

Globális klímaváltozás és a megfigyelési adatok pontossága

Domonkos Péter

Összefoglalás

E tanulmány középpontjában annak a kérdésnek a megválaszolása áll, hogy mennyire fontos a megfigyelt éghajlati adatok pontossága viszonylag gyorsan átalakuló klimatikus viszonyok idején. Ennek során áttekintjük a globális felmelegedés mértékéről szóló kiemelten fontos információkat és a közép- és dél-Európában megfigyelt regionális sajátosságokat. A megfigyelési adatok pontossága, azok tér- és időbeli reprezentativitása változatlanul fontos a légköri folyamatok és a változási tendenciák helyes értelmezéséhez. A megfigyelési adatok pontossága különösen fontos valóság-hű éghajlati modellek konstruálásához, mert a modellfejlesztés elengedhetetlen lépése a modellek megfigyelésből származó adatokon történő tesztelése. A megfigyelt globális vagy regionális trendek esetleges hibái téves irányban befolyásolhatják a modellezést. Az éghajlati trendek és éghajlati variabilitás megbízható értékeléséhez szükséges hogy a megfigyelésből származó éghajlati idősorok mentesek legyenek bármifajta nem-klimatikus hatástól, szaknyelven inhomogenitástól, amelyet a mérés technikai feltételeiben előforduló változások okozhatnak. Az éghajlati adatok hosszú történetében sokféle technikai változás történt, ezek különböző időszakokban és különböző mértékben befolyásolják az egyes megfigyelőállomásokról származó idősorok homogenitását. Olyan technikai változások is előfordultak, amelyek azonos irányú eltérést okoznak az egyes idősorokban, és ezek a regionális és globális trendek pontosságát is befolyásolják. E problémák miatt a megfigyelésből származó éghajlati idősorokat homogenizáljuk. A homogenizáláshoz statisztikai módszereket és a technikai változásokról feljegyzett dokumentumokat, szaknyelven meta-adatokat használunk fel. A statisztikai homogenizálás alapelve az azonos éghajlati zónában megfigyelt idősorok egymással történő összehasonlítása, és számos ilyen elven működő homogenizáló módszer van alkalmazásban. Módszerösszehasonlító tesztek azt mutatják, hogy a homogenizálás eredményessége széles határok között változik az adatmező tulajdonságai és az alkalmazott módszer szerint is, és leggyakrabban az ACMANT módszer szűri ki legnagyobb hatékonysággal a nem-klimatikus hatásokat. A tanulmány bemutat néhány kulcsfontosságú teszt eredményt, és tárgyalja röviden az eredményes homogenizálás elméleti feltételeit és az aktuálisan zajló új fejlesztéseket.

Kulcsszavak: globális felmelegedés, megfigyelt éghajlati trend, hőmérséklet, homogenizálás, módszer-összehasonlító teszt

Abstract

The central question of this study is the importance of the accuracy of observed climatic data in the era of relatively fast-changing climates. The review presents some key information about the actual state of global warming and its regional particularities in central and southern Europe. The accuracy and representativeness of observed climatic data both in space and time are still important, since these are necessary for the correct interpretation of the atmospheric processes and climatic tendencies. The accuracy of observed data has particular importance in climate modelling, since an indispensable step of climate model developments is the testing of models against observed climatic data. Possible errors in the global or regional trends calculated from observed data could push the model developments towards mistaken directions. The reliable evaluation of climate trends and climate variability needs observed climate data free from any non-climatic effect (called inhomogeneity), which may be caused by the changes of the technical conditions of observation. In the long history of climate observations, various changes occurred in the technical conditions, which affect the homogeneity of time series of individual observing sites in varied degrees and periods. Some changes caused biases of the same direction in individual time series, hence they also affect the accuracies of global and regional trends. For these reasons, observed time series are subjected to homogenization. Homogenization can be performed with statistical methods and by using documents of technical changes (so-called metadata). The principle of statistical homogenization is the comparison of time series within networks of the same climatic zone, and several of such homogenization methods are used in practice. Method comparison tests show that efficiency varies widely according to both dataset properties and homogenization methods, and in most cases the ACMANT method removes non-climatic effects with the highest efficiency. The study presents some test results of key importance, and briefly discusses the theoretical conditions of efficient homogenization and the ongoing developments.

Keywords: global warming, observed climatic trends, temperature, homogenization, method comparison test

1. Bevezetés

Hiába tanulmányozzuk a múltban megfigyelt éghajlatról készült adatokat és feljegyzéseket: a klíma már most sem mindenben hasonló a múltban megfigyelt éghajlathoz, és még kevésbé marad az a következő évtizedekben. E tanulmány középpontjában az a kérdés áll, hogy érdemes-e foglalkozni a múltban megfigyelt adatok pontosságával az időben gyorsan átalakuló éghajlatok korában. A kérdés megválaszolásához megvizsgáljuk hogy milyen jellegű és milyen nagyságrendű minőségi problémák fordulnak elő a múltból származó éghajlati adatokban, ezek milyen gyakorisággal fordulnak elő, és a kapcsolódó pontatlanságok mennyiben befolyásolhatják az éghajlatváltozáshoz történő alkalmazkodást. A kérdés elemzése arra az eredményre vezet, hogy a klímaadatok pontossága továbbra is nagy fontosságú, és a megfigyelésből származó idősorok éghajlati vizsgálatokhoz történő alkalmassá tételének egyik elengedhetetlen lépése az idősorok homogenizálása. A központi kérdés vizsgálatán túl, a tanulmány megjeleníti a globális klímaváltozás néhány fontos és aktuális tulajdonságát, tárgyalja az idősorok homogenizálásának elvét és alapvető tulajdonságait, valamint a homogenizálás pontosságáról szóló aktuális ismereteket. A tanulmány arra is kitér, hogy milyen, a homogenizálással összefüggő további feladatok állnak a klimatológus közösség előtt.

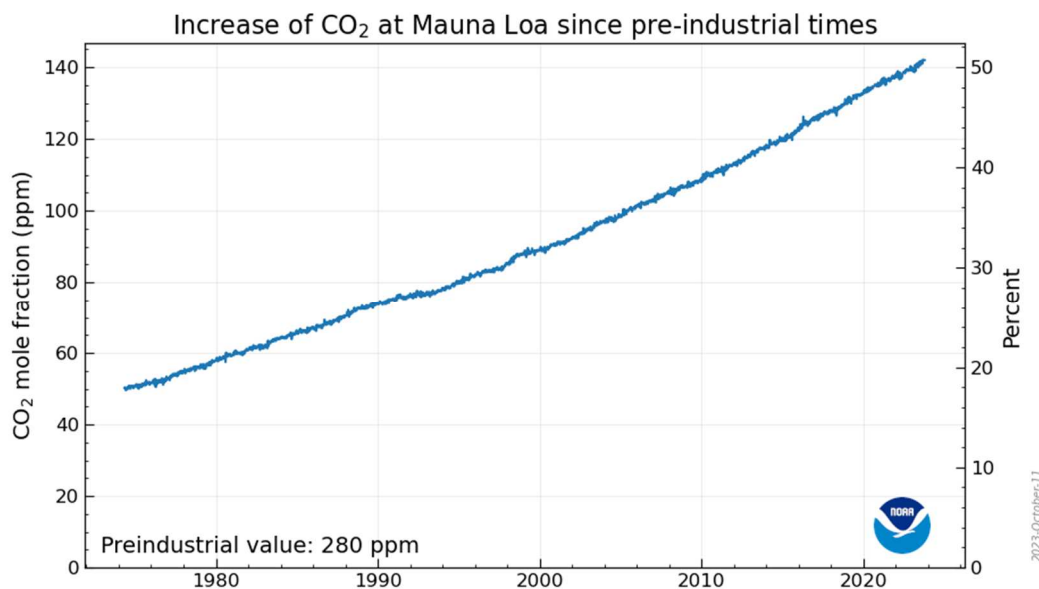
A tanulmány szerkezeti felépítése a következő: a 2. fejezet a globális felmelegedés aktuális állapotát mutatja be röviden, a 3. fejezet a megfigyelési adatok pontosságának szerepét és az adatok kutatási célra történő felhasználásának korlátait befolyásoló tényezőket elemzi, míg a 4. fejezet a homogenizálással kapcsolatos néhány elemi fogalmat, a gyakran használt homogenizáló módszerek általános tulajdonságait, valamint a leggyakrabban felmerülő problémákat tárgyalja. Az 5. fejezet néhány modern homogenizáló módszer eredményeit és korlátait mutatja be nemzetközi módszer-összehasonlító tesztek eredményei alapján, és megemlíti néhány olyan új és ígéretes módszerfejlesztést, amelyek még nem kerültek tesztelésre. Végül a 6. fejezet a jövőre vonatkozó feladatokról szól.

2. Globális felmelegedés: tények és tendenciák

2.1. Változások a légkör összetételében és a globális átlaghőmérsékletben

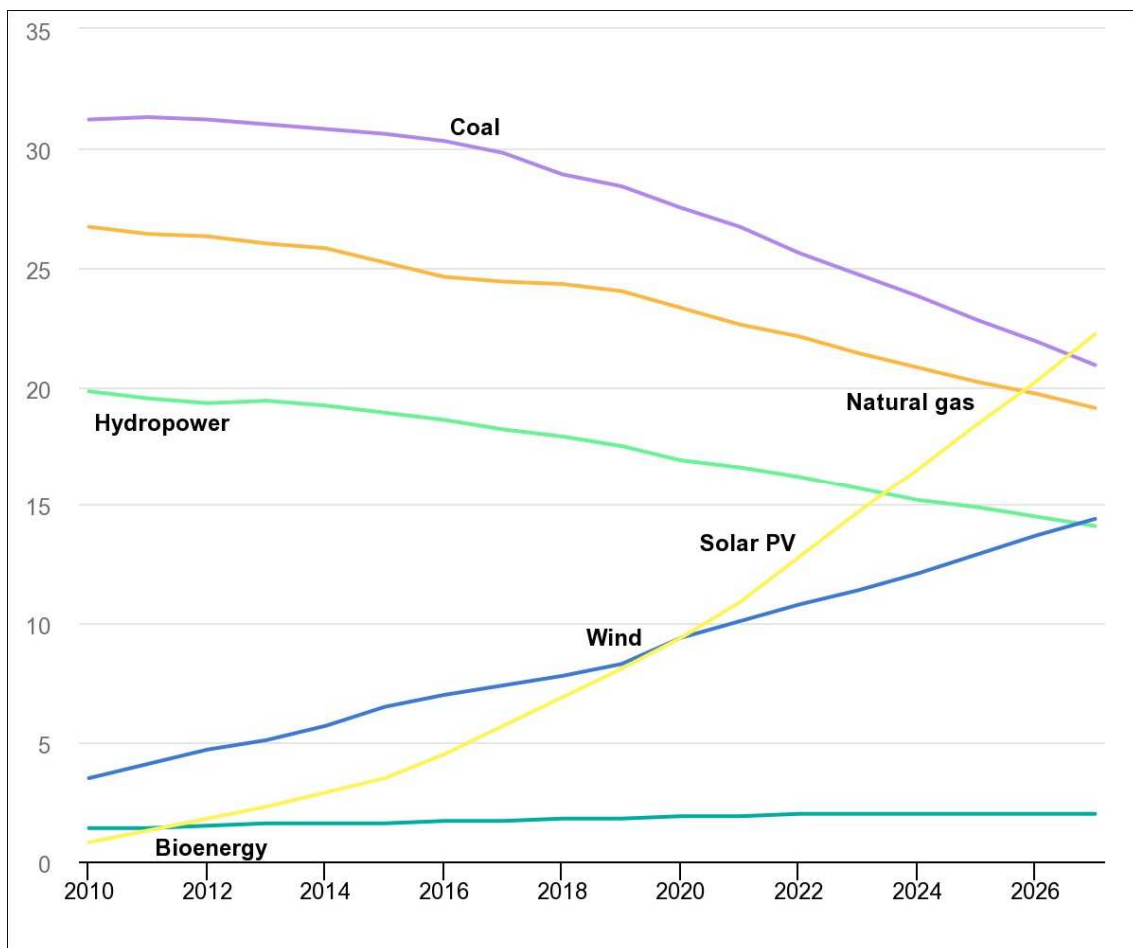
Harmincöt évvel ezelőtt, 1988-ban, az Egyesült Nemzetek Szervezetén belül alakult meg az Intergovernmental Panel on Climate Change (IPCC), abból a célból, hogy az éghajlatváltozás veszélyeiről és e veszélyek mérsékléséhez szükséges teendőkről objektíven informálja a politikusokat és a közvéleményt. Az IPCC 1990-ben adta ki első jelentését, amelyet átlag 6-7 évenként frissebb és pontosabb információkat tartalmazó újabb jelentések követtek. Mindegyik IPCC jelentésből egyértelmű, hogy az üvegházhatású gázok kibocsátásának csökkentése nélkül a földi éghajlat gyorsuló változáson megy át a XXI. század folyamán, és e változások nagy hányada kedvezőtlen hatású arra a természeti környezetre, amely megélhetésünk és életünk minőségének egyik fontos pillére. Hol tart ma az éghajlatváltozás a ~30 évvel ezelőtti becslésekhez képest?

Az 1. ábra a globális szén-dioxid koncentráció növekedési ütemét szemlélteti a NOAA egyik háttérszennyezettség-mérő állomásának (Mauna Loa Observatory) adatai alapján (1. ábra)



1. ábra. Légköri CO₂ koncentráció változása 1977 és 2022 között a háttérszennyezettség megfigyelő Mauna Loa Obszervatóriumban (Hawaii) mért adatok alapján. Forrás: <https://gml.noaa.gov/ccgg/trends/weekly.html>.

Az elmúlt 45 évből származó mérési adatok alapján a szén-dioxid koncentráció folyamatosan növekszik, sőt a növekedés ütemében is van enyhe fokozódás. Érdeemben nem jobb a helyzet a többi üvegházhatású gáz vonatkozásában sem, például a metán és a dinitrogén-oxid koncentrációi is meredeken növekednek. A szén-dioxid koncentráció utóbbi évtizedekben megfigyelt növekedési trendje szinte hajszálpontosan egyezik az első IPCC jelentésben bemutatott „Business As Usual” forgatókönyv adataival (IPCC, 1990). Tömören fogalmazva, az éghajlatváltozás megfékezésére tett eddigi erőfeszítések sikertelenek voltak. Ennek ellentmondani látszik, hogy a megújuló energiák használata gyors ütemben növekszik (2. ábra). A globális energiafelhasználás azonban szintén gyors növekedésben van, és ez egyelőre kioltja a megújuló energiák használatára történő átállás potenciális hasznát.



2. ábra. A napenergia és szélenergia felhasználás gyorsan növekszik, míg a szén alapú energiaforrások felhasználásának aránya az összes energiafogyasztáson belül csökken. Forrás: International Energy Agency (<https://www.iea.org/energy-system/renewables/solar-pv>).

A globális felmelegedés tehát folyamatban van, és annak üteme jó egyezést mutat az IPCC által is felhasznált éghajlati modelleken alapuló előrejelzésekkel (1. Táblázat).

1. Táblázat. Felszíni hőmérséklet átlagos melegedése a 2020-at megelőző 140 évben és 40 évben (forrás: IPCC, 2021).

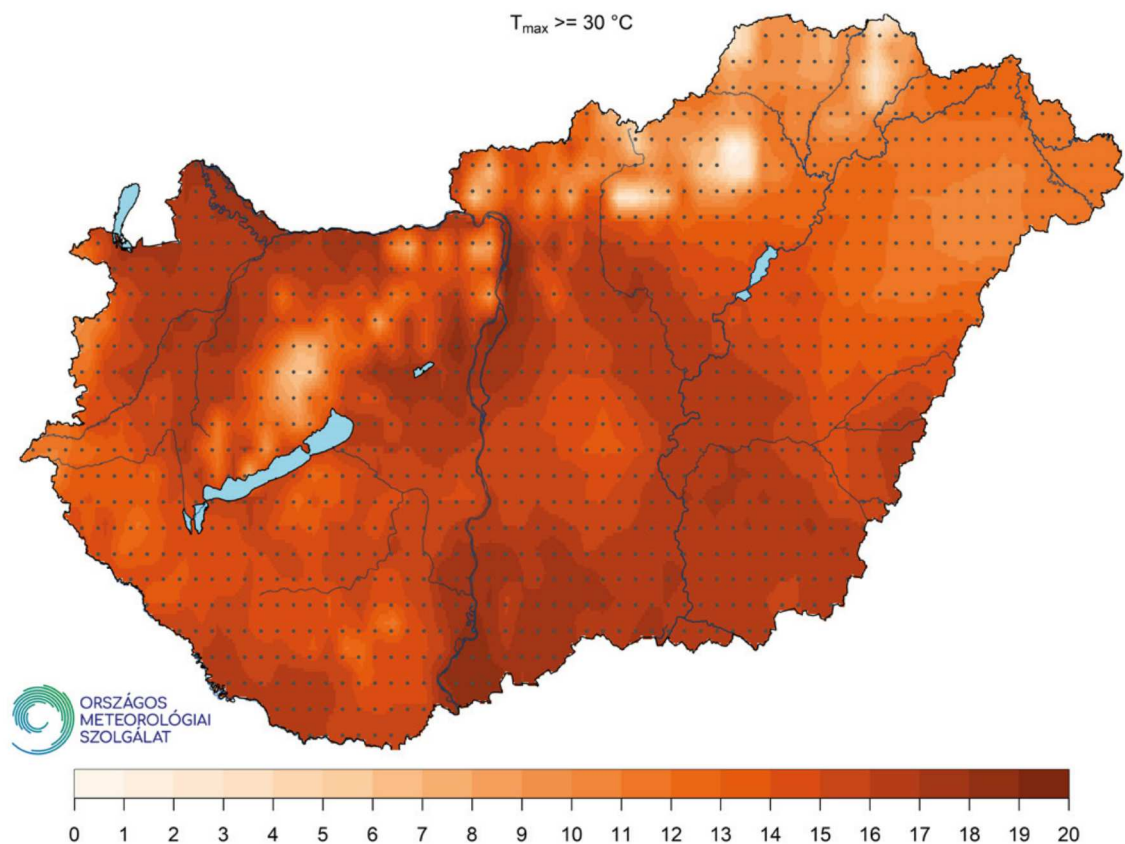
	1880-2020 (°C)	1980-2020 (°C)
Kontinentális felszínek hőmérséklete	1.50	1.18
Oceáni felszínek hőmérséklete	0.96	0.60
Globális felszínhőmérséklet	1.11	0.76

Az 1. Táblázat adatai a világ különböző pontjain működő kutatócsoportok egymástól független számításainak átlagolásán alapulnak. Az egyes kutatócsoportok által számított értékek szóródása a várakozásnak megfelelően kicsi (IPCC, 2021), hiszen a számítások múltbeli megfigyeléseken alapulnak. Az átlagosan $\sim 1^\circ\text{C} / 100$ év felmelegedési ütem nem tűnik ijesztőnek, még azzal együtt sem, hogy 1980 után lényegesen felgyorsult a melegedés. Érdekes azonban az 1. Táblázat egy részletére külön is figyelni: az 1980 utáni felmelegedés átlagos üteme kétszer olyan gyors a kontinenseken mint az óceánok felett. Az átlagos hőmérséklet megváltozását pedig a szélső hőmérsékletek változásai, és más éghajlati szélsőségek változásai kísérik.

2.2. Hőhullám, aszály, heves csapadékhullás

Az extrém éghajlati események gyakoriságának és intenzitásának változása több gondot okoz mint az átlagos hőmérsékletek emelkedése. Az extrém események statisztikai jellemzőinek megváltozása, elsősorban a hőhullámok és aszályos időszakok intenzitásának, hosszának és gyakoriságának a növekedése, valamint a hirtelen lehulló nagy csapadékok gyakoriságának és maximális mennyiségének a növekedése, elméleti meteorológiai ismeretek, éghajlati modell eredmények és megfigyelési adatok alapján is együttjárnak a globális felmelegedéssel. A számszerű jellemzés azonban itt nagyobb bizonytalansággal terhelt mint az átlagos hőmérsékletek megváltozása esetében. Ennek az extrém események nagy idő- és térbeni változékonysága, valamint a megfigyelt ritka szélsőségekhez tartozó statisztikai minták kicsi mérete az oka. A nyári hőségek gyakoribbá és intenzívebbé válását viszont alighanem már a bőrünkön is érezzük, és a változást megfigyelési adatok is alátámasztják (Lakatos és mts., 2021; Bokros & Lakatos,

2022). A 3. ábra a magyarországi hőségnapok (napi maximumhőmérséklet $\geq 30^{\circ}\text{C}$) 1901 és 2021 közötti növekedését mutatja be.



3. ábra. Hőségnapok (napi maximumhőmérséklet $\geq 30^{\circ}\text{C}$) számának változása Magyarországon az 1901-2021 időszakban (forrás: Bokros&Lakatos, 2022).

Amikor Szerző egyetemista volt (régén volt), azt tanultuk Magyarország éghajlatáról, hogy évi átlagban 12 hőségnap fordul elő a sík vidéki területeken. Nem kétséges, hogy a jelenkor éghajlata már minőségileg különbözik a múlt század nagy részében megfigyelt éghajlattól. Ez így van Magyarországon, és így van a világ más tájain is.

Egy közelmúltban készült tanulmány szerint Katalóniában 1950 és 2021 között a nyári hónapokban gyorsabb volt a felmelegedés üteme mint az év többi részében, a napi átlagos hőingás növekedett, és a napi maximumhőmérsékletek emelkedése elérte a dekádonkénti 0.43°C -ot (Prohom et al., 2023). Ez a melegedési ütem ilyen hosszúságú időszakon a globális és kontinentális átlagokat is lényegesen meghaladja. A globális átlagot meghaladó mértékű melegedés nemcsak Katalóniában, hanem az egész földközi-tengeri régióban jellemző (MedECC, 2020), és a nyári hőségperiódusok intenzívebbé válása kíséri. Míg a globális felmelegedés átlagos üteme a korábbi becsléseknek

megfelelő ütemű, a kontinens és óceáni felszínek felmelegedése közötti különbség, valamint az extrém hőségek intenzitásának növekedése nagyobbban mutatkoznak annál mint amit egyes korai modell becslések jeleztek (Kharin & Zwiers, 2000). Az 1980 utáni, fokozott mértékű dél-európai felmelegedés egyik magyarázata a légkör átlátszóságának növekedése az aeroszol koncentráció csökkenése miatt (MedECC, 2020). Egy további magyarázó tényező a fokozódó gyakoriságú nyári aszály, amely a párolgási hőelvonás csökkenését okozza. Nyáron az aszály és a hőség egymást erősítő tényezők.

A nyári csapadékhiány kialakulásában rendszerint cirkulációs tényezőknek és a helyi-regionális nedvességi anomáliáknak is van szerepük (Lyon & Dole, 1995). A relatív nedvesség anomáliáinak globális vizsgálata (Vicente-Serrano et al., 2018) azt mutatta, hogy a két tényező hozzájárulásának arányában lényeges különbségek vannak évszakok és földrajzi régiók szerint. A földközi-tengeri régió nyári éghajlatát a Hadley cella észak felé történő terjeszkedése sújtja (IPCC, 2021), és még a Kárpát-medence térségében is kimutathatók szignifikáns változások a nagytérségű cirkulációban (Mika et al., 2021). Az Ibériai-félsziget éghajlatát általános, gyakran az év nagy részére kiterjedő csapadékhiány sújtja, amely az azori szigetek akció centrumként ismert kvázi-perzisztens magasnyomás közelsége miatt van, tehát nem új jelenség, viszont az aszályhajlam súlyosbodó tendenciát mutat, és a növénytakaró degradálódását, egyes körzetekben elsivatagosodást okoz. A növénytakaró degradációja fokozza mind az aszály, mind pedig a nyári hőségek kialakulásának esélyét. A nyári hőség és aszály kapcsolatáról ausztrál kutatók tettek közzé kvantitatív analízist (Páscoa et al., 2022).

Összefoglalva, a globális felmelegedés ütemét eddig nem sikerült mérsékelni, és a felmelegedés bizonyosan folytatódni fog, mert a légkörben aktuálisan jelenlévő üvegházhatású gázok időben elnyújtva fogják kifejteni a földfelszínt melegítő hatásukat. A globális felmelegedés átlagos üteme megfelel az éghajlati modellek által előrejelzett felmelegedési ütemnek, míg a regionális és évszak szerinti különbségek kedvezőtlenebbnek mutatkoznak annál, mint amit korai becslések sugalltak. Az ember alkalmazkodóképessége első közelítésben kielégítőnek tűnhet, hiszen megfelelő hajlék, élelem, ivóvíz és energia biztosítása esetén emberek bárhol élhetnek a forró sivatagoktól a jégmezőig. Nagy problémát okoz azonban az a fejlemény, hogy a globális felmelegedés viszonylag rövid időn belül és igen nagy területeken teszi kedvezőtlenebbé, olykor hétköznapi értelemben elviselhetetlenné az életfeltételeket. A legnagyobb gondot pedig az okozza, hogy több globális probléma, különösen a túlnépesedés, az óriási gazdasági különbségek, a túlságosan éles politikai és ideológiai konfrontációk és a

klímaváltozás együtt, egymással kölcsönhatásban fogják kifejteni hatásukat. Számos jel mutat arra, hogy a civilizáció fejlődése egy nagyon nehéz időszak előtt áll, de ennek kifejtése nem tartozik e tanulmány keretei közé.

3. A megfigyelési adatok pontosságának szerepe

A különböző földrajzi régiók éghajlatának megismeréséhez nincs szükségünk nagy adatpontosságra. Sőt, mivel a klíma folyamatosan változik, valószínű, hogy a soronkövetkező évek időjárása és éghajlata még a statisztikai átlagok tekintetében is különbözni fog attól mint amit a megfigyelési adatok mutatnak.

A klímával professzionálisan foglalkozók számára viszont fontos a megfigyelési adatok pontossága, például a bekövetkezett éghajlatváltozás mértékének és ennek területi és időbeni differenciáinak az elemzéséhez. Az 1. Táblázat mutatta, hogy a globális és kontinentális skálájú hőmérsékletváltozásokat szokás két tizedesjegy pontossággal megadni, ismert pontossági korlátok ellenére is. A megfigyelt éghajlati adatok hozzáférhetősége és pontossága nagy jelentőségű az éghajlatváltozással együttjáró folyamatok helyes értelmezéséhez és hatékony alkalmazkodási stratégiák kidolgozásához.

3.1. A megfigyelési adatok pontosságának szerepe az éghajlati modellezésben

Az éghajlatkutatás, különösen pedig az éghajlat-előrejelzés egyik alapvető módszere a globális cirkulációs modellek (GCM) alkalmazása. E modellek a földi atmoszféra főbb folyamatait matematikai egyenletek formájában követik nyomon, míg a külső hatások és egyes kevésbé domináns folyamatok empirikus parametrizációval kerülnek a modellekbe. Globális és kontinentális skálájú éghajlati előrejelzések a GCM-ek közvetlen alkalmazásával (Eyring et al., 2016), míg a kisebb térskálájú előrejelzések a GCM-ek és regionális cirkulációs modellek együttes használatával (Szabó et al., 2022; Megyeri-Korotaj et al., 2023) készülnek. Az éghajlati modellek megbízható működésének egyik indikátora a megfigyelt éghajlati változások reprodukálásának minősége. Santer et al. (2021) a modellezett troposzférikus hőmérsékletek pontosságát vizsgálták 5, a „Coupled Model Intercomparison Project”-ben részt vevő GCM-ek (CMIP modellek) eredményeinek értékelésével. Az 1979 és 2020 közötti megfigyelt és modellezett adatok

összevetése azt mutatja, hogy rövid, 1-2 éves időszakokat vizsgálva olykor ugyan előfordulnak 0.5°C -ot meghaladó eltérések, de a hosszútávú trendek szinte hajszálpontosak. Itt merül fel az a kérdés, hogy hogyan hat az éghajlati modellezésre, ha a megfigyelési adatok trendje pontatlan a felhasznált adatokban lévő hibák vagy inhomogenitások miatt. Ez nemcsak azzal okozhat kárt, hogy a modellek reprodukációs képessége a ténylegesnél gyengébbnek mutatkozna, hanem hamis információt adhat a modellfejlesztők számára. Minden cirkulációs modell tartalmaz empirikusan becsült paramétereket, mert a légkör a maga teljes bonyolultságában nem képezhető le véges számú matematikai egyenletbe. A GCM-ek alkalmasságának egyik indikátora a megfigyelt éghajlatot reprodukáló képességük, annak ellenére is, hogy a jelenkori klíma pontos reprodukálása nem nyújt biztosítékot arra, hogy a modell pontosan működik megváltozott klimatikus körülmények között is. A GCM-ek jelenkori klímát reprodukáló képességének optimalizálása állandó része a modellfejlesztéseknek, ezért a megfigyelési adatok hibái téves irányban befolyásolhatják a fejlesztői munkát. Megjegyzés: a leírt példa egyszerűsítést tartalmaz abban a tekintetben, hogy míg a megfigyelési adatok a földfelszíni mérések esetében a legpontosabbak, addig a GCM adatok a szabad légkörben a legpontosabbak. Ez, a talajfelszín és a troposzféra éghajlati jellemzői közötti szoros összefüggések miatt nem változtat a lényegen: a megfigyelési adatok pontossága döntő fontosságú az éghajlati modellezés, különösen pedig a megbízható éghajlati előrejelzések készítése számára. A regionális átlagok pontosságának fontossága meghaladja a kisebb térségekre vonatkozó adatok pontossági igényét, és általában a hosszabb időszakok átlagában mutatkozó trendek pontossága a legfontosabb.

3.2. A megfigyelésből származó éghajlati adatok felhasználhatóságát korlátozó tényezők

Megfigyelt és feljegyzett éghajlati adatok nemcsak adathiba miatt lehetnek csökkent mértékben alkalmasak az éghajlatváltozás és éghajlati variabilitás vizsgálatára, hanem olyan esetekben is, amikor az adat ugyan korrekt, de valamilyen okból nem alkalmas a térbeli vagy időbeli változások reprezentálására. Három féle probléma lehetséges: i) adathiba, ii) az adat nem reprezentálja megfelelően a célul tűzött éghajlati változót, iii) régebbi megfigyelésekből származó adatok nem alkalmasak az újabb megfigyelésekkel történő összehasonlításra a mérés technikai feltételeinek időközben történt változásai miatt.

i) Adathibák forrásai a mérőműszer vagy a megfigyelést végző személy hibái, ezeken kívül pedig adattovábbítási vagy adatrögzítési hibák is előfordulnak. A professzionális megfigyelőhálózatokban az esetleges adathibák előfordulását rendszeresen ellenőrzik minőségellenőrzési (QC) folyamat keretében, és a nagyobb hibák kiszűrése gyakorlatilag 100 százalékos. Fejlett QC eljárások (Aguilar & Prohom, 2010; Durre et al., 2010) jó eséllyel szűrnek ki olyan hibákat is, amelyek mértéke ugyan kicsi, de gyakran fordulnak elő, és emiatt jelenlétük eltérést okozhat hosszabb időszakok átlagában is. Kisebb méretű, nem szisztematikus hibákat a QC eljárások nem tudnak kiszűrni, de ezek érdemben nem befolyásolják az éghajlati vizsgálatok eredményeit.

ii) Reprezentativitás hiánya. Hiába korrekt a megfigyelési és adatrögzítési folyamat, ha a megfigyelés nem arra irányult amire az adatot használni kívánjuk. Például a standard magasságban végzett hőmérsékletmérések nem alkalmasak a talajmenti fagy gyakoriságának jellemzésére. Nyilvánvaló, hogy agrometeorológiai, biometeorológiai vagy városklimatológiai kutatásokban mind a mérések reprezentativitására, mind pedig esetleges mérési hibák kiküszöbölésére a szokásosnál is nagyobb gondot kell fordítani. A globális vagy nagyobb régiók átlagaira számított éghajlati jellemzőknél más formákban jelenik meg az adat-representativitás probléma: vagy úgy, hogy az adat egy bizonyos megfigyelőhelyen pontos, de a környező térség éghajlatára kevésbé jellemző, vagy pedig úgy, hogy a megfigyelési adatok elégtelen térbeli sűrűsége miatt az adatok nem alkalmasak a terület egészének reprezentálására. Ez utóbbit mintavételi hibának is szokás nevezni (Shen et al., 2007) a globális vagy kontinentális skálájú éghajlati vizsgálatokban. Az 1980-at követő időszakban a mintavételi hiba szerepe jelentéktelen, mert műholdas mérések és modellezett megfigyelési adatok (reanalízis adatok) egészítik ki a felhasználható adatbázist. Ezzel szemben a XX. század első feléből vagy még korábbról származó regionális adatok esetében a mintavételi hiba szerepe jelentős.

iii) Időbeni összehasonlíthatóság hiánya: ezt az adatsorokban lévő nem-klimatikus változásoknak vagy még gyakrabban inhomogenitásnak is nevezzük. Szinte minden hosszantartóan végzett éghajlati megfigyelés történetében előfordulnak technikai változások, pl. állomás áthelyezés, műszertípus csere, megfigyelési időpont változása, vagy a megfigyelés környezetének változása. A globális vagy regionális éghajlati elemzések esetében azok az inhomogenitások a legfontosabbak, amelyek azonos irányú eltérést okoznak egy-egy régió nagy részében, miáltal a regionális átlagok idősorában is jelen lehetnek az éghajlati vizsgálatok pontosságát befolyásoló nem-klimatikus változások. Például a felszíni léghőmérséklet korai műszeres megfigyelésének

történetében a hőmérők elhelyezésének szabályai és a hőmérők árnyékolása többször is változott, és a tendencia világszerte abban az irányban haladt, hogy az újabb módszerek nagyobb mértékben szűrték ki a sugárzási szennyezést a korábbiaknál. Ez szisztematikus pozitív eltérést okoz a korai hőmérsékletekben amikor azokat összevetjük az 1950 után megfigyelt mérések eredményeivel (Menne et al., 2018). A nem kívánatos nem-klimatikus változásoknak legalább egy része kiszűrhető az adatokból alkalmas homogenizáló módszerek használatával. Ez az a pont ahol a tanulmány két témája, az éghajlatváltozás és a megfigyelésekből származó adatok pontossága egymáshoz kapcsolódnak.

4. Éghajlati idősorok homogenizálása

Az éghajlati idősorok homogenizálása egy fontos tekintetben különbözik a homogenizálás szó szokványos értelmezésétől: az éghajlati idősorok homogenizálása (továbbiakban: homogenizálás) során a valós éghajlati változásokat törekszünk megőrizni, miáltal az idősorok általában nem válnak matematikai értelemben homogénné. A homogenizálás egyik fő problémája a nem-klimatikus és klimatikus változások helyes szétválasztása, és ennek leggyakoribb módszere az aktuálisan vizsgált idősor (jelölt idősor) adatainak összevetése az azonos éghajlati régióba eső más megfigyelőállomások idősoraival, amelyeket szomszéd idősoroknak nevezünk. Az így végzett homogenizálást relatív homogenizálásnak nevezzük. E tanulmány kizárólag relatív homogenizálással foglalkozik.

4.1. A relatív homogenizálás alapvető fogalmai és módszerei

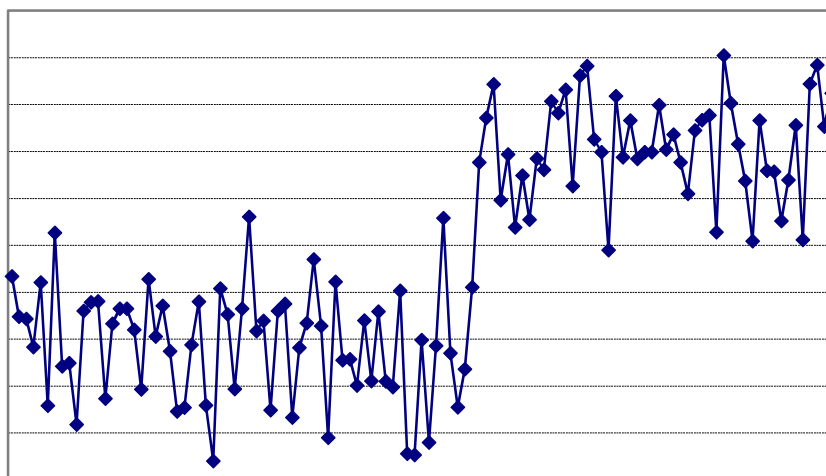
A relatív homogenizálás kiindulási gondolata az, hogy ha az esetlegesen inhomogén n elemű jelölt idősort (\mathbf{X}) össze tudjuk hasonlítani egy garantáltan homogén szomszéd idősorral (\mathbf{F}), akkor a két idősor különbségében (\mathbf{T}) a jelölt idősor inhomogenitásai láthatóvá válnak azáltal hogy a két idősor közös éghajlati variabilitása (\mathbf{U}) kiesik a különbség képzése során. Az inhomogenitások befolyását (ezt szokás állomás-hatásnak is nevezni, mert az inhomogenitások nagy hányada specifikus a megfigyelőhelyre) \mathbf{V} , az időjárás térbeli különbségeiből származó zaj idősorát pedig ϵ jelöli.

$$\mathbf{X} = \mathbf{U} + \mathbf{V} + \boldsymbol{\varepsilon} \quad (\mathbf{X} = x_1, x_2, \dots, x_i \dots x_n \quad (i = 1, 2 \dots n)) \quad (1)$$

$$\mathbf{T} = \mathbf{X} - \mathbf{F} = \mathbf{V}_X - \mathbf{V}_F + \boldsymbol{\varepsilon}' \quad (2)$$

Az (1) és (2) egyenletek azt szemléltetik, hogy az éghajlati variabilitás eltűnik a különbség képzése során, és abban a szerencsés esetben ha a felhasznált szomszéd idősor (ezt szokás referencia idősornak nevezni) homogén, vagyis $\mathbf{V}_F \equiv 0$, csupán egy nem szisztematikus zaj zavarja a jelölt idősor inhomogenitásainak felismerését.

Megjegyzendő, hogy csapadékmennyiségnél a jelölt idősor és referencia idősor hányadosát kell képezni, és e hányadosból esik ki az éghajlati variabilitás, de a következőkben csak különbség idősor kerül megemlítésre, az egyszerűség kedvéért. A különbség idősorokban tehát nincs jelen a zónára jellemző éghajlati jele, de az időjárás rövid távú térbeli különbségei, valamint esetlegesen előforduló nem szisztematikus hibák továbbra is zavarják az inhomogenitások detektálását. A jel-zaj arány nagyon különböző lehet, szerencsés esetben egyes inhomogenitások akár ránézésre is észrevehetők (4. ábra).



4. ábra. Szintetikusan generált fehér zaj idősor egy nagy méretű ugrást okozó töréspont hozzáadásával.

A legtöbb inhomogenitás a szakasz-átlag értékek hirtelen megváltozását tartalmazza, úgy, ahogyan azt a 4. ábra szemlélteti. Az ilyen változást töréspontnak nevezzük. A homogénizálási feladatok többségében a töréspontok felismerése sokkal nehezebb mint a 4. ábrán. Először azért, mert a változások zajhoz viszonyított mértéke nem szokott olyan nagy lenni mint a 4. ábra példáján, másodsor azért, mert a referencia idősorok általában nem tökéletesen homogének, és a referencia idősorok inhomogenitásai addicionális

zajként hatnak, harmadszor pedig azért, mert gyakori hogy egy jelölt idősorban több töréspont vagy esetleg más típusú inhomogenitások is előfordulnak, és az egyes töréspontok detektálásánál az egyéb, kisebb méretű inhomogenitások szintén zajként hatnak. Ezen okok miatt az esetek túlnyomó hányadában a töréspontok detektálása nem szemrevételezéssel, hanem statisztikai módszerek alkalmazásával történik. Ha rendelkezésre állnak dokumentumok az éghajlati megfigyelések során előfordult technikai változásokról (ú.n. meta-adatok), akkor azok szintén felhasználhatók a töréspontok detektálásánál. A homogenizálás sikeressége azonban nemcsak a töréspontok számának és bekövetkezési idejüknél a helyes detektálásától, hanem a referencia idősorok minőségétől és használatuk módjától, valamint az inhomogenitások korrigálására alkalmazott módszertől is függnék.

Egy relatív homogenizáló módszernek legalább három része van, ezek i) különbség idősorok képzése, ii) inhomogenitások detektálása, iii) inhomogenitások megszüntetése a megfigyelési adatok korrekciójával.

i) A referencia idősorok kiválasztásánál problémát okoz, hogy homogenizálás előtt nem tudjuk hogy mely idősorok tekinthetők legalább megközelítőleg homogénnek (ha egyáltalán vannak ilyenek). Ezért a relatív homogenizálás során általában nem elégszünk meg egyetlen szomszéd állomás idősorának a felhasználásával, és ideális esetben a felhasznált szomszéd idősorok száma 10-nél nagyobb (Menne & Williams, 2009; Squintu et al., 2019). Az azonos éghajlati régióhoz tartozó megfigyelési idősorokat általában együtt, egy közös adatmezőben homogenizáljuk, úgy hogy mindegyik idősor felveszi egyszer a jelölt idősor szerepét. A túlságosan nagy vagy túlságosan kis számú idősorot tartalmazó adatmezők használatát lehetőség szerint kerüljük (Rustemeier et al., 2017). Három féle különbség idősor képzési módot lehet alkalmazni: i) szomszéd állomások idősoraiból, azok átlagolásával szerkesztett egyetlen referencia idősor képzése (szerkesztett referencia idősor, Peterson & Easterling, 1994; Moberg & Alexandersson, 1997); ii) egy körzeten belül az összes szomszéd állomás idősorai egyenként kerülnek felhasználásra referencia idősor szerepben, erre a páronkénti összehasonlítás kifejezéssel szokás hivatkozni (Craddock, 1979; Caussinus & Mestre, 2004); iii) a szomszéd állomások csoportosításával és az azonos csoporthoz tartozó állomások idősorainak átlagolásával több szerkesztett referencia idősor képezhető (többszörös referencia idősorok használata, Szentimrey, 1999). A többszörös referencia idősorok használata kevésbé elterjedten használt a megelőzően említett két módszerhez viszonyítva. A szerkesztett referencia idősoroknak és a páronkénti összehasonlítás módszernek is vannak

sajátos előnyei és hátrányai, ezért egy viszonylag új fejlesztés a két eltérő tulajdonságú idősor összehasonlítási módszer kombinált alkalmazásával kísérletezik (Domonkos, 2021).

ii) Az inhomogenitások detektálása alatt általában a töréspontok felderítését, és előfordulásuk időpontjának meghatározását értjük. A töréspontoknál kisebb gyakorisággal, de előfordulnak olyan inhomogenitások is, amelyeknél az inhomogenitás mértéke nem hirtelen, hanem fokozatosan növekszik, és léteznek olyan inhomogenitás detektáló módszerek, amelyekkel nemcsak töréspontok, hanem lineárisan változó mértékű inhomogenitások is detektálhatók. Azonban a homogenizáló módszerek többségénél és a modern homogenizáló módszereknél minden esetben az inhomogenitás detektálás és töréspont detektálás kifejezések egymás szinonimjai. A töréspontok detektálása a szakasz átlagok statisztikai tulajdonságainak összevetésén alapul, és ennek sokféle változata van. Nem ritka, hogy egy idősorban több töréspont is van, emiatt a töréspontok együttes hatására ügyelni kell. Ennek ellenére, vannak olyan jól bevált módszerek, amelyek egy adott lépésben csak egyetlen töréspontot detektálnak (Alexandersson, 1986; Wang et al., 2007). Ezekben az esetekben az idősorokat a detektált töréspont helyén ketté vágjuk (binary segmentation, Easterling and Peterson, 1995), és a detektálás a részidősorokon folytatódik mindaddig, amíg további szignifikáns töréspontok találhatók. Szemben az eddig említettekkel, vannak olyan módszerek, amelyek egy idősor összes töréspontját egyetlen lépésben detektálják (Szentimrey, 1999; Caussinus & Mestre, 2004), sőt olyan módszer is létezik, amellyel egy adatmező idősorainak összes töréspontja egyidejűleg detektálható (Joint detection, Picard et al., 2011). Az elterjedten használt töréspont detektáló módszerek hatékonyságai között viszonylag kicsik a differenciák (Domonkos, 2011; Van Malderen et al., 2020).

iii) Az adatok korrigálásához meg kell határozni a detektált inhomogenitások nagyságát. Ez történhet a töréspontoknál előforduló szakasz-átlag értékek megváltozásának egyenkénti becslésével, térbeli interpolációval, vagy az együtt homogenizált adatmező éghajlati jel (u) és állomás-hatás (v) idősorainak egyetlen egyenletrendszerrel történő számításával. Ez utóbbi módszert ANOVA korrekciónak nevezzük Caussinus and Mestre (2004) nyomán, annak ellenére, hogy az ANOVA szélesebb körben használt jelentése más (variancia analízis). Az ANOVA korrekció modellje pontosan illeszkedik a megoldandó feladat igényeihez, és nem tartalmaz semmiféle olyan feltételt vagy paraméterbecslést, amelyek nincsenek jelen a korrekciók becslésének egyéb módszereinél (ez alól a térbeli interpolációval történő korrigálás

részben kivétel). Tömören fogalmazva, az ANOVA korrekció alkalmazása optimális (Lindau & Venema, 2018a; Domonkos & Joelsson, 2023), és kívánatos hogy az alkalmazása általánossá váljon. Az ANOVA korrekciós modellbe a detektált töréspontok dátumain kívül az ismert, állomás-specifikus technikai változásokat leíró meta-adatok dátumait is célszerű bevezetni (Domonkos, 2022a).

A homogenizálás során gyakran meg kell oldani olyan problémákat is, mint a nyilvánvalóan hibás adatok kiszűrése, évszak szerinti különbségek az inhomogenitás mértékében, vagy különböző hosszúságú, esetleg részben hiányos idősorok együttes használata. A homogenizáló eljárások nagy része a leírt három lépéshez képest több lépést tartalmaz azért is, mert az esetlegesen előforduló viszonylag nagy méretű inhomogenitások előzetes kiszűrése vagy mérséklése céljából egy homogenizáló eljárás kettő vagy több komplett homogenizálási ciklust tartalmazhat. További bonyodalom, hogy a homogenizálás nemcsak a szakasz-átlagok, hanem a valószínűségeloszlás egyéb tulajdonságainak a korrigálását is célul tűzheti. A téma iránt mélyebben érdeklődők további elemzéseket találnak egy közelmúltban publikált könyvben (Domonkos et al., 2022), illetve létezik egy magyar nyelvű tanulmány a homogenizáló módszerek 1996 és 2021 közötti fejlődéséről (Domonkos, 2022b). A 4.2. alfejezetben és az 5. fejezetben a homogenizálás téma néhány speciális kérdés tárgyalásával folytatódik.

4.2. A homogenizálás néhány speciális problémája

Sokféle homogenizálási feladat létezik, hiszen sokféle éghajlati elemnél merül fel a homogenizálás szükségessége, emellett a megfigyelési adatbázisok mérete, az idősorok hossza, a megfigyelési adatok térbeli és időbeli sűrűsége, a térbeli korrelációk szorossága és az adathiány gyakorisága is különböző lehet. Itt csak három probléma kerül tárgyalásra a gyakran felmerülő kérdések közül, ezek i) a jel-zaj arány hatása a homogenizálás eredményességére, ii) a regionális eltérést okozó inhomogenitások kezelése, és iii) a meta-adatok felhasználásának módja.

i) Nagy jel-zaj arány megkönnyíti az eredményes homogenizálást, ezt a 4. ábra is szemlélteti. A jel-zaj arány azonban lehet kicsi vagy azért mert a térbeli korrelációk alacsonyak, zajosabb különbség idősorokat okozva, vagy azért mert az idősorokban nincsenek jelentékeny méretű inhomogenitások. A megfigyelési idősorok térbeli sűrűsége, amennyiben nem mutatkozik kielégítőnek, néha növelhető azzal hogy viszonylag rövid, esetlegesen adathiányos idősorokat is felhasználunk, feltéve hogy azok

általános minősége elfogadható. Másképp fogalmazva, törekedni kell minden felhasználható információ együttes használatára, kivéve az információk túlzott bősége esetén. Amikor alacsony a jel-zaj arány, gyakori hogy nincs mód a megnövelésére, és szélső esetben jobb lehet elhagyni a homogenizálást mint megkísérelni egy túlságosan nagy bizonytalansággal járó beavatkozást a megfigyelési adatokba. A teljes kép azonban bonyolult, mert az alacsony jel-zaj arányú homogenizálás eredményességét jelentősen javíthatja meta-adatok, reanalízis adatok vagy más éghajlati adatok felhasználása. Vannak esetek, amikor az alacsony jel-zaj arány nyilvánvaló az adatmező nem kielégítő sűrűsége miatt. A jel-zaj arány azonban lehet alacsony más okokból is, például a rövid ideig fennálló vagy kis mértékű inhomogenitások is zajként szerepelhetnek, mert ezek detektálása statisztikai eszközökkel többnyire nem lehetséges. Az alacsony jel-zaj arány nem mindig felismerhető, mivel az inhomogenitások tulajdonságai (amiket nem pontosan ismerünk) befolyásolják azt. A jel-zaj arány lehet jelentősen különböző egy adatmező egyes idősorai között, egy idősor egyes szakaszai között, ezeken kívül pedig egy adott zajszint lehet elegendően kicsi egy töréspont detektálásához, míg túlságosan nagy egy másik töréspont esetében (Lindau & Venema, 2018b). A gondosan konstruált homogenizáló szoftverek alacsony jel-zaj arány esetén sem rontanak érdemben az adatminőségen: vagy a kiindulási pontatlanságokkal megegyező nagyságrendű pontatlanságokat tartalmazó eredményeket produkálnak, vagy visszautasítják automatikus homogenizálás végrehajtását.

ii) Azonos típusú inhomogenitások előfordulása egy adatmező több idősorában megnehezíti az eredményes relatív homogenizálást. A probléma azonos időpontban előforduló töréspontok esetén a legsúlyosabb, például egy új mérés technikai módszer szervezeti egységes bevezetése okozhat ilyet. A relatív homogenizálás elve az, hogy az inhomogenitás specifikus a jelölt idősorra, míg az idősorok közös változásai az éghajlati variabilitást mutatják. Ha az együtt homogenizált idősorok több mint felében azonos időszakot befolyásol egy azonos típusú inhomogenitás, akkor a relatív homogenizálás folyamata a homogén idősorokat fogja az inhomogénekhez igazítani, pont megfordítva a homogenizálás céljához képest. Szerencsére ez az eset ritka, mert az egy időben bevezetett változásokról többnyire rendelkezésre állnak a homogenizálás folyamán felhasználható meta-adatok. A homogenizálás pontosságát azonban a különböző időpontokban előforduló, azonos előjelű inhomogenitásokat okozó töréspontok is kedvezőtlenül befolyásolják, mivel a zaj és a szomszéd idősorok közeli töréspontjai együttesen hatnak a jelölt idősor töréspontjának detektálhatóságára, és

rontanak rajta. Az azonos típusú inhomogenitások hibát okoznak a regionális trendek becslésében. Legeredményesebben a páronkénti összehasonlítás és az ANOVA korrekció alkalmazásával csökkenthetők az ilyen típusú inhomogenitások hatásai, de teszt kísérletek azt mutatják, hogy meta-adatok felhasználása nélkül végzett homogenizálás esetén az ilyen típusú hibák jelentékeny hányada sajnos gyakran megmarad az adatokban.

iii) A meta-adatok felhasználását helyesebb lehetőségnek mint problémának tekinteni, noha a meta-adatok válogatása és automatikus vagy fél-automatikus homogenizáláshoz történő előkészítése többletmunkával jár. Az inhomogenitások felderítése szempontjából értékes meta-adatok leírják a technikai körülményekben bekövetkezett változás tartalmát és időpontját, de a technikai változás megfigyelési adatokra gyakorolt számszerű hatását az esetek többségében nem tartalmazzák. Kivétel ez alól, amikor az új és régi technikai feltételekkel két megfigyelési idősort hoznak létre ún. párhuzamos mérések végzésével (Böhm et al., 2010; Hannak et al., 2020), és a két idősor kiértékelése tartalmazza a technikai változás hatását. A párhuzamos mérések eredményei általában pontosabbak a statisztikai homogenizálás eredményeinél, de az egy bizonyos helyen végzett párhuzamos mérések nem minden esetben adaptálhatók más megfigyelőhelyek hasonló inhomogenitásainak a számszerűsítéséhez (Brunet et al., 2011). A párhuzamos mérésekből ismert inhomogenitás korrekciókat a statisztikai homogenizálás megkezdése előtt helyes elvégezni (a párhuzamos mérésekből ismert inhomogenitáson kívül egyéb inhomogenitások is terhelhetik az idősort, ezért a statisztikai homogenizálást nem hagyjuk el). Az egyéb meta-adatok közül azok a legfontosabbak, amelyek azonos típusú inhomogenitások több idősorban történő előfordulásáról szólnak. Ilyen esetekben szükséges lehet kézzel szerkeszteni olyan referencia idősorokat, amelyek a jelölt idősor inhomogenitásával megegyező típusú inhomogenitástól mentesek. Végül a „közönséges”, állomás-specifikus meta-adatok szintén hasznosak, és ezek automatikus homogenizálás keretén belül is felhasználhatók (Domonkos, 2022a).

5. Módszer-összehasonlító tesztek eredményei

5.1. Homogenizáló módszerek eredményességének tesztelése

A homogenizálás eredményességének értékeléséhez olyan szintetikus adatbázisokra van szükség, amelyek statisztikai tulajdonságai jól közelítik a megfigyelésből származó adatok tulajdonságait. Ez azért van így, mert a valós megfigyelési adatbázisok inhomogenitásait nem tudjuk teljes pontossággal megismerni, míg a szintetikusan fejlesztett adatbázisokba a fejlesztő maga illeszti be az ismert tulajdonságú inhomogenitásokat. A homogenizálás eredményessége sokféle statisztikai számítással jellemezhető (Venema et al., 2012; Killick, 2016), ezek közül azok tekinthetők a legfontosabbnak, amelyek közvetlenül jellemzik a homogenizált idősorok alkalmasságát az éghajlatváltozás és éghajlati variabilitás vizsgálatához. Ilyenek például a homogenizált idősorok maradék hibáit jellemző root mean square error (RMSE), átlagos abszolút hiba és a trend meredekségek eltérése a homogén adatok trendjeihez viszonyítva. Ilyen hibastatisztikákkal az egyedi idősorok és a regionális átlag idősorok homogenizálásának pontossága is jellemezhető.

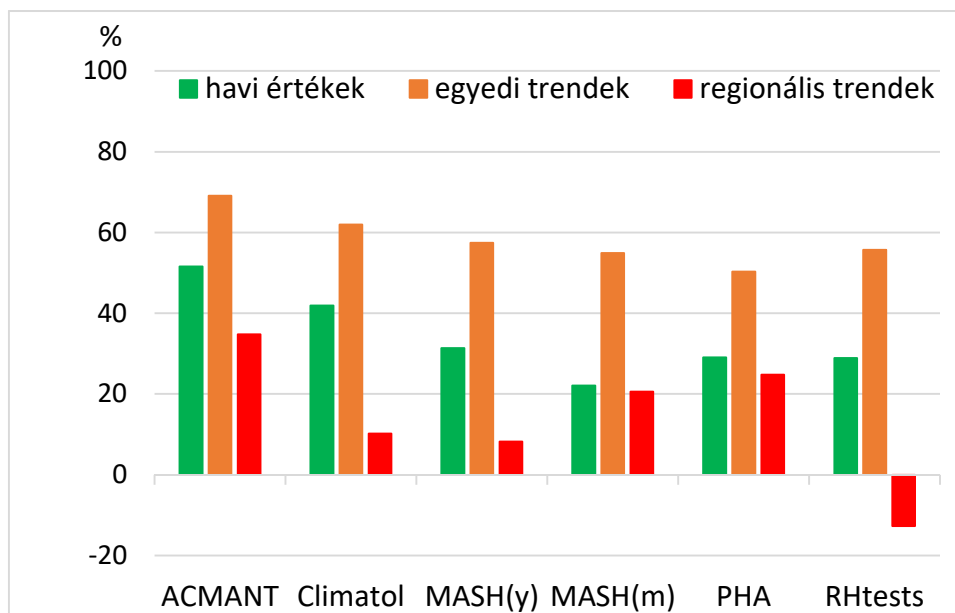
A COST ES0601 project („HOME”) tesztjeinél alkalmaztunk először olyan teszt adatbázisokat, amelyekkel komplett homogenizáló módszerek eredményessége értékelhető. A teszt adatbázis csupán 15 hőmérséklet és 15 csapadék adatmezőt tartalmazott, és gyengeség volt, hogy az adatmezők 60%-a esetében azok mindössze 5 idősort tartalmaztak. A teszteredmények azt mutatták, hogy a homogenizálás eredményessége erősen függ az alkalmazott módszertől (Venema et al., 2012). Egyes eredmények csalódást okozóan gyengék voltak, és a tesztelés tapasztalatai arra mutatnak, hogy megbízható minőségű statisztikai homogenizáláshoz feltétlenül szükséges a módszerek tesztelése. Néhány évvel a COST prozsekt befejeződése után Killick (2016) egy PhD program keretében hozott létre egy nagy méretű szintetikus adatbázist, amely az USA 4 régiójának szimulált napi hőmérsékleteit tartalmazza. A szintetikus homogén adatbázisok fejlesztése tárgyában ez a munka máig a legjelentősebb, de a független adatmezők száma (13) ebben az adatbázisban sem elegendő a regionális trendekben megmaradó hibák megalapozott értékeléséhez. Speciális célú szintetikus adatbázis fejlesztés történt az európai INDECIS prozsektben (INtegrated approach for the Development across Europe of user oriented Climate Indicators for global framework for climate Services high-priority sectors), ennek során Guijarro (2019) 9 éghajlati elem (maximumhőmérséklet, minimumhőmérséklet, csapadékmennyiség, relatív nedvesség, napos órák száma, légnyomás, szélsébség, felhőborítottság és hóvastagság) napi értékeit tartalmazó adatmezőit állította elő. A fejlesztés kíválóan szolgálta a prozsekt céljait, de az adatbázis használatánál figyelembe kell venni, hogy az adatok generálásánál

számos leegyszerűsítés történt, és a regionális trendek pontosságának értékeléséhez nem áll rendelkezésre elegendő számú adatmező. Szemben az eddig említettekkel, a spanyol MULTITEST prodzsektben (Guijarro et al., 2023) igen nagy számú szimulált adatmezőt használtunk fel a homogenizáló módszerek eredményességének értékeléséhez. Az adatmezők havi középhőmérséklet és csapadékösszeg adatokat tartalmaztak, statisztikai tulajdonságukról a kutatócsoport együttesen döntött. Az adatmezők generálását és a tesztek futtatását nagy részben José Guijarro végezte az általa kifejlesztett automatikus tesztelés programmal (Guijarro, 2011). E tesztelési módszernek van egy hátránya: az adatmezőket a program folyamatosan felülírja, azok tehát nem kerülnek megőrzésre, és így a regionális átlagok sem értékelhetők. A prodzsekt késői fázisában viszont megőrzésre került egy nagy méretű hőmérsékleti adatbázis (Domonkos et al., 2021), amely nem kevesebb mint 1900 adatmezőt tartalmaz. A MULTITEST prodzsekt eredményeinek bemutatásánál (5.2. alfejezet) elsősorban az ezen adatbázishoz tartozó eredményeket használjuk fel.

5.2. A MULTITEST prodzsekt eredményei

Nagy méretű adatbázisokon csak teljesen automatizált homogenizáló módszerek futtathatók. Mivel tesztelés nélkül a homogenizáló módszerek eredményessége bizonytalan, az interaktív módszereket is szükséges tesztelni, amely úgy történhet, hogy alap paraméterek helyettesítik a felhasználói beavatkozással megadható információkat. A MULTITEST prodzsektben 3 teljesen automatizált, 2 fél-automatikus és 1 a tesztelők által automatizált de eredetileg interaktív homogenizáló módszer összesen 13 változatát teszteltük (Guijarro et al., 2023). ACMANT (Domonkos, 2020), Climatol (Guijarro, 2021) és Pairwise Homogenization Algorithm (PHA, Menne and Williams, 2009) voltak az automatizált módszerek, MASH (Szentimrey, 1999) és RHtests (Chan & Feng, 2019) voltak a fél-automatikus módszerek és HOMER (Mestre et al., 2013) az eredetileg interaktív módszer. A MASH és az RHtests számára kizárólag a referencia idősorok megadásánál volt szükség a tesztelők beavatkozására (Domonkos et al., 2021). A megőrzött adatbázison végrehajtott homogenizálások értékeléséből kizártuk azokat a módszereket és módszer verziókat, amelyek részben vagy teljesen abszolút homogenizálás módban (különbség idősor képzése nélkül) működnek, így a MULTITEST prodzsekt ezen szegmensében nincs értékelés a HOMER módszerről és az RHtests abszolút verzióiról.

A megőrzött adatbázis 12 adat-szekcióból áll, amelyek egyenként 100-500 adatmezőt tartalmaznak. Az egyes szekciók különböznek egymástól az idősorok hosszában, a homogén adatok szimulálásának módjában, valamint a hozzáadott inhomogenitások és adathiányok statisztikai tulajdonságaiban (Domonkos et al., 2021). Négy adatszekcióban a jel-zaj arány kedvezőtlennek bizonyult, ezeknél alig vagy egyáltalán nem sikerült csökkenteni az inhomogenitásokat, a többi adatszekció homogenizálása viszont lényegesen sikeresebb volt. Az 5. ábra a teljes adatbázisra vonatkozó átlagolt eredményeket mutatja be.



5. ábra. MULTITEST eredmények 12 nagy méretű hőmérséklet teszt adatbázison: az inhomogenitások eltávolításának aránya az inhomogén adatok hibájának százalékában.

A vizsgált statisztikai tulajdonságok közül az egyedi idősorok trendjeinek homogenizálása volt a legsikeresebb, míg a regionális trendeké a legkevésbé sikeres. E tekintetben az eredmények hasonlóak a HOME eredményeihez (Venema et al., 2012). Megjegyzendő, hogy a hibák abszolút méretét tekintve a regionális trendek hibái lényegesen kisebbek az egyedi idősorokra húzott trendek hibáinál.

Az egyes homogenizáló módszereket tekintve, az 5. ábra a legsikeresebb módszer-variációk eredményeit mutatja be azokban az esetekben amikor ennek kiválasztása egyértelmű volt, míg a MASH-nál két verzió van ábrázolva, mivel az egyik az egyedi idősorok homogenizálásában, a másik a regionálisan átlagolt idősorok

homogenizálásában volt eredményesebb. A vizsgált módszerek közül ACMANT volt a legeredményesebb mind az egyedi idősorok, mind pedig a regionális átlag idősorok homogenizálásában. Második legjobbnak az egyedi idősorok homogenizálásában Climatol, míg a regionális sorok homogenizálásában PHA bizonyult. További MULTITEST eredmények (Guijarro et al., 2023) megerősítik az itt bemutatott különbségek fennállását a homogenizáló módszerek eredményességei között. ACMANT a csapadék idősorok homogenizálásában is a legeredményesebbnek bizonyult, viszont rövid és hosszú hőmérsékleti idősorok együttes homogenizálásánál Climatol azonosan pontos vagy kicsivel pontosabb eredményeket adott mint ACMANT. Néhány más módszer-összehasonlító tesztben (Killick, 2016; Guijarro, 2019) az ACMANT és Climatol módszerek technikailag azonos eredményeket adtak az egyedi idősorok homogenizálásánál.

A bemutatott MULTITEST eredmények összhangot mutatnak a módszerek elméleti tulajdonságaival. ACMANT fokozott eredményességét nagy részben magyarázza, hogy a tesztelt módszerek közül egyedül ACMANT alkalmazza az ANOVA korrekciót. A térbeli interpolációt alkalmazó Climatol eredményessége igen jó az egyedi idősorok homogenizálásánál, de regionális trendek esetleges hibáinak a korrigálására szinte teljesen alkalmatlan. Ennek az a magyarázata, hogy a Climatol módszer sok lépéses iterációval közelít a végső megoldáshoz, és ezek az iterációk a regionális átlag trendek hibáit inkább újraosztják az egyes idősorok között mintsem megszüntetnék azt. Az iteráció során az egyes idősorok variabilitása egyre hasonlóbba válik, de a konvergencia nem feltétlenül a tényleges éghajlati variabilitás felé tart (Domonkos et al., 2022). A PHA módszer egyetlen homogenizálási ciklust alkalmaz, és megfelelő jel-zaj arány esetén igen eredményes a regionális átlagok homogenizálásában. Viszonylag gyenge az eredményessége az egyedi idősorok trendjeinek korrigálásában, mert egyetlen homogenizálási ciklussal csak a viszonylag nagy méretű inhomogenitások korrigálhatók. A jel-zaj arány mértékére PHA különösen érzékeny, alacsony jel-zaj arány esetén a módszer rosszabb eredményeket ad más módszerekhez viszonyítva mind az egyedi idősorok mind pedig a regionális átlagok homogenizálásánál.

A MULTITEST eredmények azt mutatják, hogy valószínűleg ACMANT a legalkalmasabb homogenizáló módszer arra, hogy az inhomogenitások nagy részét mind az egyedi idősorokból mind pedig a regionálisan átlagolt idősorokból eltávolítsuk. Az ACMANT eredményességéhez az ensemble homogenizálás és a kétváltozós homogenizálás (amikor alkalmazható) is hozzájárulnak (Domonkos, 2020). Az

ACMANT számítógépes futtatása kevésbé időigényes a homogenizáló szoftverek többségéhez viszonyítva, és a MULTITEST prozsektben tesztelt ACMANTv4 verzió szabad hozzáférésű.

5.3. Újabb fejlesztések

A homogenizáló módszereknek vannak olyan új fejlesztései, amelyek a MULTITEST prozsekt idején még nem voltak tesztelhetők. Néhány ilyen fejlesztés rövid bemutatása következik.

Joelsson et al. (2022, 2023) a HOMER módszer automatizálásáról számoltak be. Az új módszer, amely a Bart nevet viseli, nemcsak havi, hanem napi idősorok homogenizálására is alkalmas, a Joint detection különbség idősorokon fut, és az első teszteredmények kedvezőek. A fejlesztés jelentősnek ígérkezik, mert a HOMER módszer bizonyítottan magas hatékonyságú szegmenseket tartalmaz (Mestre et al., 2013), és az R programmal előállított output produktumok hatékony interaktív homogenizálást tesznek lehetővé.

A MASHv4 verzió (Szentimrey, 2023) a variancia inhomogenitásainak olyan módon történő detektálását és korrigálását tartalmazza, amely nem érzékeny a szakasz-átlagok töréspontjai által okozott variancia változásokra. Ennek alkalmazása pontosabbá teszi a szakasz-átlagok töréspontjainak detektálását, és várhatóan pozitívan hat majd mind az átlagok mind pedig a valószínűségeloszlás egyéb statisztikai jellemzőinek a homogenizálására.

Az ACMANT fejlesztése folytatódik (Domonkos, 2021, 2022a), és az ACMANTv5 verzió már interaktív módban is futtatható. A több idősorban hasonló eltérést okozó inhomogenitások detektálása javult a kombinált idősor összehasonlítás módszer bevezetésével, és ebben a verzióban már a meta-adatok felhasználása is lehetséges, akár automatikus, akár interaktív üzemmódban.

6. Jövőben megoldandó feladatok

Törekedni kell a fejlesztések során elért eredmények gyakorlatba történő átültetésére, mert egyelőre meglehetősen gyakori, hogy kutatók olyan homogenizálási eljárásokat alkalmaznak, amelyek hatékonysága megkérdőjelezhető.

Visszatérünk kiindulási témánkhoz annak áttekintésével hogy hogyan hasznosulnak a homogenizáló módszerek a globális vagy kontinentális skálájú adatbázisok előállításában. Menne et al. (2018) előállították a Global Historical Climatology Network (GHCN) havi hőmérsékleti idősorainak egy újabb (negyedik) verzióját, és ennek során automatikus homogenizálásnak vetették alá a globális adatbázis mintegy ~26.000 idősorát. A homogenizálást a PHA módszerrel végezték, és szisztematikus, 0,1-0,3°C közötti pozitív eltérést detektáltak és korrigáltak az 1965-nél korábbi időszakok adataiban. Ez egy hatalmas előrelépés, mert korábbi globális adatbázis fejlesztéseknél homogenizálást vagy nem alkalmaztak, vagy csak korlátozott esetekben és erősen leegyszerűsített módszerek alkalmazásával homogenizáltak (Menne et al., 2012; Harris et al., 2020). Egy GHCN méretű adatbázis homogenizálása nagyon munkaigényes feladat még abban az esetben is ha a homogenizálás automatikus, de automatikus módszer alkalmazása nem minden esetben előnyös a homogenizálás eredményességére. Törekedni kell a meta-adatok felhasználására, noha ez a nemzeti meteorológiai szolgálatokkal közösen végzett munkát igényelne. Azok a meta-adatok fontosak, amelyek egy megfigyelőhálózatban történt azonos típusú technikai változásról szólnak, ezen kívül pedig azok, amelyek térben ritka adatmezők homogenizálását segítik. Másrészt, térben kevésbé sűrű adatmezők homogenizálásához a PHA módszer használata nem előnyös. Ezek figyelembevételével a globális és kontinentális idősorok homogenitása a jövőben tovább javítható. Mi várható egy precízebb homogenizálástól? Valószínű, hogy nagyobb mértékű szisztematikus pozitív eltérés válik ismertté mint az eddigiekben feltárt 0.1-0.3°C. Ez a homogenizálás tesztek azon eredményéből következik, hogy amennyiben a homogenizálatlan adatmező szisztematikus trend differenciát tartalmaz, az automatikus homogenizálások túlnyomó hányadában a kezdeti differenciák egy jelentékeny hányada megmarad. Az ezen alapuló becslés csak egy feltételezés, amelyet csak alaposabb homogenizálással lehetséges ellenőrizni. Az inhomogenitások miatti pontatlanság nem becsülhető kielégítő biztonsággal a különböző adatbázisok összehasonlításával (Rao et al., 2018), mivel az inhomogenitások hatásai valószínűleg hasonlóak az egyes adatbázisokban. Nem ad kielégítő biztonságú megoldást a konfidencia intervallumok Menne et al. (2018) által bemutatott szerkesztése sem, mert a számítások részben önkényes feltételezéseken alapulnak, és ezek nem kiszámítható mértékű hibát okoznak az intervallumok becslésében.

Természetesen a kisebb méretű adatbázisok pontos homogenizálása is fontos, és jelenlegi ismeretek szerint azokhoz is elsősorban az ACMANT módszert kellene

alkalmazni. Lehetséges azonban, hogy újabb, más földrajzi területekre érvényes vagy más éghajlati változókat szimuláló adatbázisok a MULTITEST eredményeitől lényegesen eltérő teszt eredményeket adnának. Ezt csak akkor fogjuk megtudni, ha lesznek ilyen tesztek. Tesztek időnkénti megújítása azért is szükséges, hogy ellenőrizzük az újabb módszerfejlesztések eredményességét.

A jelenlegi feltételek nem nagyon kedveznek a homogenizálás fejlődésének. Több mint 10 éve nincs olyan nemzetközi prozsekt amely célzottan támogatná ezt a kutatási témát, és a homogenizáló módszerek tesztelésében tapasztalt kutatók fokozatosan kiöregednek. José Guijarro már nyugdíjban van, Victor Venema sajnos korán elhunyt, és Szerző sem a most fiatalok gárdájához tartozik. Kik fogják folytatni?

Irodalom

Aguilar, E., & Prohom, M. (2010). *EXTRAQC quality control software*. Centre for Climate Change, Universitat Rovira i Virgili, Tarragona, Spain.
http://www.c3.urv.cat/data/Manual_rclimindex_extraQC.r.pdf

Alexandersson, H. (1986). A homogeneity test applied to precipitation data. *J. Climatol*, 6, 661–675.

Bokros K., & Lakatos M. (2022). Hőségperiódusok vizsgálata Magyarországon a XX. század elejétől napjainkig. *Léggör*, 67, 130-140.
<https://doi.org/10.56474/legkor.2022.3.2>

Böhm, R., Jones, P. D., Hiebl, J., Frank, D., Brunetti, M., & Maugeri, M. (2007). The early instrumental warm-bias: a solution for long central European temperature series 1760–2007. *Climatic Change*, 101, 41–67. <https://doi.org/10.1007/s10584-009-9649-4>

Brunet, M., Asin, J., Sigró, J., Bañón, M., García, F., Aguilar, E., Palenzuela, J. E., Peterson, T. C., & Jones, P. (2011). The minimization of the screen bias from ancient Western Mediterranean air temperature records: an exploratory statistical analysis. *Int. J. Climatol*, 31, 1879-1895. <https://doi.org/10.1002/joc.2192>, 2010

Caussinus, H., & Mestre, O. (2004). Detection and correction of artificial shifts in climate series. *JR. Stat. Soc. Ser. C Appl. Stat*, 53, 405–425.

<http://doi.org/10.1111/j.1467-9876.2004.05155.x>

Chan, R., & Feng, Y. (2019). *The RHtests homogenization package*.

<https://github.com/ECCC-CDAS/RHtests>

Craddock, J. M. (1979). Methods of comparing annual rainfall records for climatic purposes. *Weather*, 34, 332–346.

Domonkos, P. (2011). Efficiency evaluation for detecting inhomogeneities by objective homogenisation methods. *Theor. Appl. Climatol*, 105, 455–467.

<http://doi.org/10.1007/s00704-011-0399-7>

Domonkos, P. (2020). *ACMANTv4: Scientific content and operation of the software*.

71pp.

https://github.com/dpeterfree/ACMANT/blob/ACMANTv4.4/ACMANTv4_description.pdf

Domonkos, P. (2021). Combination of using pairwise comparisons and composite reference series: a new approach in the homogenization of climatic time series with ACMANT. *Atmosphere*, 12(9), 1134. <https://doi.org/10.3390/atmos12091134>

Domonkos, P. (2022a). Automatic homogenization of time series: How to use metadata? *Atmosphere*, 13(9), 1379. <https://doi.org/10.3390/atmos13091379>

Domonkos P. (2022b). *Éghajlati idősorok homogenizálása: Módszertani fejlődés 1996–2021 között*. In: Tanulmánykötet Dr. Makra László professzor 70. születésnapjára, 172–196. https://github.com/dpeterfree/ACMANT/blob/Hungarian_page/

Domonkos, P., Guijarro, J. A., Venema, V., Brunet, M., & Sigró, J. (2021). Efficiency of time series homogenization: method comparison with 12 monthly temperature test datasets. *J. Climate*, 34, 2877–2891. <https://doi.org/10.1175/JCLI-D-20-0611.1>

Domonkos, P., & Joelsson, L. M. T. (2023). ANOVA correction in relative homogenization: why it is indispensable. Submitted to *Időjárás*.

Domonkos, P., Tóth, R., & Nyitrai, L. (2022). *Climate observations: Data quality control and time series homogenization*. Elsevier, 302pp.
<https://www.elsevier.com/books/climate-observations/domonkos/978-0-323-90487-2>

Durre, I., Menne, M. J., Gleason, B. E., Houston, T. G., & Vose, R.S. (2010). Comprehensive automated quality assurance of daily surface observations. *J. Appl. Meteorol. Climatol*, 49, 1615-1633. <https://doi.org/10.1175/2010JAMC2375.1>

Easterling, D. R., & Peterson, T. C. (1995). A new method for detecting undocumented discontinuities in climatological time series. *Int. J. Climatol*, 15, 369–377.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, *Geosci. Model Dev*, 9, 1937–1958.
<https://doi.org/10.5194/gmd-9-1937-2016>, 2016

Guijarro, J. A. (2011). Influence of network density on homogenization performance. In (eds. Lakatos, M., Szentimrey, T., & Vincze, E.) Seventh Seminar for Homogenization and Quality Control in Climatological Databases, *WMO-WCDMP 78*, 11-18, Geneva, Switzerland.

Guijarro, J. A. (2019). *Recommended homogenization techniques based on benchmarking results. WP-3 report of INDECIS project*.
http://www.indecis.eu/docs/Deliverables/Deliverable_3.2.b.pdf

Guijarro, J. A. (2021). *Homogenization of climatic series with Climatol*.
https://www.climatol.eu/homog_climatol-en.pdf

Guijarro, J. A., López, J. A., Aguilar, E., Domonkos, P., Venema, V. K. C., Sigró, J., & Brunet, M. (2023). Homogenization of monthly series of temperature and precipitation: Benchmarking results of the MULTITEST project. *Int. J. Climatol*, *43*, 3994–4012. <https://doi.org/10.1002/joc.8069>

Hannak, L., Friedrich, K., Imbery, F., & Kaspar, F. (2020). Analyzing the impact of automatization using parallel daily mean temperature series including breakpoint detection and homogenization. *Int. J. Climatol*, *40*, 6544–6559. <https://doi.org/10.1002/joc.6597>

Harris, I., Osborn, T. J., Jones, P., & Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Sci. Data*, *7*, 109. <https://doi.org/10.1038/s41597-020-0453-3>

Intergovernmental Panel on Climate Change (1990). *FAR Climate Change: Scientific Assessment of Climate Change*.

Intergovernmental Panel on Climate Change (2021). *Climate Change 2021: The physical science basis*. Working Group I contribution to the IPCC Sixth Assessment Report.

Joelsson, L. M. T., Engström, E., & Kjellström, E. (2023). Homogenization of Swedish mean monthly temperature series 1860–2021. *Int. J. Climatol*, *43*(2), 1079–1093. <https://doi.org/10.1002/joc.7881>

Joelsson, L. M. T., Sturm, C., Södling, J., Engström, E., & Kjellström, E. (2022). Automation and evaluation of the interactive homogenization tool HOMER. *Int. J. Climatol*, *42*(5), 2861–2880. <https://doi.org/10.1002/joc.7394>

Kharin, V. V., & Zwiers, F. W. (2000). Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere-ocean GCM. *J. Climate*, *13*, 3760–3788.

Killick, R. E. (2016). *Benchmarking the performance of homogenisation algorithms on daily temperature data*. PhD thesis, University of Exeter, UK.
<https://ore.exeter.ac.uk/repository/handle/10871/23095>

Lakatos M., Bihari Z., Izsák B., Marton A., & Szentes O. (2021). Megfigyelt éghajlati változások Magyarországon. *Légkör*, 66/3, 5-11.

Lindau, R., & Venema, V. K. C. (2018a). On the reduction of trend errors by the ANOVA joint correction scheme used in homogenization of climate station records. *Int. J. Climatol*, 38, 5255–5271. <http://doi.org/10.1002/joc.5728>

Lindau, R., & Venema, V. K. C. (2018b). The joint influence of break and noise variance on the break detection capability in time series homogenization. *Adv. Stat. Clim. Meteorol. Oceanogr*, 4, 1–18. <https://doi.org/10.5194/ascmo-4-1-2018>

Lyon, B., & Dole, R. M. (1995). A diagnostic comparison of the 1980 and 1988 U.S. summer heat wave-droughts, *J. Climate*, 8(6), 1658-1675. [https://doi.org/10.1175/1520-0442\(1995\)008<1658:ADCOTA>2.0.CO;2](https://doi.org/10.1175/1520-0442(1995)008<1658:ADCOTA>2.0.CO;2)

MedECC (2020). *Climate and environmental change in the Mediterranean Basin – current situation and risks for the future*. First Mediterranean Assessment Report (eds. Cramer, W., Guiot, J., & Marini, K.). Union for the Mediterranean, Plan Bleu, UNEP/MAP, Marseille, France, 632 pp. <https://doi.org/10.5281/zenodo.4768833>

Megyeri-Korotaj, O. A., Bán, B., Suga, R., Allaga-Zsebeházi, G., & Szépszó, G. (2023). Assessment of climate indices over the Carpathian Basin based on ALADIN5.2 and REMO2015 regional climate model simulations. *Atmosphere*, 14, 448.
<https://doi.org/10.3390/atmos14030448>

Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E., & Houston, T. G. (2012). An overview of the Global Historical Climatology Network-Daily database. *J. Atmos. Oceanic Technol*, 29, 897-910. <https://doi.org/10.1175/JTECH-D-11-00103.1>

Menne, M. J., & Williams Jr, C. N. (2009). Homogenization of temperature series via pairwise comparisons. *J. Climate*, 22, 1700–1717.

<https://doi.org/10.1175/2008JCLI2263.1>

Menne, M. J., Williams, C. N., Gleason, B. E., Rennie, J. J., & Lawrimore, J. H. (2018). The Global Historical Climatology Network monthly temperature dataset, version 4. *J. Climate*, 31, 9835–9854. <https://doi.org/10.1175/JCLI-D-18-0094.1>

Mestre, O., Domonkos, P., Picard, F., Auer, I., Robin, S., Lebarbier, E., Böhm, R., Aguilar, E., Guijarro, J. A., Vertacnik, G., Klancar, M., Dubuisson, B., & Štěpánek, P. (2013). HOMER: homogenization software in R – methods and applications, *Időjárás*, 117, 47–67.

Mika, J., Károssy, Cs., & Lakatos, L. (2021). Variations in the Peczely macro-synoptic types (1881–2020) with attention to weather extremes in the Pannonian Basin.

Atmosphere, 12, 1071. <https://doi.org/10.3390/atmos12081071>

Moberg, A., & Alexandersson, H. (1997). Homogenization of Swedish temperature data. Part II: Homogenized gridded air temperature compared with a subset of global gridded air temperature since 1861. *Int. J. Climatol*, 17, 35-54.

Páscoa, P., Gouveia, C. M., Russo, A., & Ribeiro, A. F. S. (2022). Summer hot extremes and antecedent drought conditions in Australia. *Int. J. Climatol*, 42(11), 5487–5502. <https://doi.org/10.1002/joc.7544>

Peterson, T. C., & Easterling, D. R. (1994). Creation of homogeneous composite climatological reference series. *Int. J. Climatol*, 14, 671–679.

Picard, F., Lebarbier, E., Hoebeke, M., Rigai, G., Thiam, B., & Robin, S. (2011). Joint segmentation, calling, and normalization of multiple CGH profiles. *Biostatistics*, 12, 413–428. <https://doi.org/10.1093/biostatistics/kxq076>

Prohom, M., Domonkos, P., Cunillera, J., Barrera-Escoda, A., Busto, M., Herrero-Anaya, M., Aparicio, A., & Reynés, J. (2023). CADTEP: A new daily quality-controlled and homogenized climate database for Catalonia (1950–2021). *Int. J. Climatol*, 43, 4771–4789. <https://doi.org/10.1002/joc.8116>

Rao, Y., Liang, S., & Yu, Y. (2018). Land surface air temperature data are considerably different among BEST-LAND, CRU-TEM4v, NASA-GISS, and NOAA-NCEI. *J. Geophys. Res. Atmos*, 123, 5881–5900. <https://doi.org/10.1029/2018JD028355>

Rustemeier, E., Kapala, A., Meyer-Christoffer, A., Finger, P., Schneider, U., Venema, V., Ziese, M., Simmer, C., & Becker, A. (2017). AHOPS Europe – a gridded precipitation data set from European homogenized time series. In (eds. Szentimrey, T., Lakatos, M., & Hoffmann, L.) Ninth Seminar for Homogenization and Quality Control in Climatological Databases. *WMO-WCDMP*, 85, 88–101, Geneva, Switzerland.

Santer, B. D., Po-Chedley, S., Mears, C., Fyfe, J. C., Gillett, N., Fu, Q., Painter, J. F., Solomon, S., Steiner, A. K., Wentz, F. J., Zelinka, M. D., & Zou, C-Z. (2021). Using climate model simulations to constrain observations. *J. Climate*, 34, 6281–6301. <https://doi.org/10.1175/JCLI-D-20-0768.1>

Shen, S. S. P., Yin, H., & Smith, T. M. (2007). An estimate of the sampling error variance of the gridded GHCN monthly surface air temperature data. *J. Climate*, 20, 2321–2331. <https://doi.org/10.1175/JCLI4121.1>

Squintu, A. A., van der Schrier, G., Brugnara, Y., & Klein Tank, A. (2019). Homogenization of daily temperature series in the European Climate Assessment & Dataset. *Int. J. Climatol*, 39, 1243–1261. <https://doi.org/10.1002/joc.5874>

Szabó, A. I., Breuer, H., Ács, F., Belda, M., & Feddema, J. (2022). Projected changes in Feddema climate characteristics in the Larger Carpathian Region by the end of the 21st century. *Int. J. Climatol*, 42(11), 5732–5747. <https://doi.org/10.1002/joc.7558>

Szentimrey, T. (1999). Multiple Analysis of Series for Homogenization (MASH). In (eds. Szalai, S., Szentimrey, T., & Szinell, Cs.) Second Seminar for Homogenization of Surface Climatological Data. *WMO-WCDMP*, 41, 27-46, Geneva, Switzerland.

Szentimrey, T. (2023). Development of new version MASHv4.01 for homogenization of standard deviation. In (eds. Lakatos, M., Szentimrey, T., & T. Puskás, M.) Eleventh Seminar for Homogenization and Quality Control in Climatological Databases, *WMO-WCDMP*, 87, 8-13, Budapest, Hungary.

Van Malderen, R., Pottiaux, E., Klos, A., Domonkos, P., Elias, M., Ning, T., Bock, O., Guijarro, J. A., Alshawaf, F., Hoseini, M., Quarello, A., Lebarbier, E., Chimani, B., Tornatore, V., Zengin Kazancı, S., & Bogusz, J. (2020). Homogenizing GPS integrated vapor time series: benchmarking break detection methods on synthetic datasets. *Earth and Space Sci*, 7(5):e2020EA001121. <https://doi.org/10.1029/2020EA001121>

Venema, V., Mestre, O., Aguilar, E., Auer, I., Guijarro, J. A., Domonkos, P., Vertacnik, G., Szentimrey, T., Štěpánek, P., Zahradníček, P., Viarre, J., Müller-Westermeier, G., Lakatos, M., Williams, C. N., Menne, M., Lindau, R., Rasol, D., Rustemeier, E., Kolokythas, K., Marinova, T., Andresen, L., Acquafredda, F., Fratianni, S., Cheval, S., Klancar, M., Brunetti, M., Gruber, C., Duran, M. P., Likso, T., Esteban, P., & Brandsma, T. (2012). Benchmarking monthly homogenization algorithms. *Climate of the Past*, 8, 89-115. <https://doi.org/10.5194/cp-8-89-2012>

Vicente-Serrano, S. M., Nieto, R., Gimeno, L., Azorin-Molina, C., Drumond, A., El Kenawy, A., Dominguez-Castro, F., Tomas-Burguera, M., & Peña-Gallardo, M. (2018). Recent changes of relative humidity: Regional connections with land and ocean processes. *Earth System Dynamics*, 9(2), 915-937. <https://doi.org/10.5194/esd-9-915-2018>

Wang, X. L., Wen, Q. H., & Wu, Y. (2007). Penalized maximal t test for detecting undocumented mean change in climate data series. *J. Appl. Meteor. Climatol*, 46, 916-931. <https://doi.org/10.1175/JAM2504.1>

Mika János és én soha nem voltunk közvetlen munkatársak szervezeten belül, mégis nem kevesebb mint 10 olyan publikációnk született (1992-2000), amelyeket közösen írtunk. Lelkesek voltunk, szerettük a szakmát, és szerettünk együtt dolgozni. A 90-es évek elején János már tapasztalt kutató volt, én pedig még csak kezdő. János gyakran munkaidőn túl is bent maradt velem a munkahelyén, azért hogy engem a kutatói munka fortélyaira tanítson. A János iránt érzett hálával és kis nosztalgiával gondolok vissza azokra az időkre.