x-re vonatkozott. A következő fejezetben látni fogjuk, hogyan kell eljárni az ettől különböző esetekben.

## 2. Az általános adatasszimilációs módszer

Mint ismeretes, az adatasszimiláció (vagy szűkebben: az objektív analízis) során a mérési adatokból kell előállítani a numerikus prognosztikai modellek számára pontos és kellően kiegyensúlyozott kezdeti feltételeket. Mivel a modelleket általában rácshálózatokon oldjuk meg (a számítások egy része a spektrális modellek esetében is a rácshálózaton történik), ezért általában véges (de igen nagy; általában több millió) dimenziójú vektor valószínűségi változóra vonatkozó becslésről van szó (a sztochasztikus mező esetét illetően lásd pl. Bennett (1992) és Christakos (1992)). A feladat megoldását alapvetően az nehezíti, hogy általában a megfigyelések terének dimenziója jóval alacsonyabb a rácspont (modell) tér dimenziójánál és így a feladat reménytelenül alulhatározott. (Csak az érdekesség kedvéért jegyezzük meg, hogy bár az objektív analízis feladata általában alulhatározott, bizonyos térrészekben és bizonyos meteorológiai elemekre lehet túlhatározott is.) A feladat egyúttal egy inverz feladat is, aminek keretében a mérésekből következtetünk vissza a modell-állapotra (az analízis során lényegében egy alkalmas modell-állapot előállítására törekszünk, ami egyúttal kiindulási állapot a modell további integrálásához). Inverz feladatok nem feltétlenül rosszul meghatározottak, de az objektív analízis vagy adatasszimiláció esetén bizonyosan azok. Hogy jobban illusztráljuk a problémát, az alábbiakban néhány olyan kérdést teszünk fel, amelyek a probléma néhány alapvető vonását mutatják (a példák egyáltalán nem törekednek teljességre; továbbiak minden nehézség nélkül hozhatók a különböző mérési szakterületekről). Megfelelő magyar terminológia híján, forward modellnek nevezzük a leképezést a modell térből a megfigyelési térbe vagy térre.

*Rosszul megfogalmazott kérdések (példák inverz feladatokra):*

1A. A modell tér egy rácscellájában ismerünk egy mérési adatot (pl. egy nyomási értéket). Mi a nyomás értéke a rácscella csúcsaiban?

Válasz: ?

2A. Műholdon mértük egy légoszlop fölött a kilépő sugárzás spektrális intenzitás eloszlását. Mi a hőmérséklet és légnedvesség eloszlása az adott légoszlopban?

Válasz: ?

3A. Global Positioning System (GPS) adatokból rendelkezésre áll egy légoszlop kihullható vízmennyisége. Mi a specifikus nedvesség eloszlása a légoszlopban?

Válasz: ?

*Könnyen megválaszolható kérdések (példák forward modellekre):*

1B. Egy rácscella csúcsaiban ismerjük egy meteorológiai elem értékeit (pl. a nyomásét). Mi lehet ezen elem értéke a rácscella belsejében lévő megfigyelési pontban.

Válasz: Interpoláljuk (pl. lineárisan) a rácsponti értékeket a kérdéses pontba. Ez a legtöbb esetben jó eredményt ad a gyakorlatban.

2B. Ismerjük a hőmérséklet és a légnedvesség eloszlását egy légoszlopban (és néhány felszíni adatot is). Mit mérhet egy meteorológiai műhold kimenő sugárzásként?

Válasz: Vegyünk elő egy jobb sugárzási modellt és helyettesítsük be az ismert adatokat. Az eredmény egy sor gyakorlati problémában alkalmazható.

3B. Ismerjük a specifikus nedvesség eloszlását egy légoszlopban. Mit mérhettek a GPS műholdak kihullható vízmennyiségként?

Válasz: Integráljuk a specifikus nedvesség profilt a vertikálisban.

*Jól megfogalmazott kérdések (példák inverz feladatok megoldására):*

1C. Egy rácscellában van egy mérési adatunk. Milyen értékeket tekintenénk a rácscella csúcsaiban a legvalószínűbb értékeknek?

Válasz: Azon értékeket, amelyeket a megfigyelési helyre visszainterpolálva a legközelebb kerülnénk a megfigyelési adathoz. Ha több ilyen készlet van, akkor tekintsük ezek közül azt, amelyik a legközelebb van pl. az éghajlathoz, vagy még jobb, az adott időpontra vonatkozó numerikus előrejelzéshez (ezeket a segédadatokat úgy tekintjük, hogy a rácscella csúcsaiban adottak).

2C. Műholdon mértük egy légoszlop fölött a kimenő sugárzás spektrális intenzitás eloszlását. Mi a hőmérséklet és a nedvesség legvalószínűbb eloszlása a légoszlopban?

Válasz: Az a legvalószínűbb eloszlás, amelyikből visszaszámolt sugárzás a legközelebb van a megfigyelthez. Ha több ilyen van, akkor vegyük ezek közül azt, amelyik a legközelebb van valamilyen kisegítő adathoz (pl. előrejelzett vagy éghajlati értékhez).

3C. GPS mérésekből rendelkezésre áll egy kihullható vízmennyiség adat. Mi a specifikus nedvesség legvalószínűbb profilja az adott légoszlopban?

Válasz: Az, amelyikből visszaszámolt kihullható vízmennyiség a legközelebb van a mért értékhez. Ha több ilyen van, érdemes az adott időpontra vonatkozó numerikus előrejelzéshez legközelebbit venni.

A fenti kérdések feltevésénél és a válaszok adásánál nem törekedtünk igazán pontosságra, lényegében csak egy minőségi képet akartunk felvázolni. A konkrét feladatok esetén a rendelkezésre álló mérési adatok, háttér információk (pl. klíma vagy numerikus előrejelzés) és a forward modell(ek) minősége dönti el, hogy a feladat a fenti keretben a gyakorlat számára elégséges pontossággal megoldható-e. A bemutatott három példában például ma már lehetséges a felhasználók bizonyos köre számára kielégítő pontosságú feladatmegoldás. A forward modellek fejlesztése és hibáik meghatározása az adott terület specialistáinak lenne a feladata. Igen komoly elmaradások vannak ezen a téren és egy sor észlelő eszközre a forward modellek nem kellő minőségűek és hibáikról a legtöbb esetben semmilyen információ nem áll rendelkezésre. Más esetekben léteznek forward modellek, de azok számítástechnikailag annyira költségesek, hogy lehetetlen az alkalmazásuk egy operatív rendszerben. Így pl. a sugárzásterjedésre léteznek az ún. "vonalas" modellek, de egy ilyen futtatása több erőforrást igényel, mint egy numerikus prognosztikai modellé. Nagy elmaradás van a radarokra vonatkozó forward modellek fejlesztésében, pedig a radar szolgáltathatna nagy mennyiségű információt a mezoskálájú numerikus modellek számára. Ugyanakkor pozitív példaként tekinthető és az optimális interpoláció módszerének világméretű elterjedéséhez döntő módon hozzájárult, hogy a korábbi éghajlati mezőkről áttértek a numerikusan előrejelzett mezőkre, mint háttér információra (Rutherford, 1972).

Természetesen az előzőekben inverz feladatok megoldására bemutatott megközelítés teljes mértékben formalizálható a Bayes-féle elmélet keretében. Ezt az eljárást nem visszük végig, de megtalálható mind az alapvető geofizikai (pl. Tarantola and Valette, 1982) és meteorológiai szakirodalomban (pl. Lorenc, 1986 és WMO, 1997). A normális eloszlásokat feltételező esetben a minimalizálandó functionál az alábbi módon írható fel:

$$J(x)=\frac{1}{2}\|Q^{-1}(x-x_b)\|^2+\frac{1}{2}\|R^{-1}(y_0-F(x))\|^2 \tag{12}$$

ahol x - a keresett analízis érték változója (vektor),

$x_b$ - az x-re vonatkozó háttér információ,

$Q^{(-1)}$ - a háttér információs precíziós (pontossági) operátora

(a háttér információ kovariancia mátrixának inverze),

$y_0$ - a megfigyelések vektora,

F - a forward operátor,

$R^{(-1)}$ - a megfigyelések és a forward modell pontossági operátora.

A feladat - természetesen - (12) minimalizálása. A minimumra vonatkozó feltételből közvetlenül felírható - lineáris F esetén - az analízisre vonatkozó megoldás (ami a legjobb lineáris torzítatlan becslés):

$$x_0 = x_b + K(y_0 - Fx_b) \tag{13}$$

ahol K az un. innovácios mátrix és

$$\mathbf{K} = \mathbf{QF}^T(\mathbf{FQF}^T + \mathbf{R})^{-1} \tag{14}$$

vagy

$$\mathbf{K} = (\mathbf{Q}^{-1} + \mathbf{F}^T\mathbf{R}^{-1}\mathbf{F})^{-1}\mathbf{F}^T\mathbf{R}^{-1} \tag{15}$$

Nem nehéz észrevenni, hogy az innovációs mátrixra adott két különböző felírás két különböző módszert sugall a feladat megoldására. Be lehet látni, hogy ezek egymás Lagrange-i duáljai.

A fenti összefüggések alapján nem nehéz felírni az analízis hibájának a kovariancia mátrixát sem:

$$\mathbf{A} = (\mathbf{I} - \mathbf{KF})\mathbf{Q} = \left(\mathbf{Q}^{-1} + \mathbf{F}^T\mathbf{R}^{-1}\mathbf{F}\right)^{-1} \tag{16}$$

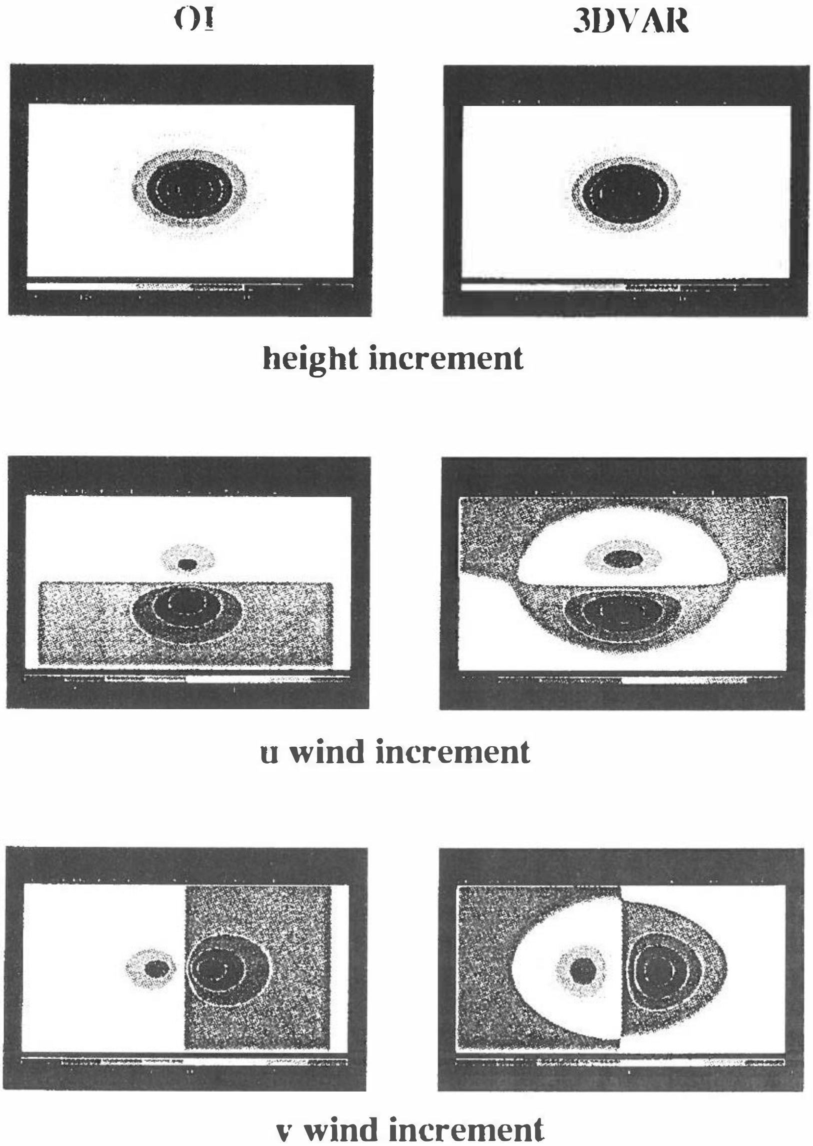
Azonban a (13)-16) egyenletek alapján látható egyszerűség igencsak látszólagos: az "ördög a részletekben van". Abban az esetben, amikor csak egyetlen időpont adatai alapján végzünk analízist és az F lényegében egy lineáris interpolációs operátor (a konvencionális adatokra) vagy egy valamivel bonyolultabb, de linearizálható operátor (műholdas és radar adatokra), akkor az analízis módszert 3-dimenziós variációs analízis módszernek nevezzük. Ez a séma most rendkívül gyorsan terjed az operatív alkalmazásban. Abban az esetben, ha az idő koordináta is szerepet játszik az analízisben, azaz ha egy adott időszak adatai kerülnek felhasználásra az analízisben és ha az abszolút pontosnak feltételezett numerikus előrejelző módszer is szerepel a forward operátorok között, akkor 4-dimenziós variációs analízisről szoktunk beszélni. Ha pedig a modell hibájának a változásait is szerepeltetjük az időbeli analízis sémában, akkor az ismert Kálmán szűrőhöz jutunk (Jazwinsky, 1970). Közbeeső megoldásként (az előrejelzési hiba kovariancia mátrixának időbeli előrejelzése helyett) adaptív szűrőkkel is folynak kísérletek (lásd pl. Hoang et al., 1997). Viszont már a legegyszerűbb 3-dimenziós variációs módszer esetén jelentős nehézségeket kell leküzdeni a feladat hatalmas méretei, a megoldandó egyenletek gyenge kondicionáltsága és az adathibák okozta komplikációk miatt. A következő fejezetben egy real time környezetben működő 3-dimenziós variációs módszert mutatunk be.

## *3. Egy 3-dimenziós variációs analízis módszer*

Az FSL egyike a National Oceanic and Atmospheric Administration (NOAA) kutató laboratóriumainak. A laboratórium fő feladata a tudományos eredmények vizsgálata és esetleges átvitele a gyakorlatba az operatív időjárási és oceanológiai előrejelzések javítása érdekében. A laboratórium egyik kiemelt témája a Mesoscale Analysis and Prediction System (MAPS) numerikus előrejelző rendszer fejlesztése. A MAPS újabb és újabb változatai az FSL-ben történő hosszabb idejű kipróbálás és tesztelés után operatív alkalmazásra kerülnek a National Centers for Enviromental Prediction (NCEP)-ben Rapid Update Cycle (RUC) néven. A MAPS/RUC rendszer nem más, mint egy olyan numerikus előrejelző rendszer, ahol óránként kerül sor analízisek készítésére és a ciklus fenntartásához szükséges 1-órás előrejelzések mellett 3-óránként készül egy 12-órás előrejelzés (az NCEP-ben), ill. 3-óránként készül egy 36 órás előrejelzés (az FSL-ben). Ez egyúttal azt is jelzi, hogy a MAPS/RUC különböző változatai real time alkalmazásban vannak az FSL-ben is, de a National Weather Service (NWS) szabályainak megfelelően a hivatalos előrejelzések készítése csak az NCEP-ben történhet.

A MAPS/RUC rendszer legnagyobb finanszírozója és felhasználója az USA repülési hatósága (Federal Aviation Administration) és a fő felhasználási terület az un. air traffic management, azaz az egyre telítettebb légkör optimális elosztása, különös tekintettel az időjárási viszonyokra. A rendszer másik fő felhasználási területe a nowcasting, azaz a veszélyes időjárási események analízise és rövid időtartamra szóló előrejelzése.

A rendszernek az FSL-en és az NCEP-en kívül egy példánya működik a Wisconsini Egyetemen, ahol műholdas adatok asszimilációjával kísérleteznek segítségével. Egy másik példányát a rendszernek most telepíti a Deutches Zentrum für Luft- und Raumfahrt (DLR, Oberpfaffenhofen, Germany), repülést segítő célokra.




1. ábra

Magassági inkrementum által generált szél inkrementumok: keresztkorrelációs függvény alkalmazásával az optimális interpoláció módszerében (első oszlop) és relaxált geosztrofikus kényszer segítségével a variációs analízisben (második oszlop

A MAPS/RUC rendszernek számos változata létezik. Az éppen aktuális változat részletes leírása megtalálható a következő címen: http://maps.fsl.noaa.gov. Ugyanott elérhetők az aktuális futtatásokból készített térképek és további információk a rendszer által használt adatokról, a futtatási rendszerről, a rendszer történetéről és a kapcsolódó irodalomról.

A 3-dimenziós variációs analízis módszer fejlesztésének alapvető célja, hogy megnyissuk az utat a távérzékelési eszközökkel mért adatoknak a mezoskálájú analízisben és numerikus előrejelzésben történő felhasználásához. Első lépésként a cél egy olyan analízis módszer kidolgozása, amivel legalább elérhető a jelenleg is alkalmazásban lévő optimális interpolációs módszer pontossága és kiegyensúlyozottsága. Az alábbiakban csak egy rövid leírást adunk az FSL-ben kidolgozott variációs módszerről, a további részleteket illetően lásd Dévényi (1996) és Dévényi and Benjamin (1998).

A minimalizálandó funkcionál ebben az esetben

$$J(x)=\frac{1}{2}(x_b-x)^T B^{-1}(x_b-x)+\frac{1}{2}(y_0-Hx)^T(O+F)^{-1}(y_0-Hx) \qquad (17)$$

alakú, ami - természetesen - nagyon hasonló a (12)-ben adott egyenlethez.

Az analízis változók (az **x**-ek):

- magasság,
- szélmező,
- virtuális potenciális hőmérséklet,
- kondenzációs nyomás (szokás emelési kondenzációs szintnek is nevezni).

A kontrol változók (amik szerint minimalizálunk):

- magasság,
- áramfüggvény, 7
- sebességi potenciál,
- virtuális potenciális hőmérséklet,
- kondenzációs nyomás.

Az analízist a megfigyelési inkrementumokra (megfigyelés mínusz háttér információ a megfigyelési pontban) végezzük és eredményül az analízis inkrementumot (az analízis eltérése a háttér információtól) kapjuk. Az analízis a MAPS/RUC rendszerben használt hibrid (sigma és virtuális potenciális hőmérséklet) koordináta rendszer által meghatározott modell szinteken van végrehajtva. Az analízist több lépésre bontottuk. Az első lépésben a tömeg és a szélmező többváltozós analízisét végezzük. Az eredményezett magassági inkrementumok segítségével pontosítjuk a hőmérsékleti háttér mezőt és az új háttér mező alapján számoljuk a virtuális potenciális hőmérséklet megfigyelési inkrementumait. Ezekre vonatkozóan egyváltozós analízist végzünk. Hasonlóképpen egyváltozós analízis keretében kezeljük a nedvességi változót. Az összes analízis inkrementum birtokában meghatározzuk a végső analízis mezőket, majd ezeket alávetjük a hibrid rendszerben használatos vertikális kiigazítási eljárásnak.

A B (háttér hiba) kovariancia mátrixot közelítő alakban használjuk. Első lépésben elkészítjük a

$$\mathbf{B} = \mathbf{C}\mathbf{C}^{T} \tag{18}$$

felbontást, majd a továbbiakban a **C** mátrixot tovább bontjuk:

$$\mathbf{C} = \mathbf{E}\mathbf{D}_x\mathbf{C}_x\mathbf{D}_y\mathbf{C}_y\mathbf{D}_p\mathbf{C}_p \tag{19}$$

(19)-ben E a háttér hiba szóródásának (jelenleg diagonális) mátrixa, a $D_q$-k skálázó mátrixok, a Cq-k pedig 1-dimenziós digitális (diszkrét) filterek, amelyek általános alakja analitikus formában:

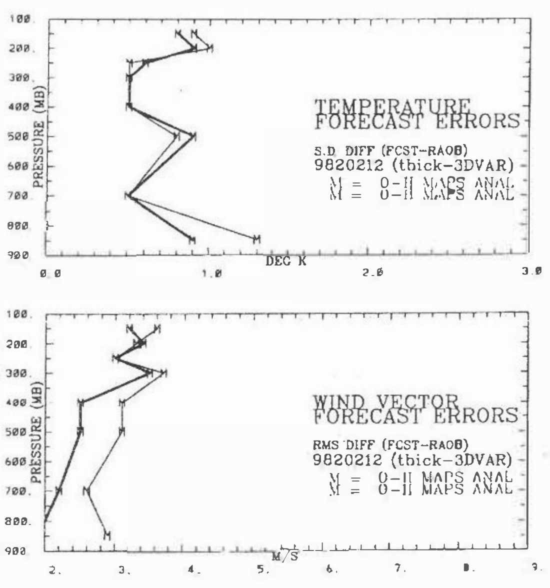
$$\mathbf{C}_q = \left[1 - \alpha\frac{\partial^2}{\partial q^2} + \beta\frac{\partial^4}{\partial q^4} - \gamma\frac{\partial^6}{\partial q^6} + \ldots\right]^{-1} \tag{20}$$

A (17) funkcionál minimalizálása a konjugált gradiens módszert használjuk a Polak-Ribiere formában. Ennek előnye, hogy működik arra az esetre is, amikor a költségfüggvény nem kvadratikus. Ilyen eset áll elő akkor, amikor a költségfüggvény megfigyelési részt tartalmazó tagját módosítjuk a durva megfigyelési hibák kiszűrése céljából és ezzel egyesítjük az adatellenőrzési és analízis fázisokat egyetlen lépésben. Prekondicionálisra a C mátrixot használjuk, azaz az analízis változókat megfelelő transzformációnak vetjük alá. Ez lehetővé teszi, hogy a háttér információ súlyát az egység márix adja meg.

Jelenleg ún. relaxált geosztrófikus kényszert alkalmazunk a variációs analízisben, de van lehetőség a relaxált termikus szél összefüggés alkalmazására is (ez utóbbi a jelenlegi algoritmus kisebb átrendezését igényli). Az 1. ábrán mutatjuk be, hogy az optimális interpolációban használt kereszt-korreláció és a variációs analízisben használt direkt relaxált kényszer mit eredményez a szélmező inkrementumon. Ez egy elemi és kötelező ellenőrzés az analízis módszerek fejlesztése során. Az ábrán az első sor a magasság perturbációját mutatja úgy, hogy a baloldali oszlop vonatkozik az optimális interpolációra, a jobboldali pedig a variációs analízisre. Az ábrán azt is láthatjuk, hogy ha mindkét esetben egyforma a magasság perturbációja. A dolog úgy is felfogható, mint egyetlen magassági inkrementum által indukált analízis inkrementum bemutatása. Könnyű belátni, hogy nem alkalmaznánk valamilyen kényszert a magassági (tömeg) és szélmezők között, akkor a magassági adat csak a magassági analízisre lenne hatással és az ábra két alsó sora üresen állna. Viszont megfelelően alkalmazott kényszer hatására szélmező is generalható a magassági adatok alapján, mint az az ábrán bemutatásra került. Meglepő, hogy a kereszt-korrelációval (optimális interpoláció) és a büntetés tagon keresztül (variációs módszer) generált analízis inkrementumok mennyire közel vannak egymáshoz. Ez segíti azt, hogy a variációs analízisből indított előrejelzések közel legyenek az optimális interpolációs analízisből indítottakhoz.

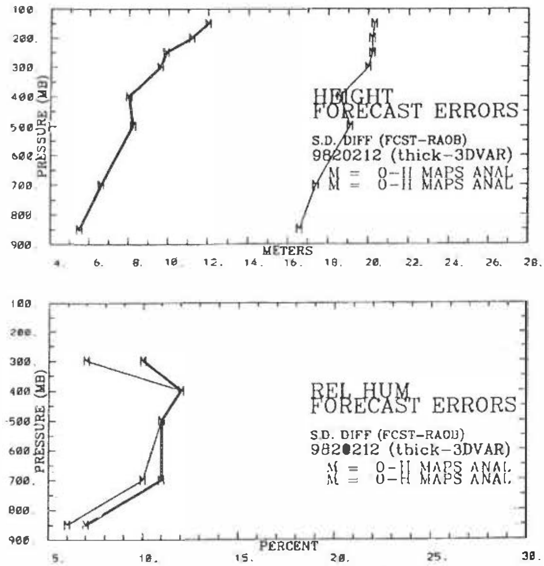
A forward modell a 3-dimenziós analízisben jelenleg lineáris interpoláció. Azonban egyváltozós és egydimenziós variációs analíziseket készítünk kísérleti jelleggel műholdas (GOES) és GPS adatok asszimilálására. Ezek alkamas teszt-terepet biztosítanak a megfelelő forward modellek kipróbálásához.

A következő ábrákon egy tipikus helyzetet mutatunk be az analízis módszerek összehasonlítására. A kísérleteket valós idejű adatokkal és a rendszert 3-órás asszimilációs ciklussal készítettük. A MAPS 60km-es felbontású változatát futtattuk párhuzamosan két különböző analízis módszerrel: az egyik változat az optimális interpolációs analízist tartalmazta, a másik pedig a variációst. Ezen kívül minden más azonos volt, még arra is ügyeltünk, hogy a számításokat ugyanazon a gépen végezzük, nehogy eltérő aritmetikai megoldások az eredményt befolyásolják. Az analízisek időpontja: 1998. júl. 21. 12 UTC. A 2. és 3. ábra az analízisek távolságát mutatja a rádioszondás adatoktól. A vékony vonal vonatkozik az optimális interpolációra, a vastag pedig a variációs módszerre. Látható, hogy túlnyomóan a variációs analízis van közelebb az analízisekhez. Megjegyezzük, hogy itt nem beszélhetünk verifikációról, mivel ugyanezeket a rádioszondás adatokat felhasználtuk az analízis során is, tehát a független mintavétel követelménye nem teljesül. Mindenképpen kívánatos viszont, hogy az analízis jól "feküdjön" az adatokra, különösen az általunk megcélzott nowcasting jellegű felhasználások esetén. A következő két ábrán viszont már verifikáltuk a 12-órás előrejelzéseket. A vékony és vastag vonalak használata itt azonos az előző esettel. Úgy tűnik, hogy a variációs analízisre alapozott 12-órás analízisek valamivel jobbak, mint az optimális interpolációval kapottak. Ennek kapcsán fontos arra emlékezni, hogy az analízisnek az adatokhoz való közelségéből egyáltalán nem következik az előrejelzéseknek a verifikációs adatokhoz való közelsége. Nagy szerepet játszik itt a mezőknek a megfelelő spektrális tulajdonsága és - elsősorban - kiegyensúlyozottsága. Egy analízis rendszert természetesen esetek sokaságán szükséges tesztelni. A rendelkezésünkre álló eddigi statisztikai eredményeket máshol fogjuk a közeljövőben publikálni.



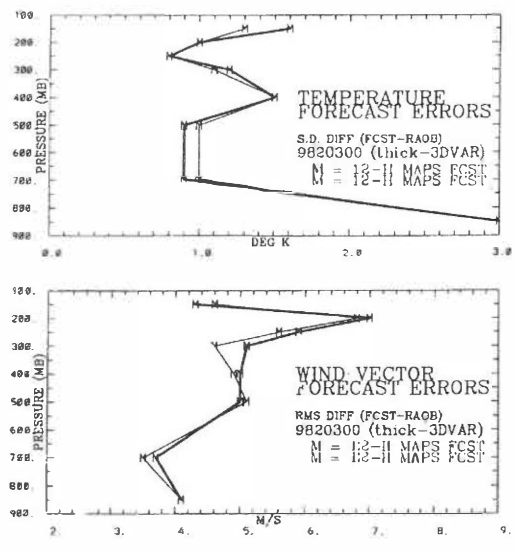

2. ábra

Az optimális interpoláció módszerével (vékony vonal) és a variációs analízissel (vastag vonal) készített hőmérséklet és szél analízisek távolsága az adatoktól



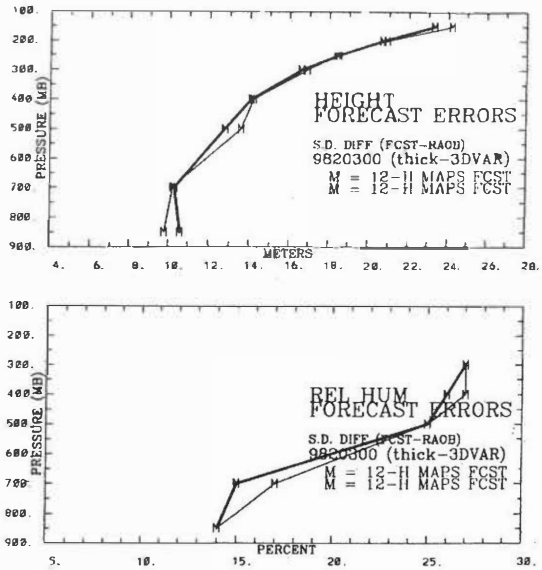

3. ábra

Az optimális interpoláció módszerével (vékony vonal) és a variációs analízissel (vastag vonal) készített magasság és relatív nedvesség analízisek távolsága az adatoktól




4. ábra

Az optimális interpolációs (vékony vonal) ill. a variációs analízis (vastag vonal) alapján futtatott 12 órás előrejelzések verifikációja (magasság és relatív nedvesség).




5. ábra

Az optimális interpolációs (vékony vonal) ill. a variációs analízis (vastag vonal) alapján futtatott 12 órás előrejelzések verifikációja (magasság és relatív nedvesség).

### *Összefoglalás*

A dolgozatban bemutattuk a Bayes-féle becslés alapgondolatát és annak a meteorológiai analízisben való alkalmazhatóságát egy egyszerű példán illusztráltuk. A példa jól mutatta, hogy a Bayes-féle elmélet keretében milyen lehetőség van különböző forrásokból származó információk optimális kombinálására. Az egyszerű példát felhasználtuk annak illusztrálására is, hogy statisztikai becslési problémákat hogyan lehet visszavezetni optimalizációs feladat megoldására.

Bemutattuk a direkt és inverz feladatokat és egy ötletet az inverz feladatok megoldására. Ezek a feladatok nagy szerepet játszanak a modern adatasszimilációs problémák megoldásában, mivel a meteorológiai adatasszimiláció (szűkebben az objektív analízis) tipikusan egy inverz feladat.

Egy működő 3-dimenziós analízis módszerként bemutattuk a MAPS/RUC rendszerre kifejlesztett variációs analízis módszert. A módszer működésének eredményességét egy eseten illusztráltuk.

Összefoglalóan érdemes elmondani, hogy a Bayes-féle elméleten alapuló módszerek adják jelenleg a legjobb keretet az adatasszimilációs módszerek fejlesztéséhez. Bár jelentős kezdeti sikerek vannak, a további fejlődéshez egy sor problémát kell megoldani. Ezek közül a legfontosabbak listája következik alább: - a háttér adatok hibáinak meghatározása, - a háttér adatok hibáinak jó és számítástechnikailag gazdaságos közelítései (pl. digitális szűrök), - nem homogén és nem izotrop háttér hiba-kovarianciák hatékony kezelése (pl. tenzoriális filterek), - hatékony módszerek a Kálmán szűrő esetében a filter divergencia elkerülésére, - megfelelő forward modellek kidolgozása, ill. a létezők tökéletesítése, - a forward modellek hibáinak meghatározása (különösen a numerikus prognosztikai modell hibái és általában a korrelált hibák esete fontos), - megfelelő optimalizációs algoritmusok kidolgozása, - hatékony prekondicionáló módszerek származtatása, - újfajta számítástechnikai megközelítés (pl. párhuzamos programozás).

A fenti problémák egy részének eredményes megoldása már újabb jelentős lökést adhat a Bayes-elméleten alapuló adatasszimilációs módszereknek az operatív meteorológiában történő szélesebb elterjedéséhez.

## Köszönetnyilvánítás

A szerző köszönetét fejezi ki az FSL MAPS csoportjában dolgozó kollégáinak a folyamatos támogatásért és Dr. Jim Pursernek és Dr. Dave Parrishnek az NCEPből a rendkívül hasznos beszélgetésekért és segítségükért.

## Irodalom

*Box, G. E. P., and G. C. Tiao,* 1973: Bayesian Inference in Statistical Analysis. Wiley and Sons, Inc., pp 588.

*Bennett, A. F.,* 1992: Inverse Methods in Physical Oceanography. Cambridge University Press, Cambridge, pp 346.

*Christakos, G.,* 1992: Random Field Models in Earth Sciences. Academic Press, New York, pp. 474.

*Dévényi, D.,* 1996: A 3-diemnsional variational analysis for MAPS/RUC. Proceedings of 11th Conf. on Numerical Weather Prediction, Norfolk, Virginia, August 19-23, Amer. Meteor. Soc., 110-113.

*Dévényi, D., és Gulyás, O.,* 1988: Matematikai Statisztikai Módszerek a Meteorológiában. Tankönyvkiadó, Budapest, 443 old.

*Dévényi, D., and S. G. Benjamin,* 1998: Application of a three-dimensional variational analysis in RUC-2. Proceedings of 12th Conf. on Numerical Weather Prediction, Phoenix, Arizona, 11-16 January, Amer. Meteor. Soc., 37-41.

*Gandin, L. S.,* 1963: Objektivnij Analiz Meteorologicseszkich Polej. Gidrometeoizdat, Leningrad. Angol kiadas: Objective Analysis of Meteorological Fields. Tranlated by Israel Program for Scientific Translations.

*Hoang, S., R. Baraille, O. Talagrand, X. Carton, and P. De Mey,* 1997:Adaptive filtering: application to satellite data assimilation in oceanography. Dyn. Atmos. Oceans, 27, 257-281.

*Jazwinsky, A. H.,* 1970: Stochastic Processes and Filtering Theory. Academic Press, New York, pp. 376.

*Lorenc, A.,* 1986: Analysis methods for numerical weather prediction.Q. J. R. Meteorol. Soc., 112, 1177-1194.

*Sasaki, Y.,* 1955: An objective analysis based on the variational principle. J. Meteorol. Soc. Japan, 33, 262-275.

*Sivia, D. S.,* 1996: Data Analysis. A Bayesian Tutorial. Clarendon Press, Oxford, pp. 189.

*Rutherford, I. D.,* 1972: Data assimilation by statistical interpolation for forecast error fields. J. Atmos. Sci., 29, 809-815.

*Tarantola, A., and B. Valette,* 1982: Inverse problems = Quest for information. J. Geophysics, 50, 159-170.

*WMO,* 1997: Data Assimilation in Meteorology and Oceanography: Theory and Practice. A Collection of Papers Prsented at the WMO Second International Symposium on Assimilation of Observations in Meteorology and Oceanography, 13-17 March 1995, Tokyo, Japan, J. Meteorol. Soc. Japan, 75, No IB, pp. 496.

## Gulyás Ottó publikációi

1964 Stacionárius sztochasztikus folyamatok. *TKI Szemináriumi Közlemények*, 1964.

1967 Some remarks on the sampling theorem. *Proceedings of the of the Fourth Colloquium on Information Theory*, Debrecen, 1967.

1969 A potenciálfüggvényes módszer (összefoglalás). *TKI Szemináriumi Közlemények*, 1969.

1969 Alakfelismerési feladatok. *TKI Szemináriumi Közlemények*, 1969.

1969 Potenciálfüggvényes tanulás program. *TKI Szemináriumi Közlemények*, 1969 (Társszerző: Molnár L.)

1969 Tanuló algoritmusok I. *TKI Intézeti Tanulmány*, 1969 (Társszerzők: Németh J., Rét A., Hoffer A., Esze T., Molnár L., Battistig Gy., Csibi S.)

1970 A potenciálfüggvényes tanuló algoritmus általánosításáról és konvergencia sebességéről. *TKI Kézirat*, 1970.

1970 Об ошибке при усечении ряда в теореме отсчетов. *Proceedings of the Fourth Colloquium on Microwave Communication*, Budapest, 1970.

1970 Tanuló algoritmusok alkalmazása meteorológiai előrejelzésre. *KEI-TKI Intézetközi Tanulmány*, 1970. (Társszerzők: Pápainé Szalay G., Molnár L.)

1970 Tanuló algoritmusok felhasználása meteorológiai előrejelzésre. *VI. Magyar Automatizálási Konferencia Kiadványa*, Budapest, 1970.

1970 Tanuló és feismerő algoritmusok II. *TKI Intézeti Tanulmány*, 1970. (Társszerzők: (Bak M., Balogh B., Battistig Gy., Csibi S., Gábor Gy., Győrfi L., Molnár L., Nagy A., Pápainé Szalay G., Pick R., Rét A.)

1971 A konvektiv aktivitás előrejelzése tanuló algoritmusok fekhasználásával. *Időjárás*, 1971, 3-4. (Társszerzők: Molnár L., Szalay G.)

1971 Elektrokardiogrammok automatikus kiértékelése. *Számítástechnikai és kibernetikai módszerek alkalmazása az orvostudományban és biológiában* konferencia kiadványa, Szeged, 1971. (Társszerzők: Bak J., Kobzos L.)

1971 Mintavételi tételek konvergenciája és sorcsonkítási hibája. *TKI Évkönyve*, 1971

1971 Tanuló-felismerő algoritmusok III. *TKI Intézeti Tanulmány*, 1971. (Társszerzőkkel)

1971 Tanuló-felismerő algoritmusok, kisszámítógépes alkalmazásokra. A *Számítástechnika 71'* konferencia kiadványa, Esztergom, 1971. (Társszerzők: Bak J., Csibi S., Kobzos L., Molnár L.)

1972 An iterative method for estimating regression coefficients. *Paper for the CISM*, Udine, 1972.

1972 EKG regisztrátumok analízisének számítógéppel történő automatizálása. *TKI Szemináriumi Közlemények*, 1972. (Társszerzők: Bak J., Ghyczy K., Lamm Gy.)

1972 Elektrokardiogrammok számítógépes értékelésének néhány kérdése. *Ssámitástechnikai és kibernetikai módszerek alkalmazása az orvostudományban és biológiában* konferencia kiadványa, Szeged, 1972. (Társszerzők: Bak J., Ghyczy K., Lamm Gy.)

1972 Применение алгоритмов обучения в метеорологии для предсказания конвективной активности. *Acta Cybernetica*, 1972. (Társszerzők: Molnár L., Szalay G.)

1972 On extended potential function type learning algorithm and their convergence rate. *Problems of Control and Information Theory*, 1972.

1972 Számítógépes diagnosztikai állomás matematikai és software kérdései. *TKI Évkönyve*, 1972. (Társszerzőkkel)

1973 О методе выделения признаков в распознавания образов. *Számítástudományi Konferencia Közleményei*, Székesfehérvár, 1973. (Társszerző: Faragó T.)

1974 A regresszió típusú extrapoláció, módszer, programok és alkalmazási példák. *OMSZ Meteorológiai Tanulmányok*, 1974, N 2, 1-48. (Társszerző: Faragó T.)

1974 A regresszió típusú extrapoláció. *OMSZ Meteorológiai Tanulmányok*, 1974, N 2, 1-48. (Társszerző: Faragó T.)

1974 A valószínűség sűrűségfüggvény becslése és illesztése. *OMSZ Meteorológiai Tanulmányok*, 1974, N 1, 1-67. (Társszerző: Faragó T.)

1974 A valószínűségi sűrűségfüggvény becslése és illesztése. Módszerek, programok és alkalmazási példák. *OMSZ Meteorológiai Tanulmányok*, 1974, N 3, 1-87. (Társszerző: Faragó T.)

1974 Az alakfelismerés néhány matematikai kérdése és alkalmazása, *Kand. Ért.*, 1974 ,1-127.

1974 Some methods of expansions type feature extraction in pattern recognition. (Az alakfelismerés néhány sorfejtéses lényegkiemelési módszere.) *Proc. of the Prague Symposion on Asymptotic Statistics*, 1974 ,69-87. (with T. Faragó)

1975 Az analógia elvén alapuló prognosztikai módszerek matematikai modellje. *Időjárás*, 1975, 79, N 3, 166-177. (Társszerzők: Faragó T. és Kaba M.)

1975 Jogi problémák vizsgálata és megoldásuk kibernetikai módszerekkel. *ELTE JTE jegyzet*, Budapest, 1975, 166-189.

1975 Kisérletek jogesetek matematikai módszerrel történő elemzésére. *Jogtudományi Közlöny*, XXX., 1975, N 6, 319-328. (Társszerzők: Bárdos, Bárdosné)

1975 Tanuló algoritmusok felhasználása jogesetek számítógépen történő elemzésére. *ELTE JTE jegyzet*, Különkiadvány, Budapest, 1975, 22-38.

1976 A tanulóalgoritmus program rendszer. Beszámoló jelentés Az Országos Meteorológiai Szolgálat adatfeldolgozó és adattároló rendszere alapjainak kialakítása R-50 kategóriájú elektronikus számítógépre, *OMFB tanulmány*, OMSZ, Budapest, 1976, 7-52. (Társszerzők: Bak Miklósné és Szádeczky Kardoss G.)

1976 Az analógia fogalma és felhasználása meteorológiai típusok képzésére. Beszámoló jelentés Az Országos Meteorológiai Szolgálat adatfeldolgozó és adattároló rendszere alapjainak kialakítása R-50 kategóriájú elektronikus számítógépre, *OMFB tanulmány*, OMSZ, Budapest, 1976, 119-144.

1976 Meteorológiai mezők természetes ortogonális sorfejtése. Beszámoló jelentés Az Országos Meteorológiai Szolgálat adatfeldolgozó és adattároló rendszere alapjainak kialakítása R-50 kategóriájú elektronikus számítógépre, *OMFB tanulmány*, OMSZ, Budapest, 1976,53-89. (Társszerző: Bartholy J.)

1977 Az analógia fogalma és felhasználása típusok képzésére I. The concept of analogy and its utilization for forming of types I. *Időjárás*, 1977, 81, N 1, 11-18.

1977 Az analógia fogalma és felhasználása típusok képzésére II. The concept of analogy and its utilization for forming of types II. *Időjárás*, 1977, 81, N 6, 341-351. (Társszerzők: Bartholy J. és Légrády G.)

1977 Meteorológiai mezők sorfejtése típusok átlagai szerint. Beszámoló jelentés "Az Országos Meteorológiai Szolgálat adatfeldolgozó és adat tároló rendszere alapjainak kialakítása R-50 kategóriájú elektronikus számítógépre" *OMFB tanulmány*, OMSZ, Budapest, 1977, II. kötet, 74-116. (Társszerző: Bartholy J. és Kaba M.)

1977 Mintavételi tétel homogén, izotrop mezőkre és a sorcsonkítási hiba vizsgálata. Program a nyomásmezők korrelációs viszonyainak elemzéséhez. Az Országos Meteorológiai Szolgálat adatfeldolgozó és adattároló rendszere alapjainak kialakítása R-50 kategóriájú elektronikus számítógépre, *OMFB tanulmány*, OMSZ, Budapest, 1977. II. kötet, 117-131, 132-143. (Társszerző: Légrády G.)

1978 Valószínűségszámítás és matematikai statisztika. Egyetemi jegyzet, *Tankönyvkiadó*, Budapest, 1978.

1978 Meteorológiai mezők számítógépes analízise. *Pályázati tanulmány* az MTA 1978. évi kutatási jutalmára, 1-129. (Társszerzők: Bartholy J., Kaba M. és Légrády G.)

1978 Statisztikus alakfelismerés. Fejezetek a matematikai statisztika alkalmazásaiból, *BJMT jegyzet* [IV. fejezet], 1978. (Társszerző: Révész P.)

1978 Valószínűségszámítás és matematikai statisztika. *Egyetemi jegyzet, Tankönyvkiadó*, Budapest,1978, 1-168.

1979 Az analógia elvén alapuló prognosztikai módszerek matematikai modellje. *Időjárás*, 1979, N 3, 166-177. (Társszerzők: Kaba M. és Faragó T.)

1979 Short range forecast of precipitation with the aid of learning algorithms. (A csapadék rövidtávú előrejelzése tanuló algoritmusok segítségével.) *Időjárás*, 1979, 83, N 2, 101-106. *Meteorológiai előrejelzések*. Meteorológiai Tudományos Napok '78, *OMSZ Hiv. Kiadv.*, XLVIII. kötet, Budapest, 1979, 80-87. (Társszerzők: Bak J. és Tänczer T.)

1979 Terresztikus hatások távprognosztikai alkalmazhatóságának statisztikus vizsgálatai. (Statistical investigations of the applicability of natural effects in long-range forecasting.) Meteorológiai előrejelzések. Meteorológiai Tudományos Napok '78, *OMSZ Hiv. Kiadv.*, XLVIII. kötet, Budapest, 1979, 125-142. (Társszerzők: Bartholy J., Kaba M. és Légrády G.)

1980 A method of analysing types using analogy indices, *Acta Climatologica*, 1980, XV-XVI, 1.-4. 11-17. (Társszerző: Bartholy J.)

1980 Konvektív felhők jégeső-veszélyességének felismerése. *OMSZ Kiadvány* a Meteorológiai Napok 79' alkalmából, 1980. (Társszerzők: Wirth E., Győre S., Kőhegyi I., Légrádi G.)

1981 Fejezetek a matematikai statisztika meteorológiai alkalmazásaiból. (Methods of mathematical statistics in meteorology.) *OMSZ Meteorológiai Tanulmányok*, 1981, N 34, 1-84.

1983 Évszakos bontású makroszinoptikus típusok kialakítáasa clusteranalízissel az atlanti-európai térségre. (Determination of seasonal macrosynoptic types for the atlantic-european region by cluster analysis.) *OMSZ Meteorológiai Tanulmányok*, 1983, N 39, 1-80. (Társszerzők: Ambrózy P. és Bartholy J.)

1984 Meteorológiai megfigyeléssorozatok szélső értékeinek statisztikája. *Időjárás*, 1984, Vol. 86, N 2, 101- 108. (Társszerzők: Balog M. és Szentimrey T.)

1984 A system of seasonal macrocirculation patterns for the Atlantic-European region. *Időjárás*, 1984, Vol. 88, N 3, 121-133. (with A. Ambrozy and J. Bartholy)

1984 Оценка распознаваемости в задачах дистационных исследовании. *Исследование Земли из космоса*, 1984, N 3, 82-88. (Соовтор: Т. Фараго)

1984 Többsávos digitális műholdképek számítógépes clusterezése. *Időjárás*, 1984, Vol. 88, N 3, 161-173. (Társszerzők: Ketskeméty L. és Korándi M.)

1984 Vizsgálatok a paradicsom érésüteménének és a termésmennyiség előrejelzéséhez. *OMSZ Beszámoló Kötet,* 1984, 144-150.

1985 Matematikusok a meteorológiában. *Légkör*, 1985, XXX., N 2, 17-19.

1985 Komplexen integrált konzervipari számítógépes termelésirányítást előkészítő kutatások. *OMFB tanulmányok*, 10 kötet: 1977-1985 (Témavezető: dr. Antal Emánuel).

1986 A folytonos paraméterű folyamatok diszkretizálásának módszerei. Idősorok analízise (Szerk.: Tusnády G. és Ziermann M.), IV. fejezet, *Műszaki Könyvkiadó*, Budapest, 1986, 119-139.

1986 Meteorológiai idősorok periodicitásának elemzése I. *Időjárás*, 1986, Vol. 90, N 1, 14-23. (Társszerzők: Hamed, A. F., Ketskeméty L.)

1986 Osztályozási módszerek. Többváltozós statisztikai analízis (Szerk.: Móri F. T., Székely J.G.), X. fejezet, *Műszaki Könyvkiadó*, Budapest, 1986, 173-190.

1987 Néhány szó a meteorológia népszerűsítéséről. *Légkör*, 1987,XXXII., N 2, 22-25.

1987 Véletlenszerűek-e a légköri folyamatok, *Légkör*, 1987, XXXII., N 1, 7-10. (Társszerző: Götz G.)

1988 Matematikai statisztikai módszerek a meteorológiában. *Tankönyvkiadó*, Budapest, 1988. (Társszerző: Dévényi D.)

1988 Meteorológiai idősorok periodicitásának elemzése II. *Időjárás*, 1988, Vol. 92, N 1, 38-45. (Társszerzők: Hamed, A. F., Szentimrey T.)