

Nagy adatbázisok alkalmazása és kezelése a jövőbeli éghajlatváltozás kutatásában

Szépszó Gabriella (szepszo.g@met.hu)

Krüzselyi Iлона, Szabó Péter, Zsebeházi Gabriella

Országos Meteorológiai Szolgálat
Éghajlati Osztály, Klímamodellező Csoport



Big Data Day
2013. szeptember 12.

TARTALOM

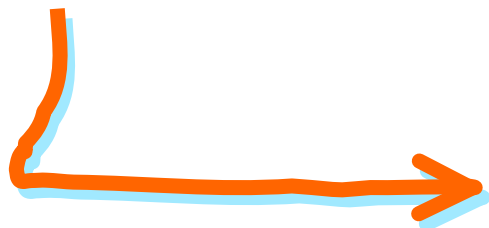
1. Motiváció
2. Éghajlati modellezés
3. BigData vonatkozások
4. Felhasználás, kitekintés

Motiváció

- Az éghajlatváltozás hatásainak ismerete és az azokra való felkészülés fontos a társadalom és a gazdaság szempontjából
- Modellekkel lehetséges a jövőbeli éghajlati viszonyok becslése
- A meteorológiai modellezés számításigényes folyamat
- Elengedhetetlen a szimulációs bizonytalanságok számszerűsítése → ez még költségesebbé teszi a modellezést
- Egy éghajlati szimulációban rengeteg adat keletkezik

Motiváció

- Az éghajlatváltozás hatásainak ismerete és az azokra való felkészülés fontos a társadalom és a gazdaság szempontjából
- Modellekkel lehetséges a jövőbeli éghajlati viszonyok becslése
- A meteorológiai modellezés számításigényes folyamat
- Elengedhetetlen a szimulációs bizonytalanságok számszerűsítése → ez még költségesebbé teszi a modellezést
- Egy éghajlati szimulációban rengeteg adat keletkezik



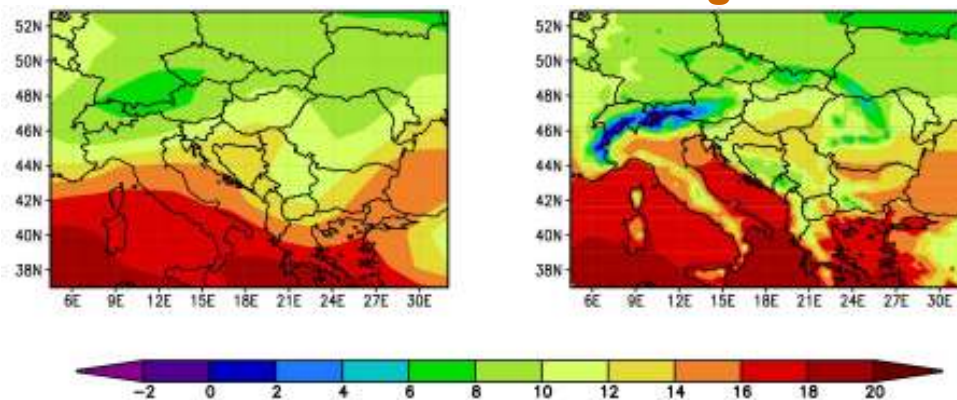
BigData vonatkozások:

1. Számítási kapacitás
2. Adattárolás

Regionális sajátosságok

- Globális modellek: 250-100 km-es rácssűrűség
- Finomítás (leskálázás): regionális éghajlati modellekkel
 - Kisebb terület, nagyobb felbontás: 10-25 km
 - Fizikai folyamatok és felszíni jellemzők pontosabb leírása
 - Tartományon kívüli folyamatok figyelembevétele

Átlaghőmérséklet [$^{\circ}\text{C}$], 1961–1990
Globális
Regionális



A robotika 3 törvénye

Éghajlati szimulációk készítésének 3 törvénye

Éghajlati szimulációk készítésének 3 törvénye

0. Az éghajlat csak egy hosszabb időszak átlagában értelmezhető

Éghajlati szimulációk készítésének 3 törvénye

0. Az éghajlat csak egy hosszabb időszak átlagában értelmezhető
1. A modelleket először a múltra vonatkozóan tesztelik (összehasonlítás mérésekkel) → az éghajlati szimulációk a múltban kezdődnek

Éghajlati szimulációk készítésének 3 törvénye

0. Az éghajlat csak egy hosszabb időszak átlagában értelmezhető
1. A modelleket először a múltra vonatkozóan tesztelik (összehasonlítás mérésekkel) → az éghajlati szimulációk a múltban kezdődnek
2. Elvárt pontosság: egy hosszabb időszak éghajlati viszonyainak reprezentálása → a szimulációk több (legalább három) évtizeddel a jelen állapot előtt kezdődnek

Éghajlati szimulációk készítésének 3 törvénye

0. Az éghajlat csak egy hosszabb időszak átlagában értelmezhető
1. A modelleket először a múltra vonatkozóan tesztelik (összehasonlítás mérésekkel) → az éghajlati szimulációk a múltban kezdődnek
2. Elvárt pontosság: egy hosszabb időszak éghajlati viszonyainak reprezentálása → a szimulációk több (legalább három) évtizeddel a jelen állapot előtt kezdődnek
3. Projekciók a jövőre: az éghajlatváltozási jel szignifikanciája – a vizsgált jövőbeli és a referencia-időszak megfelelő távolsága → a projekciók gyakran több, mint egy évszázadot ölelnek fel

Gyakorlati példa

A légkör esetében megoldandó nem-lineáris parciális differenciál-egyenletrendszer:

Mozgásegyenletek

$$\frac{d\bar{v}}{dt} = -\frac{1}{\rho} \nabla p + \bar{g} - 2\bar{\Omega} \times \bar{v} + \bar{F} + \bar{S}$$

Tömeg-megmaradás

$$\frac{d\rho}{dt} = -\rho \cdot \operatorname{div} \bar{v}$$

Energia-megmaradás

$$\frac{dQ}{dt} = c_p \frac{dT}{dt} - \alpha \frac{dp}{dt}$$

Nedvesség tömeg-megmaradása

$$\frac{dq}{dt} = -\frac{1}{\rho} \cdot M$$

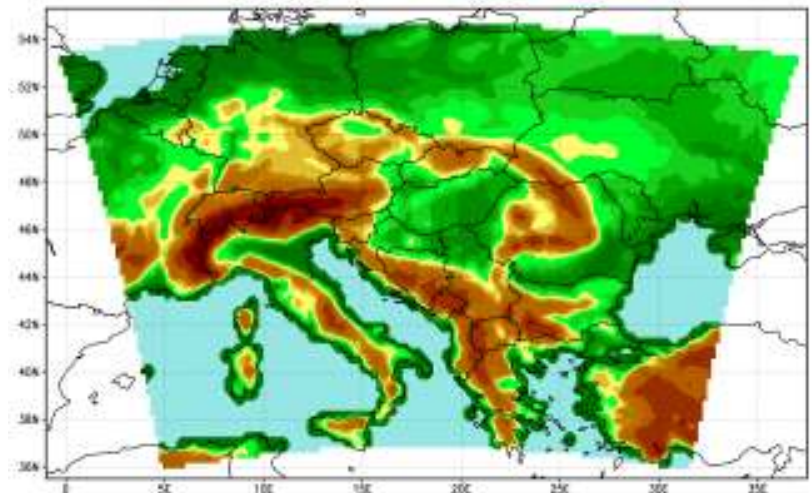
Állapotegyenlet

$$p = \rho RT$$

A szimulációk adatmennyisége

- Az OMSZ-ban két jövőbeli időszakot vizsgálunk: 2021–2050 és 2071–2100; a referencia-időszak: 1961–1990
- Éghajlati szimulációk az **1951–2100 időszakra** (globális modellek: 1860-tól)
- Az OMSZ-ban néhány évvel ezelőtt egy nagyobb tartományra és **25 km-es felbontás** alkalmazására volt lehetőség:

81 pont
x
101 pont
x
20 szint
=
163 620 rácspont



A szimulációk adatmennyisége

- A meteorológiai előrejelzéseknél legalább **5 változót** (hőmérséklet, nedvesség, légnyomás, áramlási sebesség horizontális komponensei) írunk le a kiválasztott tartomány minden rácspontjában
- Az előrejelzés (a differenciál-egyenletek integrálása) **kis lépésekben** történik:

A szimulációk adatmennyisége

- A meteorológiai előrejelzéseknél legalább **5 változót** (hőmérséklet, nedvesség, légnyomás, áramlási sebesség horizontális komponensei) írunk le a kiválasztott tartomány minden rácspontjában
- Az előrejelzés (a differenciál-egyenletek integrálása) **kis lépésekben** történik:

Időjárás: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
0h 1h

24 óra:
720 lépés

A szimulációk adatmennyisége

- A meteorológiai előrejelzéseknek legalább **5 változót** (hőmérséklet, nedvesség, légnyomás, áramlási sebesség horizontális komponensei) írunk le a kiválasztott tartomány minden rácspontjában
- Az előrejelzés (a differenciál-egyenletek integrálása) **kis lépésekben** történik:



24 óra:
720 lépés



150 év:
~40 millió lépés

A szimulációk adatmennyisége

- Lépésenként összesen ~800 000, szimulációnként 3×10^{13} adat
- Mit tartunk meg? Szempontok:
 - Mit szeretnénk vizsgálni?
 - Fel szeretnénk-e még használni az ebből kiinduló futtatásokhoz (további leskálázáshoz)?
 - Szükséges-e részleges/teljes újrafuttatás?



3-6 óránkénti közbülső file-ok és nyers outputok: ~1 TB

- Tárolni kell még:
 - Határfeltételek (globális mezők)
 - Egyéb változók
 - Feldolgozott adatok



4-5 TB

Az OMSZ számítástechnikai kapacitása

Szuperszámítógép

- Szempontok:
 - Éghajlati modellezés céljára dedikált szuperszámítógép (ne zavarja az operatív alkalmazásokat)
 - Gyors számítási sebesség
- SGI Altix 3700 cluster
- 92 Intel Itanium-2 CPU, 2 GB RAM/CPU, 900 GFlop/s
- MPI kommunikáció (26 Gbps), PBS job scheduler
- Az előbbi szimuláció 5-6 hónap alatt fut le 32 processzoron

Archiváló rendszer

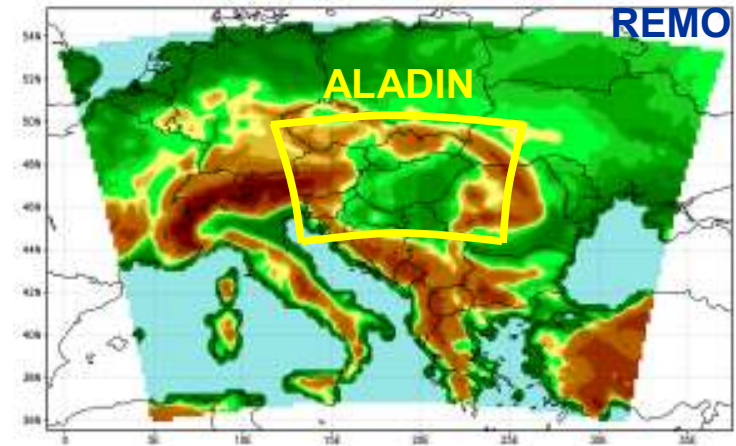
- Szempont: viszonylag gyors elérés
- Szalagos archiváló rendszer
- Kapacitás: ~200 TB
- Folyamatos bővítés



A szimulációk bizonytalanságai

- Többféle bizonytalanság a szimulációkban: modellek eltérései, emberi tevékenység kiszámíthatatlansága
- Számszerűsítés: **több modellkísérlet** együttese
- Az OMSZ-ban kutatások két adaptált modellel:

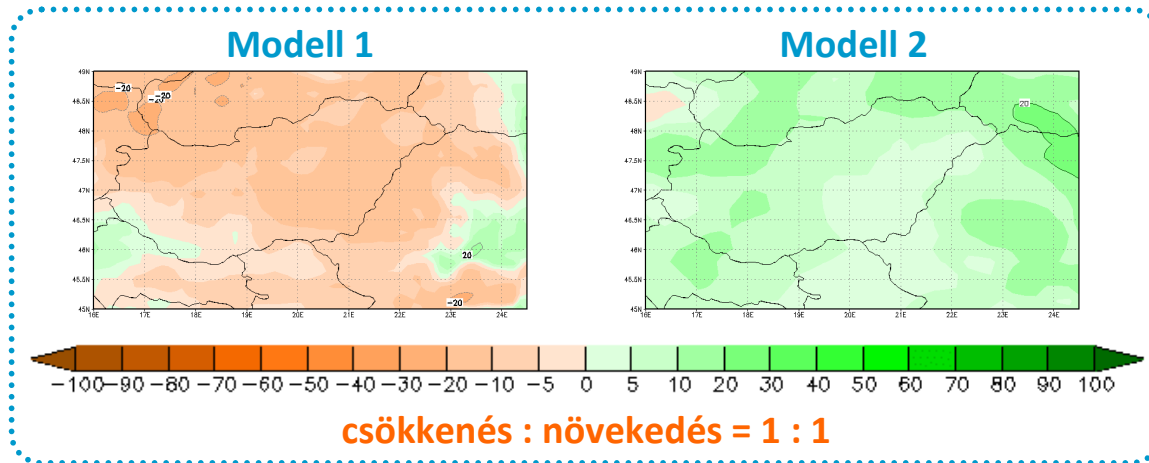
Modell	ALADIN	REMO
Felbontás	10 km	25 km
Időszak	1961–2100	1951–2100
Forgatókönyv	A1B	A1B



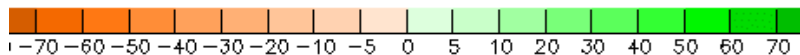
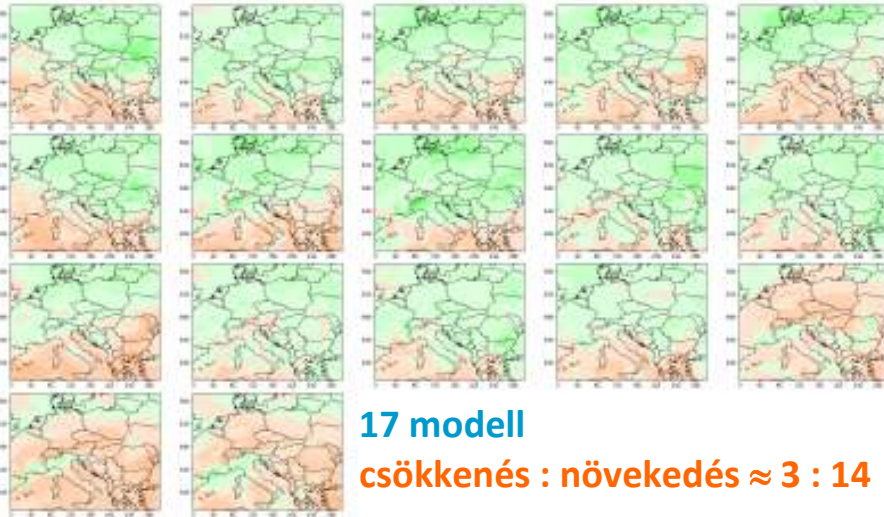
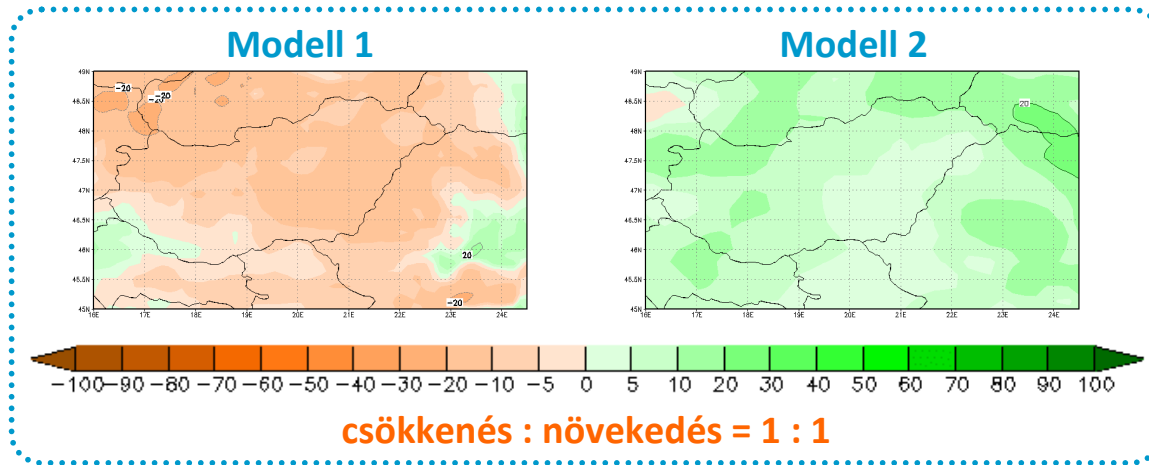
+ Európai modelleredmények felhasználása (25 km-es felbontás)

Téli csapadékváltozás [%], 2021–2050

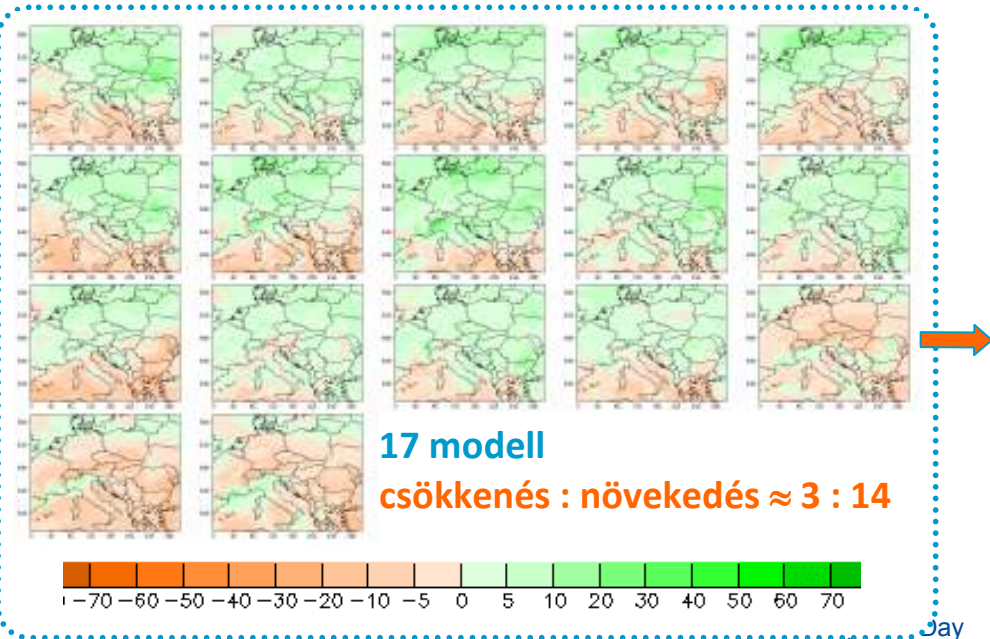
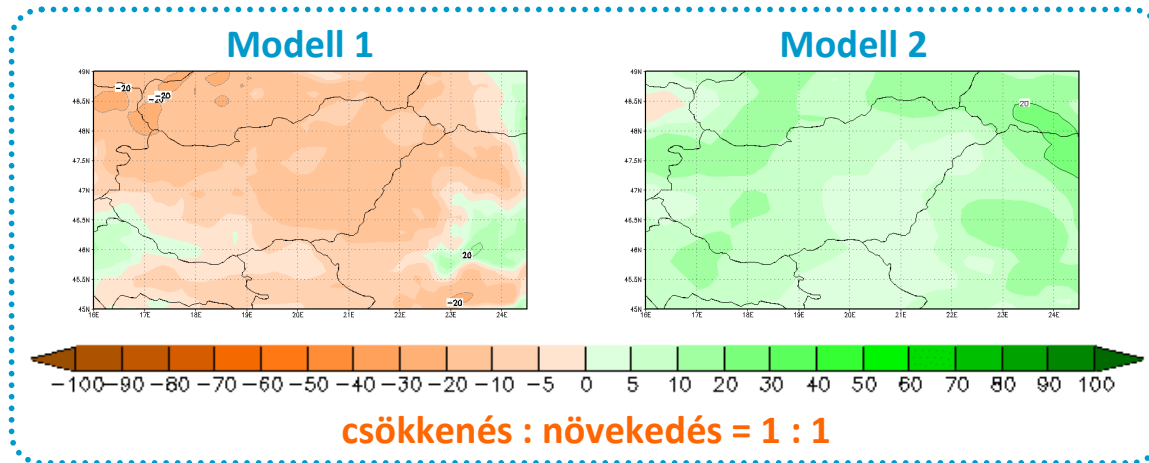
Téli csapadékváltozás [%], 2021–2050



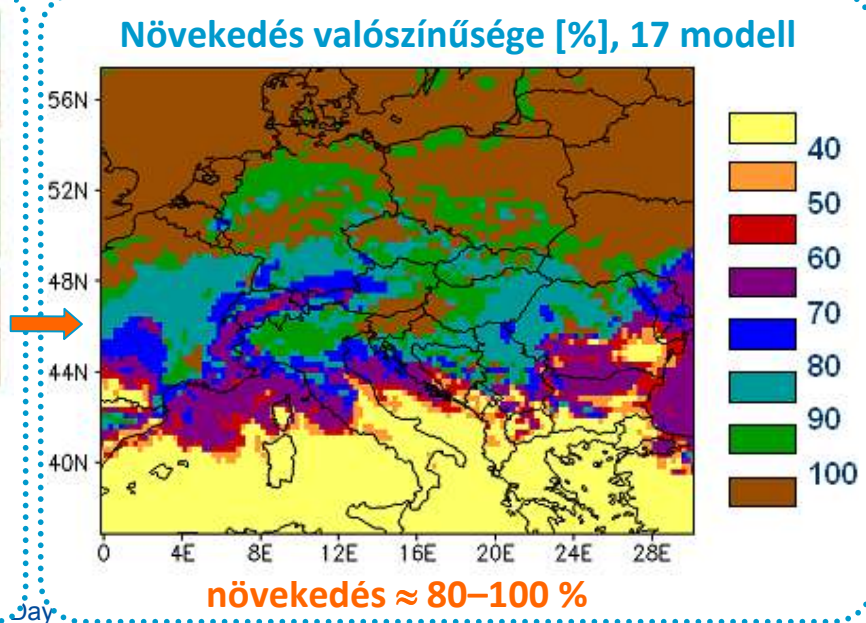
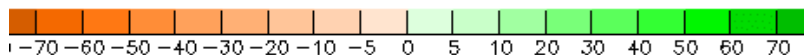
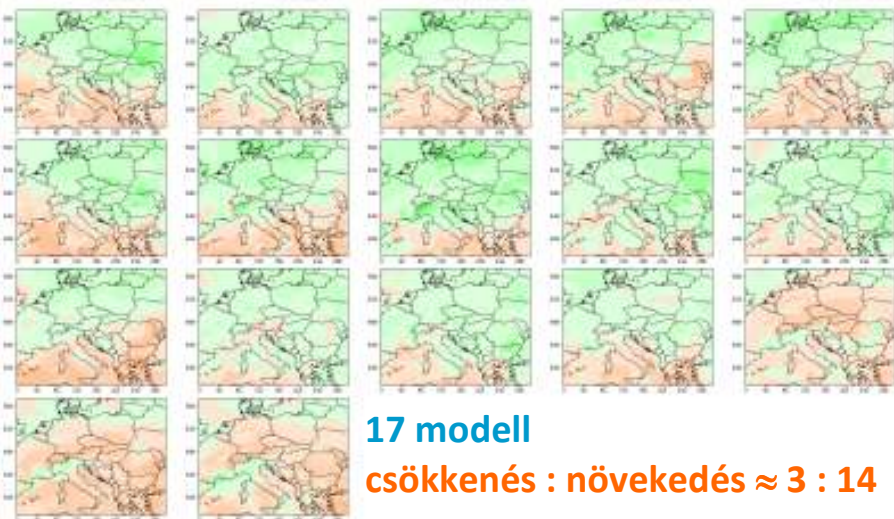
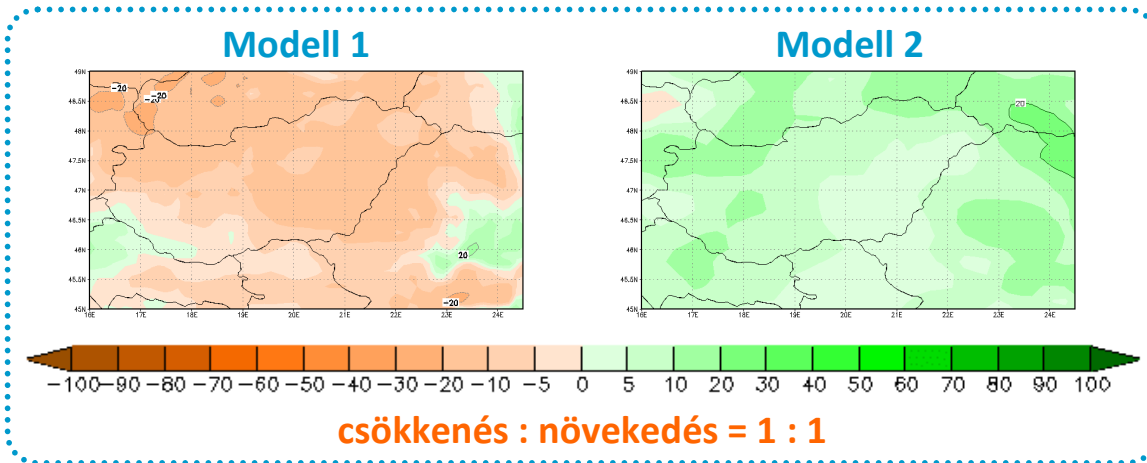
Téli csapadékváltozás [%], 2021–2050



Téli csapadékváltozás [%], 2021–2050



Téli csapadékváltozás [%], 2021–2050



Felhasználás

- **Objektív** döntéshozatal: számszerű **hatásvizsgálatok**
- Kiindulás: **modelleredmények + bizonytalanságok**
- Példa: városklimatológiai hatásvizsgálatok dinamikus leskálázással (egyben témánk a VKSZ12-ben)
 - Tesztkísérletek Budapestre 1 km-es felbontással
 - Újabb modellszimuláció, megfelelő frekvenciájú és felbontású input-adatigény, nagyfelbontású outputok
 - Az eredmények felhasználhatók a hosszútávú tervezésben

Felhasználás

- **Objektív** döntéshozatal: számszerű **hatásvizsgálatok**
- Kiindulás: **modelleredmények + bizonytalanságok**
- Példa: városklimatológiai hatásvizsgálatok dinamikus leskálázással (egyben témánk a VKSZ12-ben)
 - Tesztkísérletek Budapestre 1 km-es felbontással
 - Újabb modellszimuláció, megfelelő frekvenciájú és felbontású input-adatigény, nagyfelbontású outputok
 - Az eredmények felhasználhatók a hosszútávú tervezésben

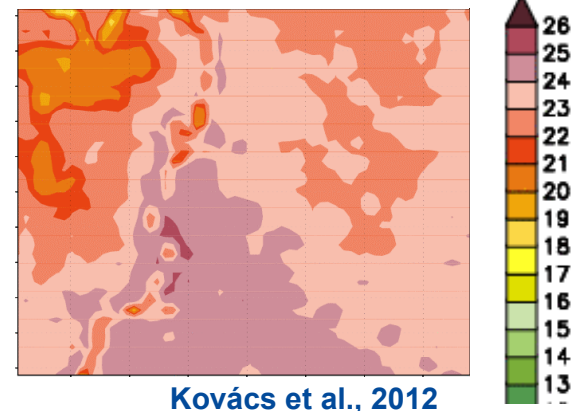
Nyári átlaghőmérséklet [°C]
Budapest, 2001–2010
Klímamodell, 10 km



Felhasználás

- **Objektív** döntéshozatal: számszerű **hatásvizsgálatok**
- Kiindulás: **modelleredmények + bizonytalanságok**
- Példa: városklimatológiai hatásvizsgálatok dinamikus leskálázással (egyben témánk a VKSZ12-ben)
 - Tesztkísérletek Budapestre 1 km-es felbontással
 - Újabb modellszimuláció, megfelelő frekvenciájú és felbontású input-adatigény, nagyfelbontású outputok
 - Az eredmények felhasználhatók a hosszútávú tervezésben

**Nyári átlaghőmérséklet [°C]
Budapest, 2001–2010**
Leskálázás, 1 km



Kitekintés

- További modellkísérletek:
 - Finomabb, 10 km-es felbontáson
 - Nagyobb tartományon
 - Új kibocsátási forgatókönyvekkel
 - Dinamikai leskálázás a felszín leírására
 - A jelenlegi számítási kapacitásunk már nem elegendő az összes tervezett kísérlet elkészítésére
 - Kapacitás-bővítés szükséges – jó lehetőség a pályázati konstrukció
- némelyik kísérlet megvalósítása 3 évbe is telhet

Köszönöm szépen
a figyelmet!

Email:
szepszo.g@met.hu

<http://www.met.hu/RCM>

