NEROGRAV

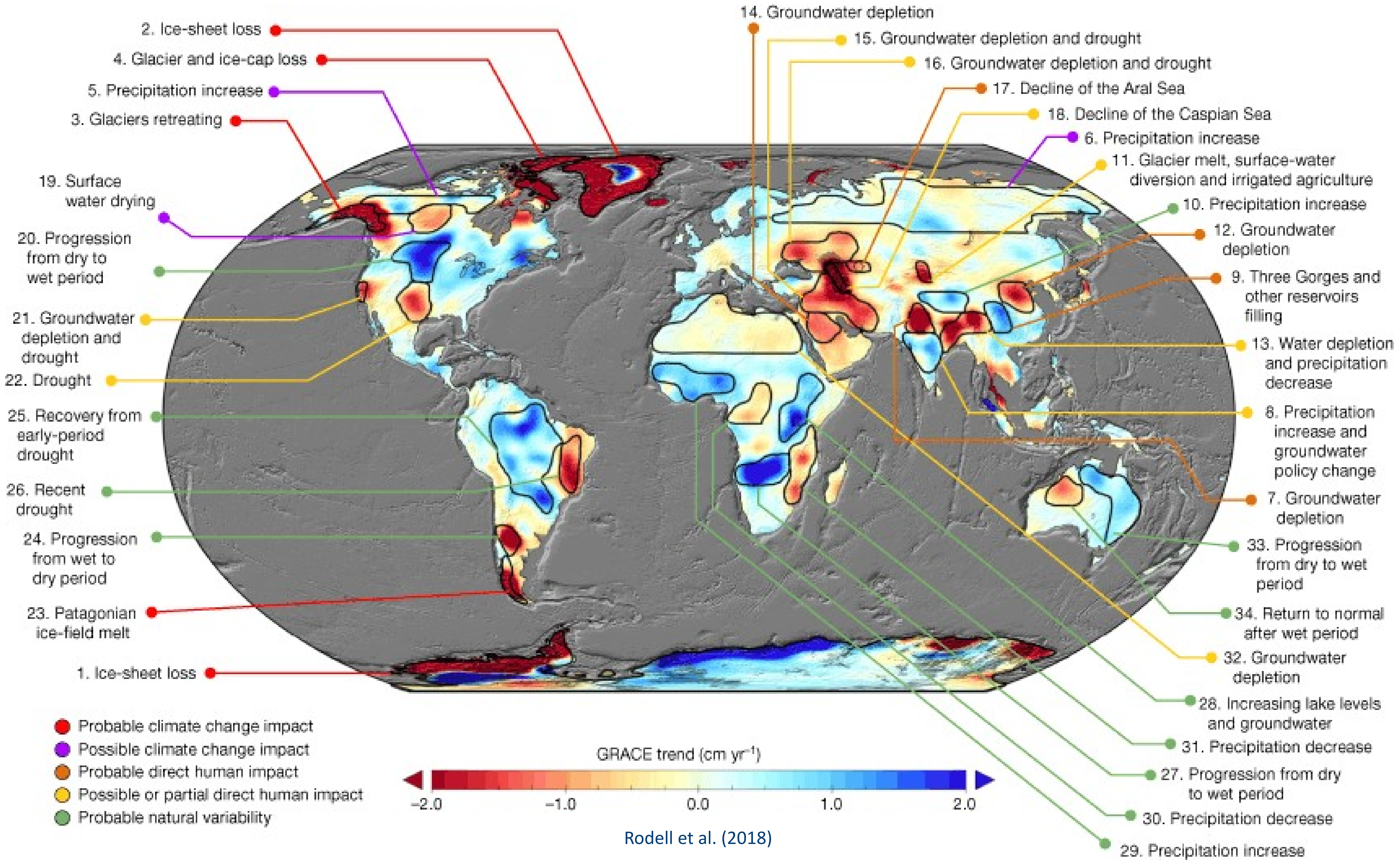
New Refined Observations of Climate Change from Spaceborne Gravity Missions

International Spring School
Neustadt an der Weinstraße, Germany, March 10-14, 2025

Lecture
GRACE/GRACE-FO Data for Model Assimilation and Service Applications

Anne Springer (University of Bonn), Yorck Ewerdwalbesloh

TU berlin · Freie Universität Berlin · Technische Universität München · GFZ Helmholtz Centre for Geosciences · Universität Bonn · hcu HafenCity Universität Hamburg



Rodell et al. (2018)

GRACE/-FO observations for operational applications at local scales?

short-term extreme event prediction

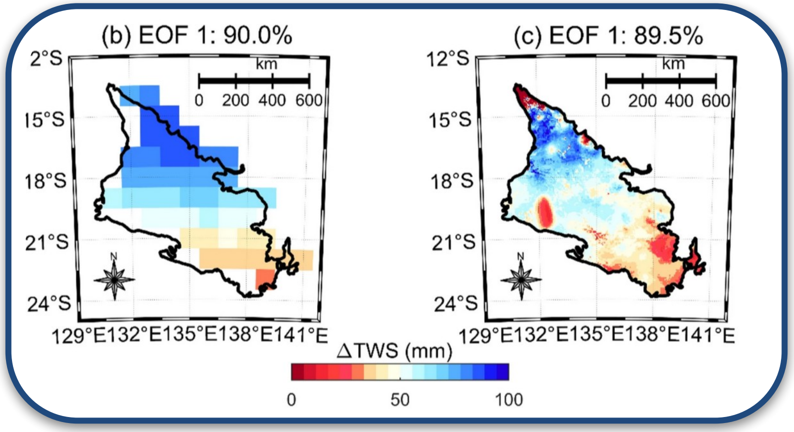
irrigation planning

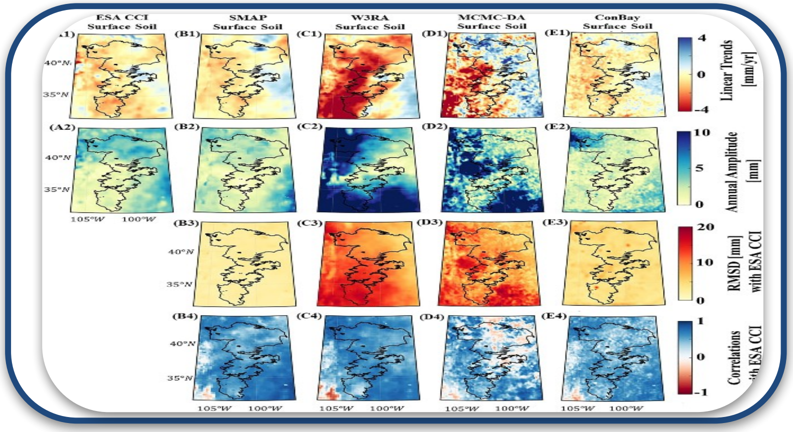
monitoring of local groundwater depletion

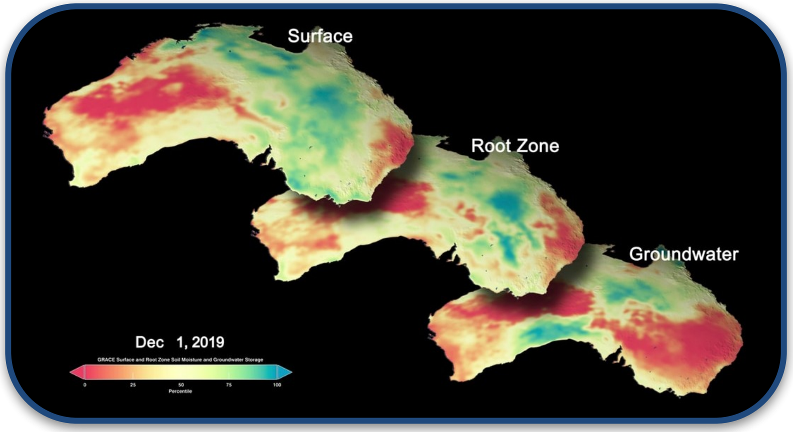
landslide risk assessment

limited spatial resolution
limited temporal resolution and latency
cannot separate surface and subsurface water storage changes

How can we increase the spatial and temporal resolution?







statistical downscaling

- Deep learning methods
- Random forest
- Support vector machines
- Bayesian methods
- Regression
- …

➔ **training** GRACE TWSA towards independent high-resolution data

merging approaches

- Bayesian merging
- Optimal interpolation
- Wavelet based merging
- Deep learning fusion
- …

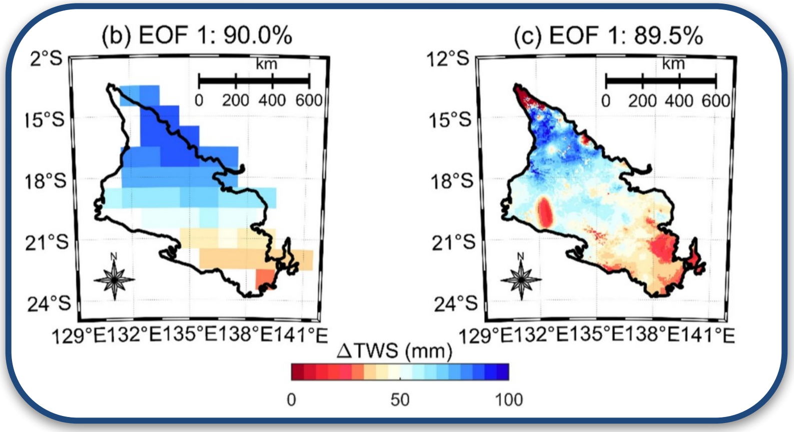
➔ **merging** GRACE TWSA with independent high-resolution data

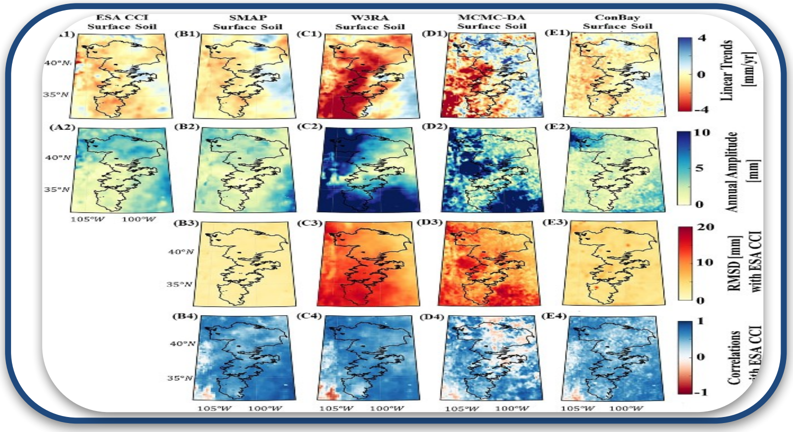
data assimilation

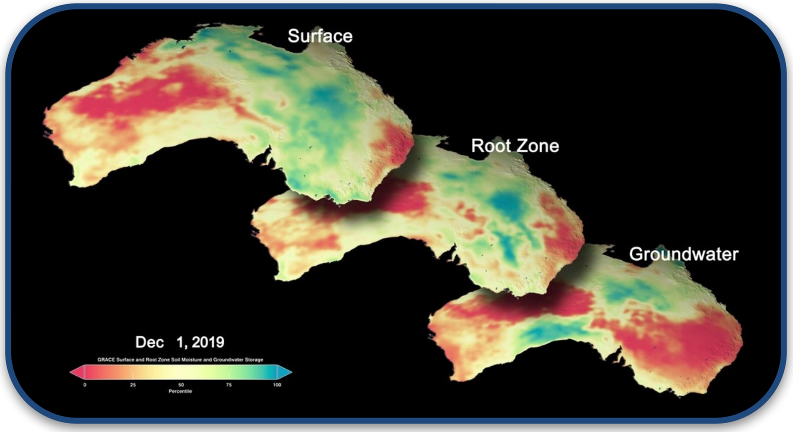
- Kalman Filter-Based methods
- Particle Filter methods
- Variational data assimilation
- …

➔ **integrating** GRACE TWSA into hydrology or land surface models

How can we increase the spatial and temporal resolution?







statistical downscaling

- ✔ provides high-resolution TWSA
- ✔ computational efficient
- ✔ independent of physical models
- ✗ no individual water storages
- ✗ no increase of temporal resolution

merging approaches

- ✔ combines strength of mult. data sets
- ✔ improves spat. and temp. res.
- ✔ provides different water storages
- ✗ inconsistencies between data sets
- ✗ depends on data availability

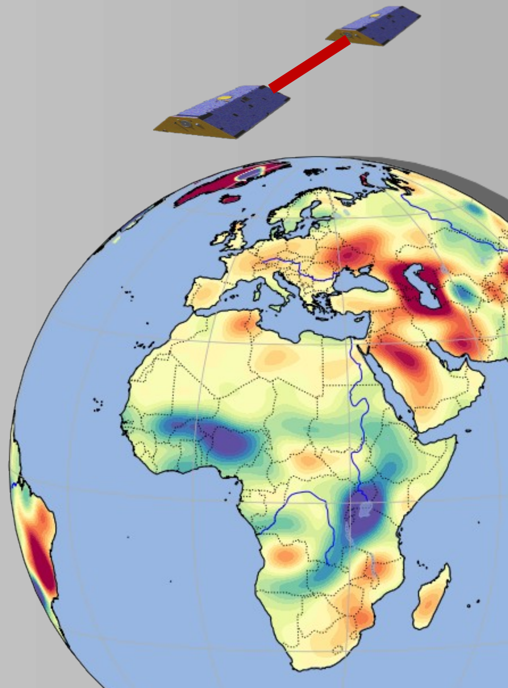
data assimilation

- ✔ physically consistent
- ✔ improves spat. and temp. res.
- ✔ explicitly accounts for uncertainty
- ✔ improves model prediction
- ✔ improves individual water storages and fluxes
- ✗ complex implementation
- ✗ computationally expensive

➔ **validation remains a challenge for all three approaches!**

Today we will learn …

- What do we need to set up a GRACE/-FO data assimilation framework?
- Which are the specific challenges of GRACE/-FO data assimilation frameworks?
- Which applications exist for GRACE/-FO data assimilation frameworks?



Assimilation of GRACE Terrestrial Water Storage Data into a Land Surface Model: Results for the Mississippi River Basin

Benjamin F. Zaitchik

Assimilation of terrestrial water storage from GRACE in a snow-dominated basin

B. A. Forman,[1,2,3] R. H. Reichle,[1] and M. Rodell[4]

Assimilation of gridded terrestrial water storage observations from GRACE into a land surface model

Manuela Girotto[1,2], Gabriëlle J. M. De Lannoy[1,2], Rolf H. Reichle[1], and Matthew Rodell[1]

Improving estimates of water resources in a semi-arid region by assimilating GRACE data into the PCR-GLOBWB hydrological model

Natthachet Tangdamrongsub[1,2], Susan C. Steele-Dunne[3], Brian C. Gunter[1,4], Pavel G. Ditmar[1], Edwin H. Sutanudjaja[5], Yu Sun[1], Ting Xia[6], and Zhongjing Wang[6,7]

Improving drought simulations within the Murray-Darling Basin by combined calibration/assimilation of GRACE data into the WaterGAP Global Hydrology Model

M. Schumacher[a,b,c,*], E. Forootan[d], A.I.J.M. van Dijk[c], H. Müller Schmied[e,f], R.S. Crosbie[g], J. Kusche[b], P. Döll[e,f]

Determining water storage depletion within Iran by assimilating GRACE data into the W3RA hydrological model

M. Khaki[a,*], E. Forootan[b], M. Kuhn[a], J. Awange[a], A.I.J.M. van Dijk[c], M. Schumacher[d], M.A. Sharifi[e,f]

GRACE Improves Seasonal Groundwater Forecast Initialization over the United States

Augusto Getirana,[a,b] Matthew Rodell,[a] Sujay Kumar,[a] Hiroko Kato Beaudoing,[a,b] Kristi Arsenault,[a,c] Benjamin Zaitchik,[d] Himanshu Save,[e] and Srinivas Bettadpur[e]

Estimating Terrestrial Snow Mass via Multi-Sensor Assimilation of Synthetic AMSR-E Brightness Temperature Spectral Differences and Synthetic GRACE Terrestrial Water Storage Retrievals

Jing Wang[1], Barton A. Forman[1], Manuela Girotto[2], and Rolf H. Reichle[3]

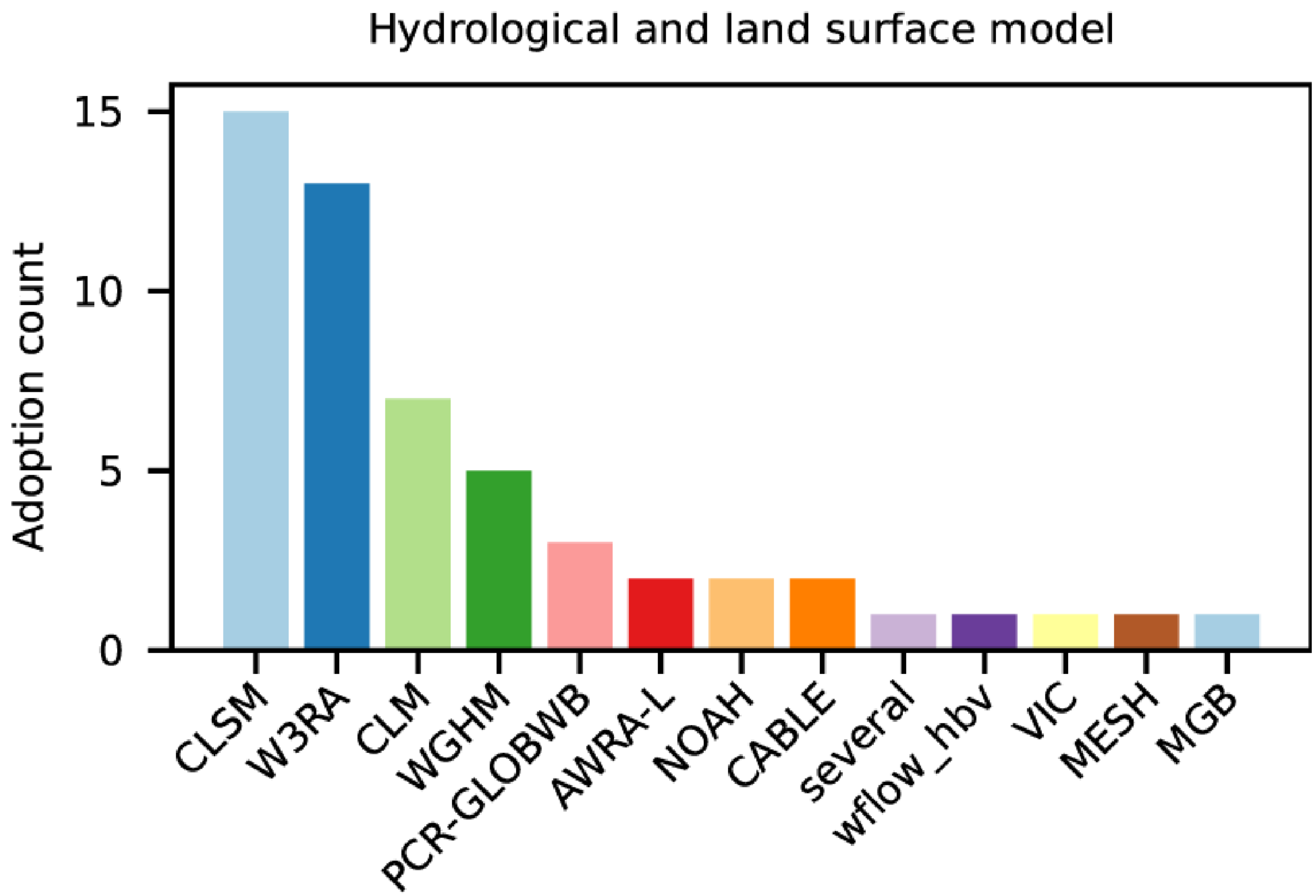
Revising precipitation – water storages – vegetation signatures with GRACE-based data assimilation

Helena Gerdener[a,*], Jürgen Kusche[a], Kerstin Schulze[a], Gohar Ghazaryan[b,c,d], Olena Dubovyk[b,c,e]

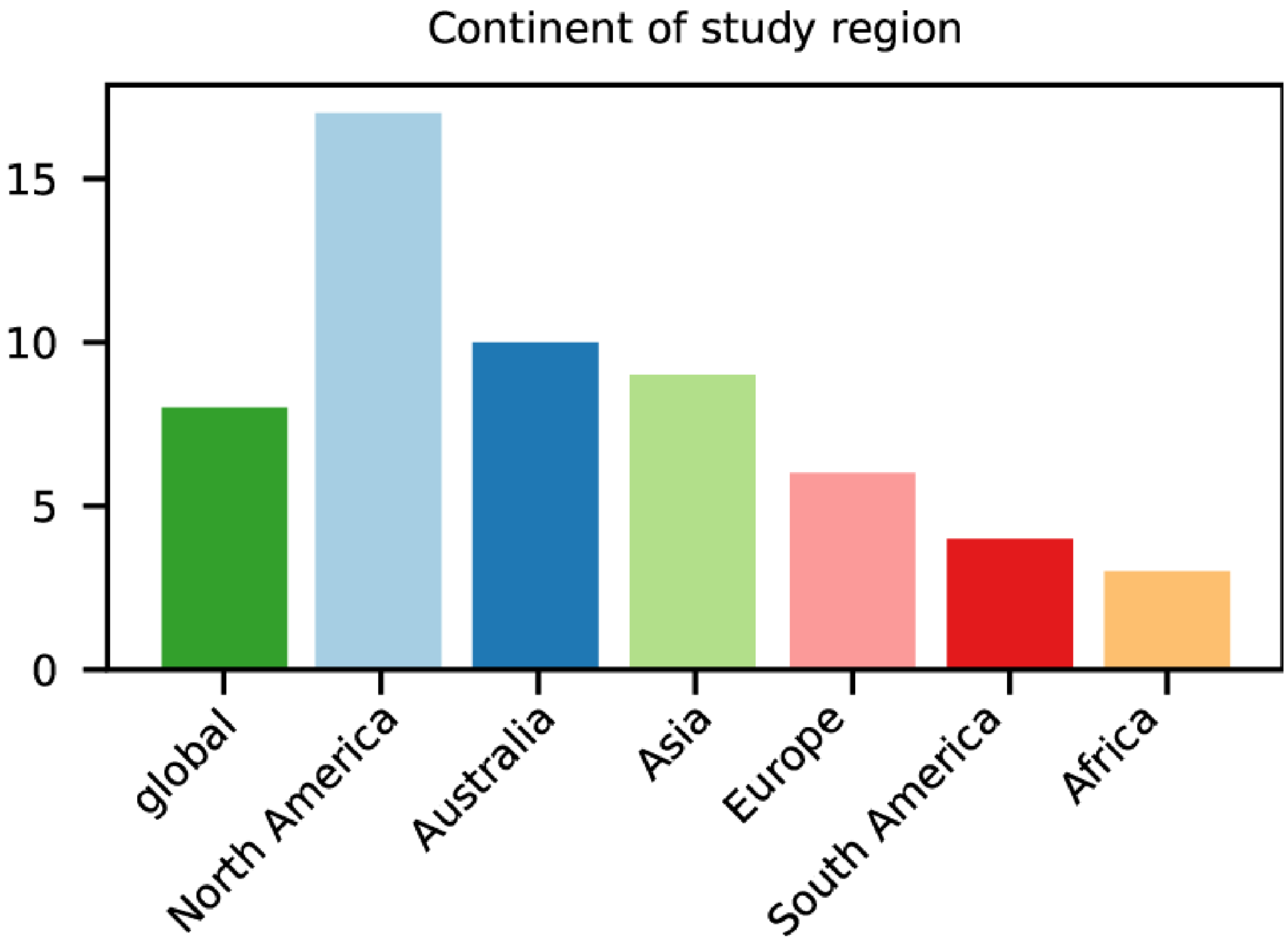
Assimilating GRACE Into a Land Surface Model in the Presence of an Irrigation-Induced Groundwater Trend

Wanshu Nie[1], Benjamin F. Zaitchik[1], Matthew Rodell[2], Sujay V. Kumar[2], Kristi R. Arsenault[2,3], Bailing Li[2,4], and Augusto Getirana[2,4]

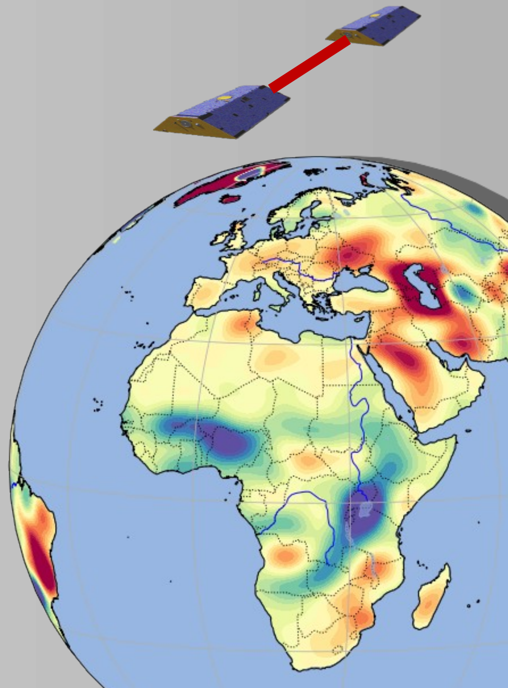
Which model is used?



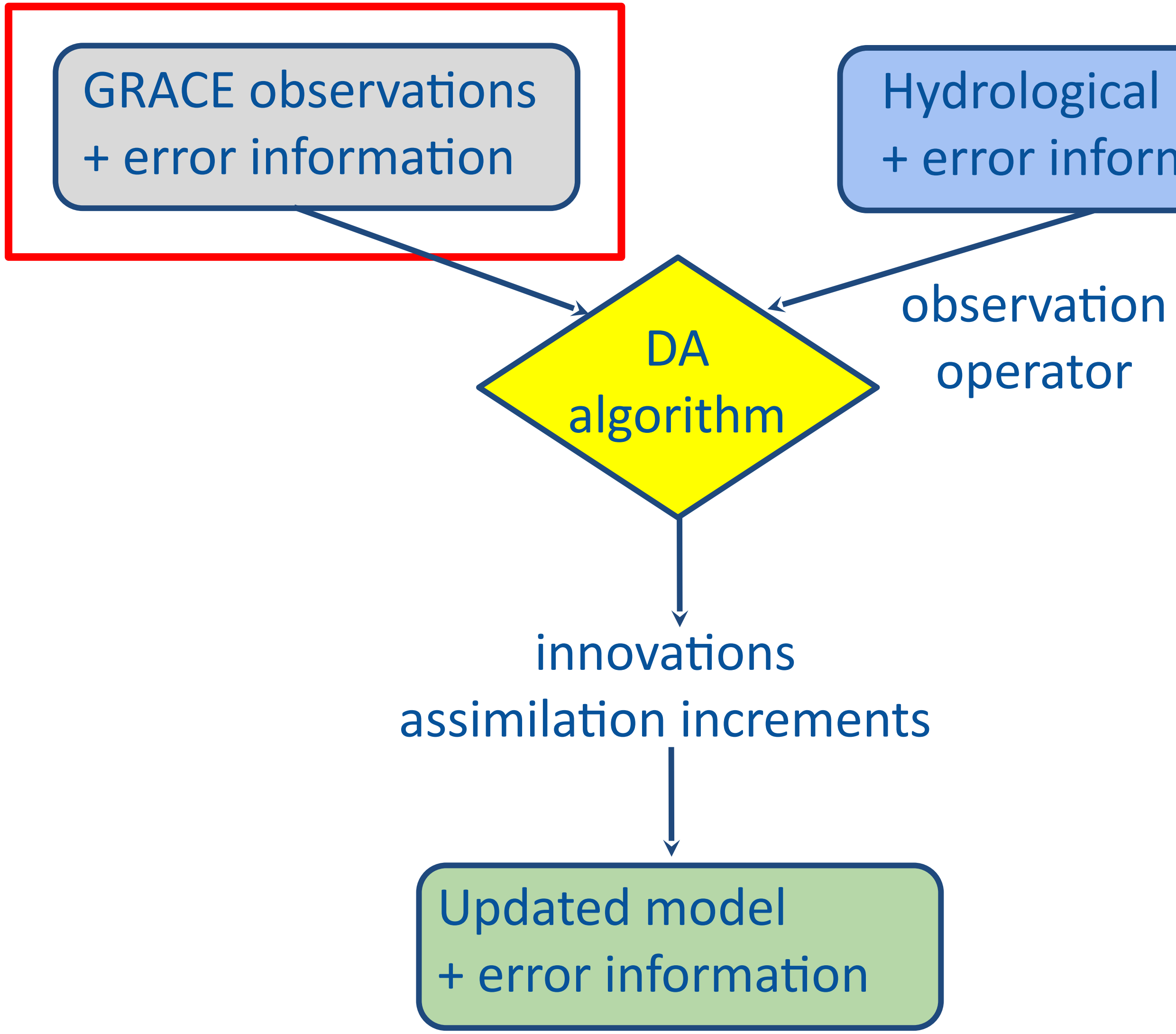
Which study region is selected?



Let us start with a recipe for GRACE/-FO data assimilation!



Set up of a data assimilation framework



Terminology

- DA … data assimilation
- TWSA … terrestrial water storage anomalies
- GRACE … GRACE + GRACE-FO
- KF … Kalman Filter
- EnKF … Ensemble Kalman Filter

Recap GRACE/-FO observations

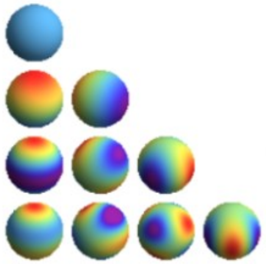
http://icgem.gfz-potsdam.de/series

Spherical harmonics

Level 2

- parameters in spectral domain
- unconstrained solution
- anisotropic error

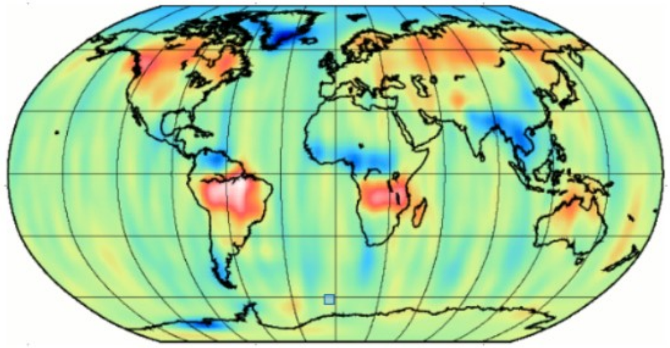
$$
\begin{array}{ccccccc}
 & & & \overline{C}_{00} & & & \\
 & & \overline{S}_{11} & \overline{C}_{10} & \overline{C}_{11} & & \\
 & \overline{S}_{22} & \overline{S}_{21} & \overline{C}_{20} & \overline{C}_{21} & \overline{C}_{22} & \\
\overline{S}_{33} & \overline{S}_{32} & \overline{S}_{31} & \overline{C}_{30} & \overline{C}_{31} & \overline{C}_{32} & \overline{C}_{33}
\end{array}
$$



http://thegraceplotter.com/

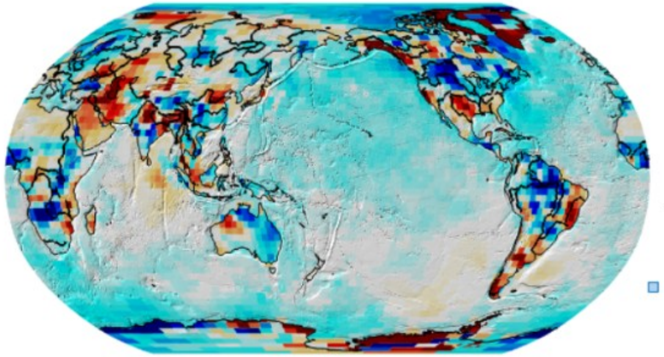
Level 3

- regular grid
- filtering necessary
- Leakage effect



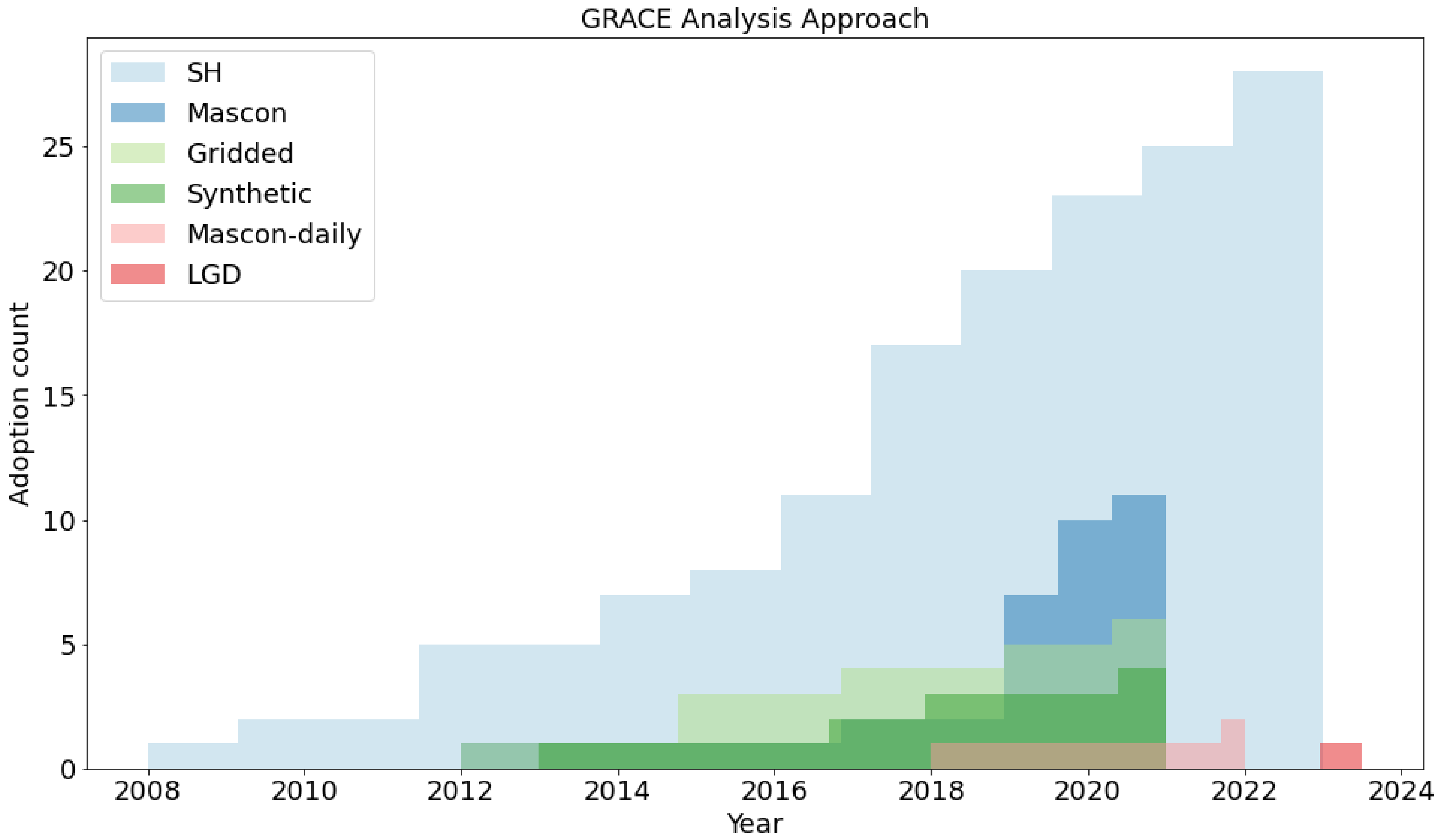
Mascons

http://www2.csr.utexas.edu/grace/RL06_mascons.html



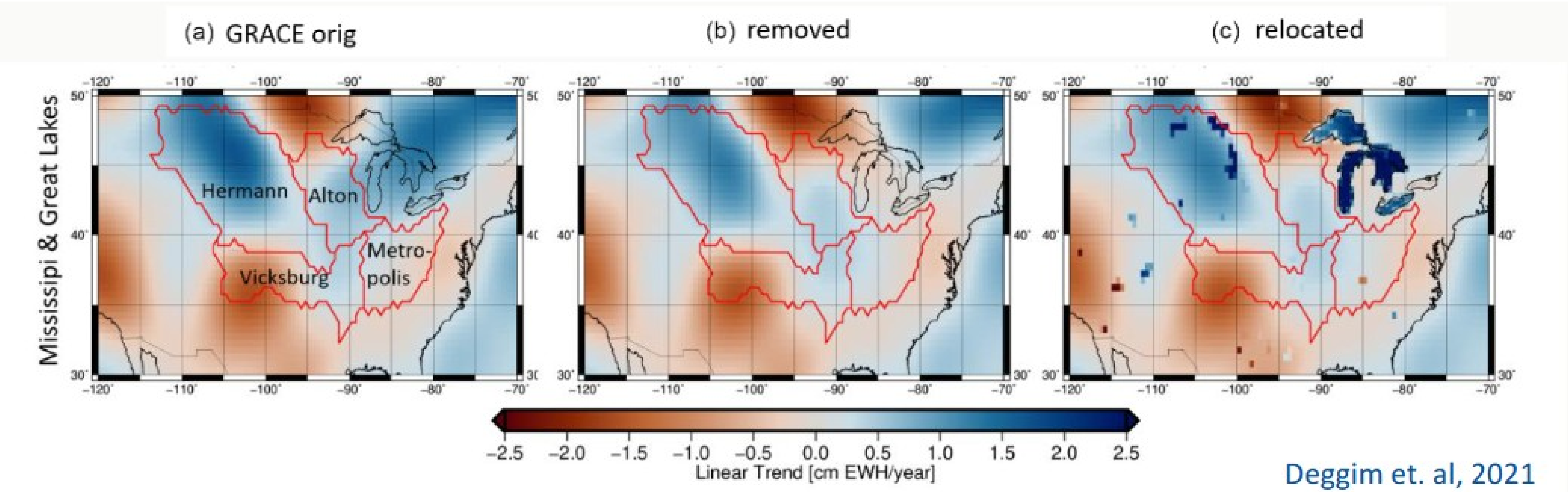
- parameters in space domain
- spatial constraints applied
- leakage effects avoided

Which GRACE data set do we use for DA?



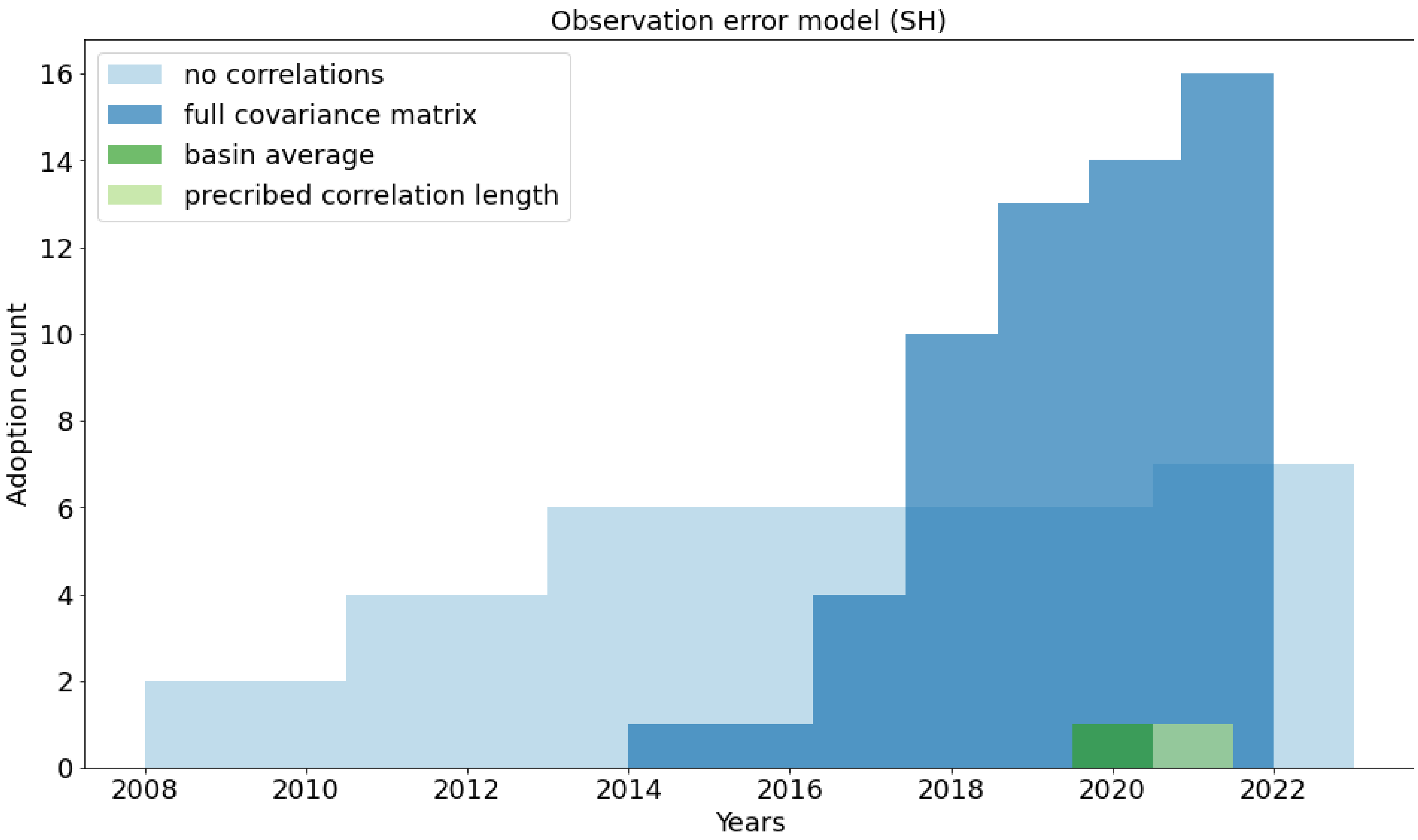
Geophysical corrections

- Glacial isostatic adjustment (GIA)
- Lakes / reservoirs
- Earthquakes
- Glaciers

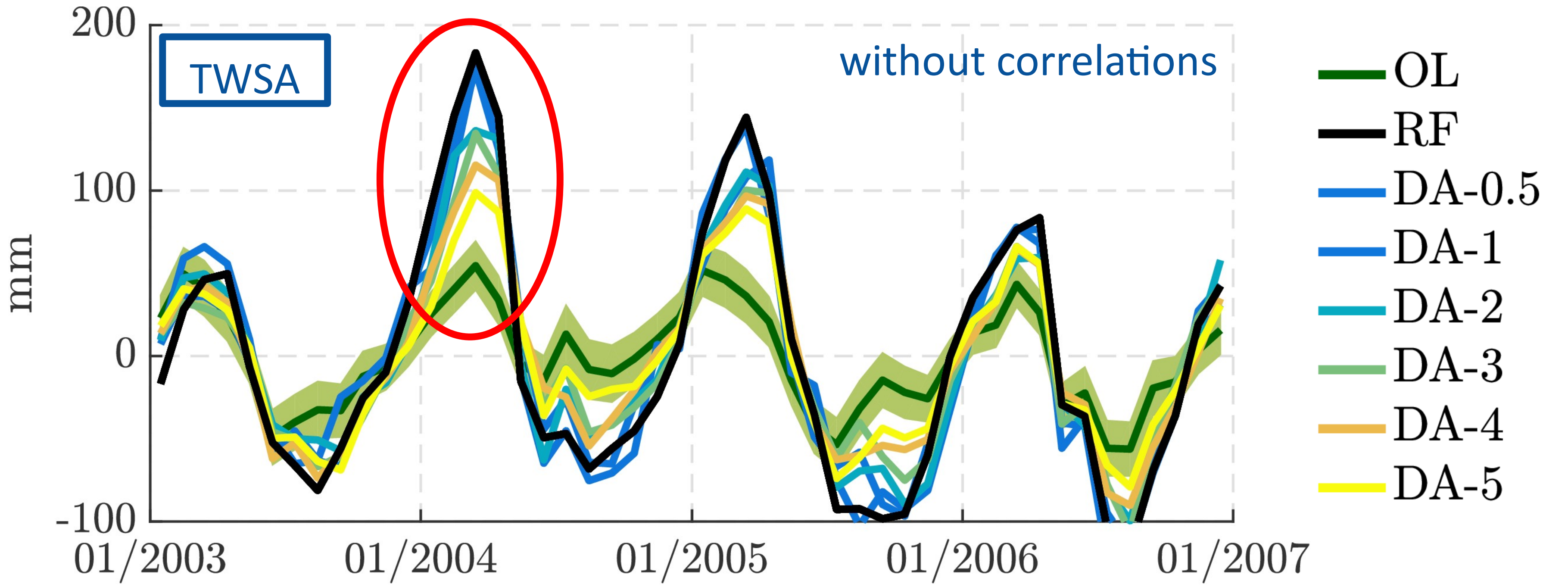


Deggim et. al, 2021

Which kind of observation errors are considered?

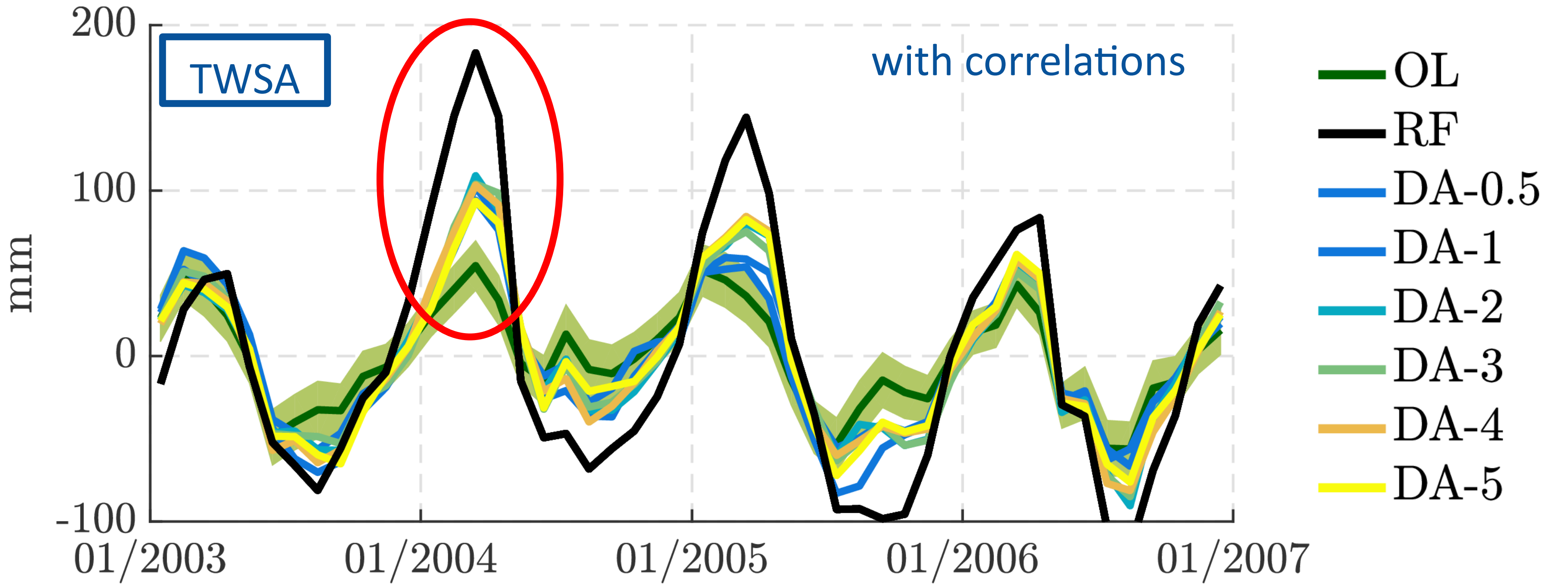


Observation grid and observation error model



OSSE experiment with CLM3.5 over Europe (here: Daugava, Narva, Neva)

Observation grid and observation error model

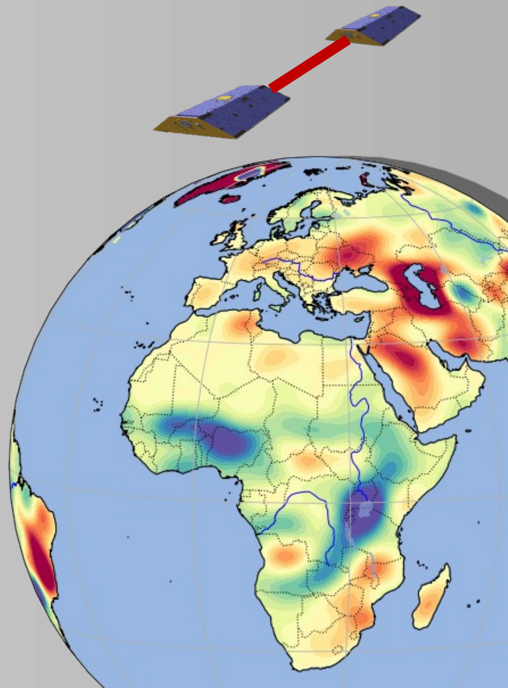


OSSE experiment with CLM3.5 over Europe (here: Daugava, Narva, Neva)

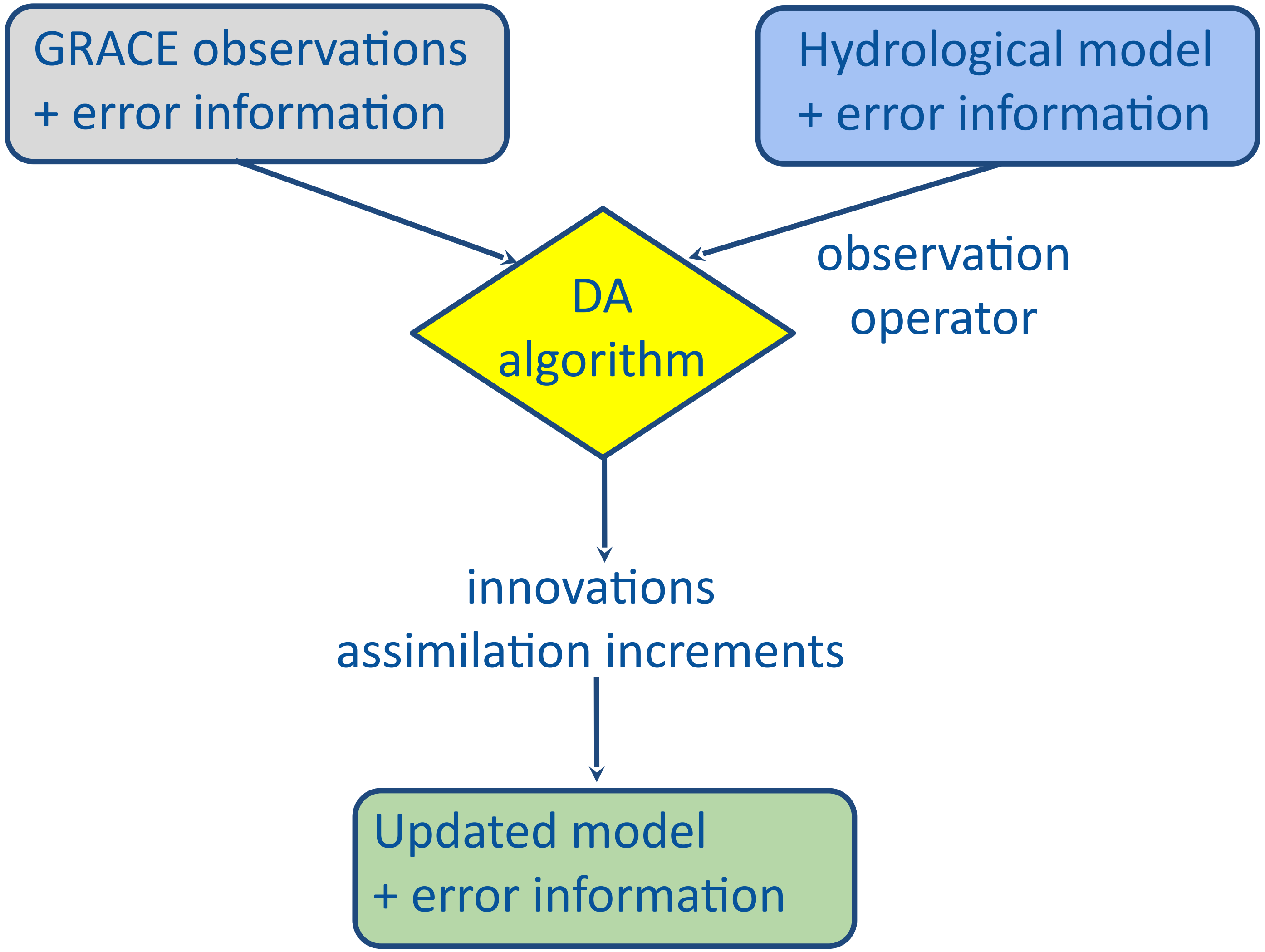
➢ **Spatial correlations need to be taken into account!**

Recap GRACE/-FO observations

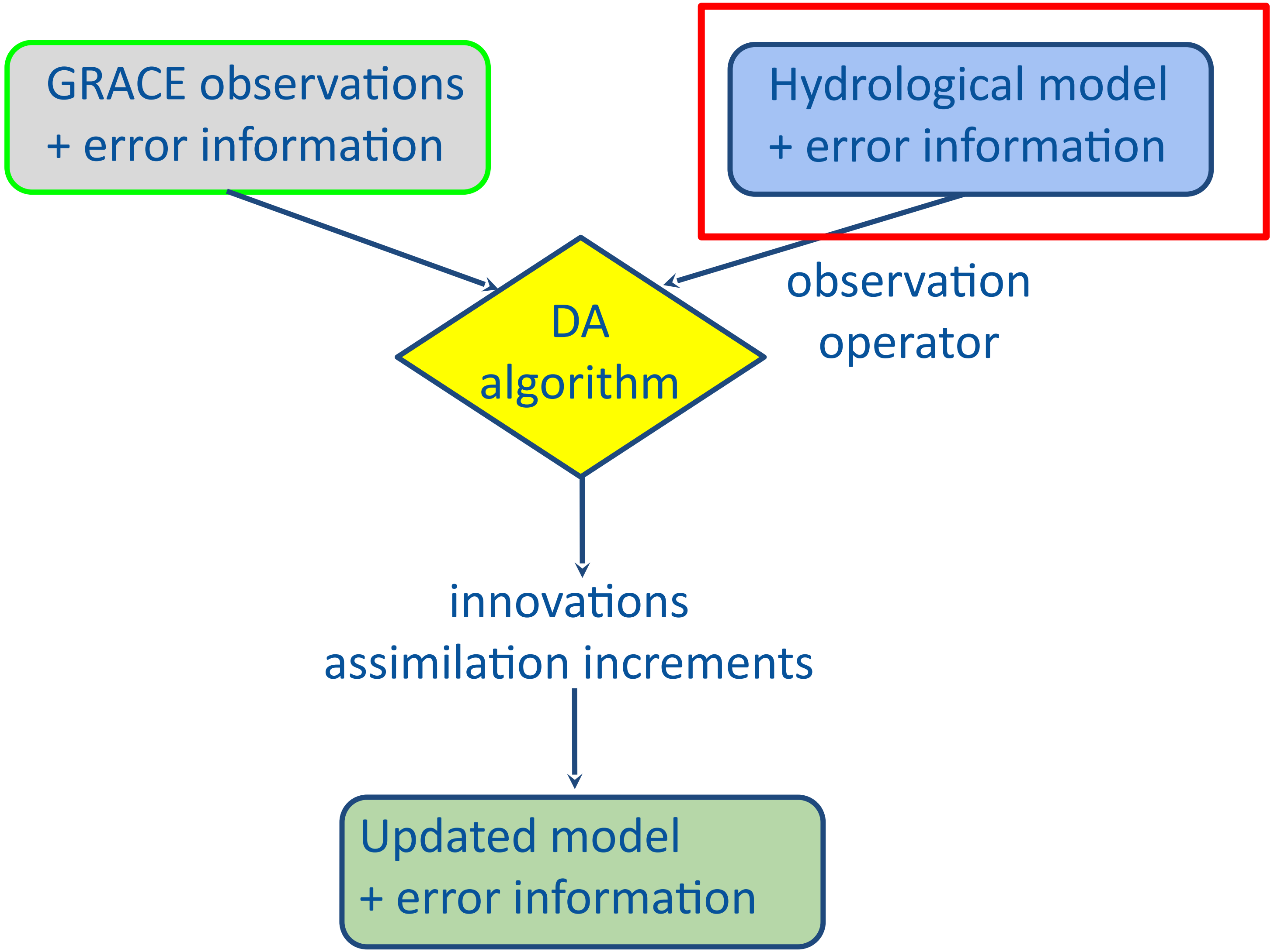
- Long-term mean has to be removed consistently with the hydrological model used for DA
- Geophysical corrections need to be applied carefully
- Realistic error estimates including correlations are important

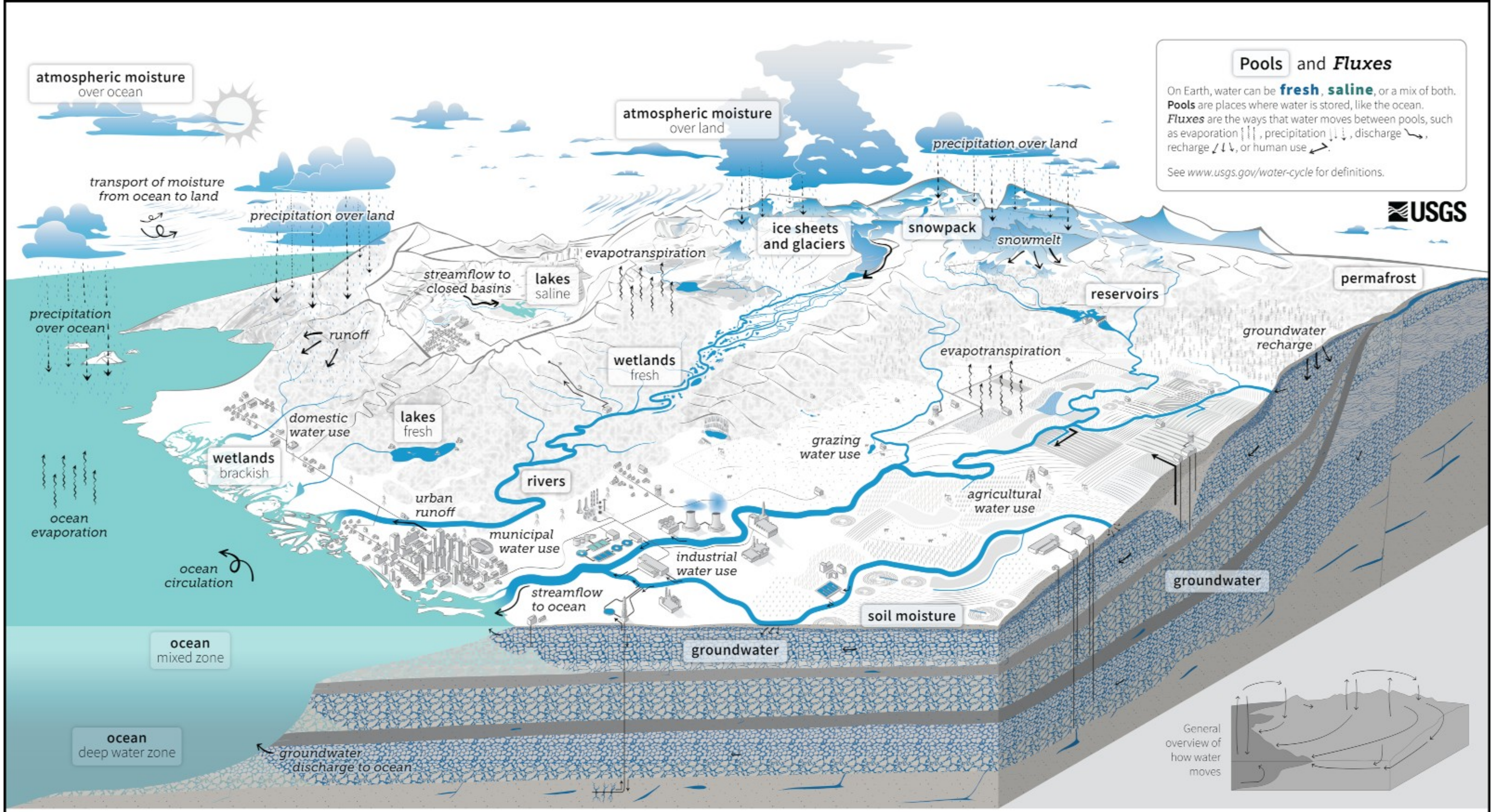


Set up of a data assimilation framework



Set up of a data assimilation framework





The Water Cycle

The water cycle describes where water is on Earth and how it moves. Water is stored in the atmosphere, on the land surface, and below the ground. It can be a liquid, a solid, or a gas. Liquid water can be fresh, saline (salty), or a mix (brackish). Water moves between the places it is stored. Water moves at large scales and at very small scales. Water moves naturally and because of human actions. Human water use affects where water is stored, how it moves, and how clean it is.

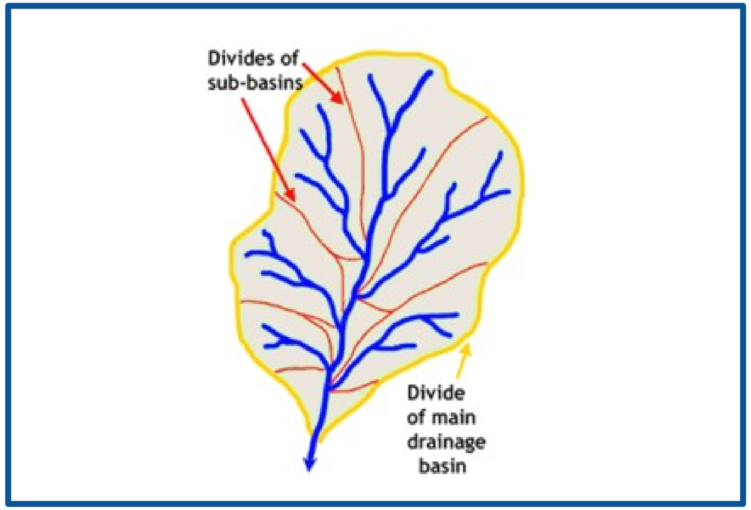
Pools store water. 96% of all water is stored in **oceans** and is saline. On land, saline water is stored in **saline lakes**. Fresh water is stored in liquid form in **freshwater lakes**, artificial **reservoirs**, **rivers**, and **wetlands**. Water is stored in solid, frozen form in **ice sheets** and **glaciers**, and in **snowpack** at high elevations or near the Earth's poles. Water vapor is a gas and is stored as **atmospheric moisture** over the ocean and land. In the soil, frozen water is stored as **permafrost** and liquid water is stored as **soil moisture.** Deeper below ground, liquid water is stored as **groundwater** in aquifers, within cracks and pores in the rock.

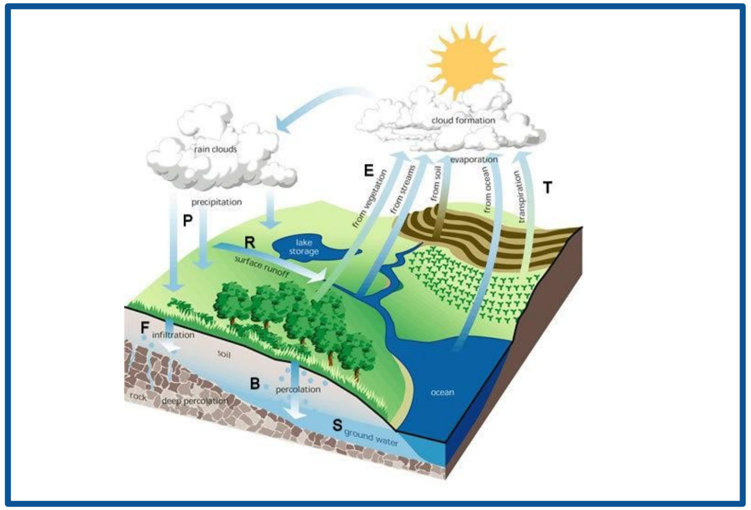
Fluxes move water between pools. As it moves, water can change form between liquid, solid, and gas. **Circulation** mixes water in the oceans and transports water vapor in the atmosphere. Water moves between the atmosphere and the surface through **evaporation**, **evapotranspiration**, and **precipitation**. Water moves across the surface through **snowmelt**, **runoff**, and **streamflow**. Water moves into the ground through infiltration and **groundwater recharge**. Underground, groundwater flows within aquifers. It can return to the surface through natural **groundwater discharge** into rivers, the ocean, and from **springs**.

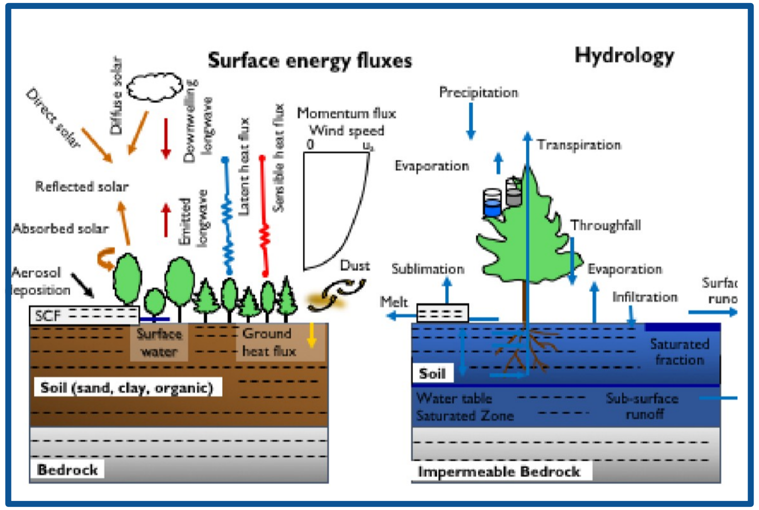
We alter the water cycle. We redirect rivers. We build dams to store water. We drain water from wetlands for development. We use water from rivers, lakes, reservoirs, and groundwater aquifers. We use that water to supply our **homes** and **communities.** We use it for **agricultural** irrigation and **grazing** livestock. We use it in **industrial** activities like thermoelectric power generation, mining, and aquaculture. The amount of water that is available depends on how much water is in each pool (water quantity). It also depends on when and how fast water moves (water timing), how much water we use (water use), and how clean the water is (water quality).

We affect **water quality**. In agricultural and urban areas, irrigation and precipitation wash fertilizers and pesticides into rivers and groundwater. Power plants and factories return heated and contaminated water to rivers. Runoff carries chemicals, sediment, and sewage into rivers and lakes. Downstream from these sources, contaminated water can cause harmful algal blooms, spread diseases, and harm habitats. **Climate change** is affecting the water cycle. It is affecting water quality, quantity, timing, and use. It is causing ocean acidification, sea level rise, and more extreme weather. By understanding these impacts, we can work toward using water sustainably.

Hydrological models and land surface models





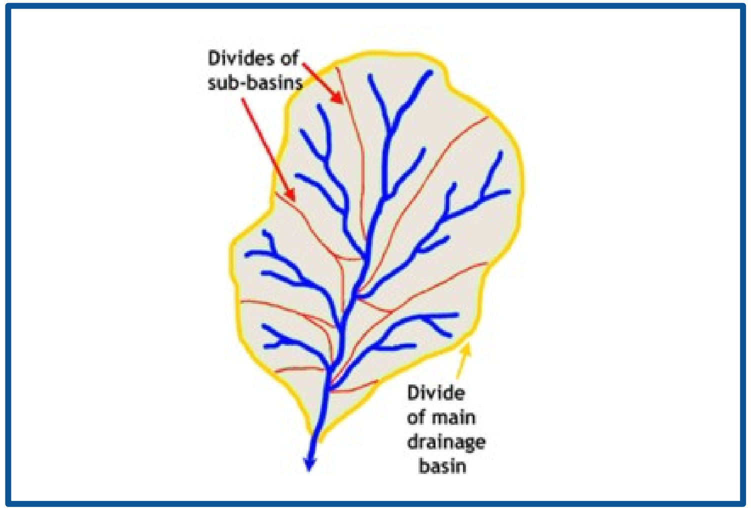


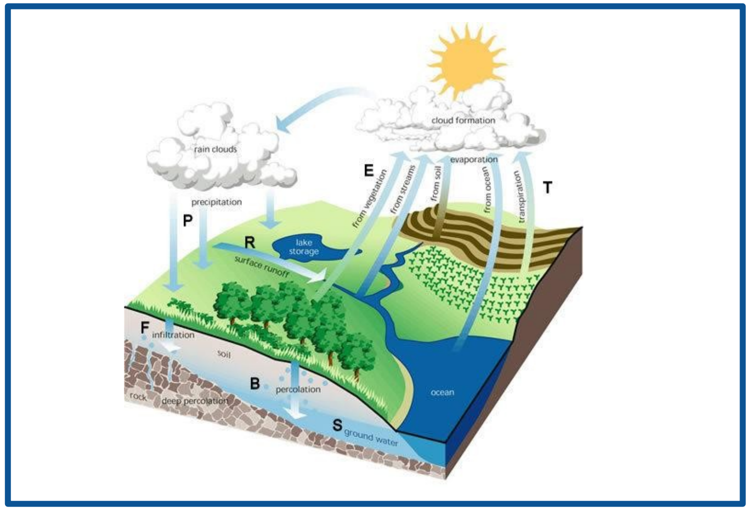
Lumped rainfall-runoff models:
establish a relationship between rainfall and runoff for predicting floods in particular in limited-gauged catchments.

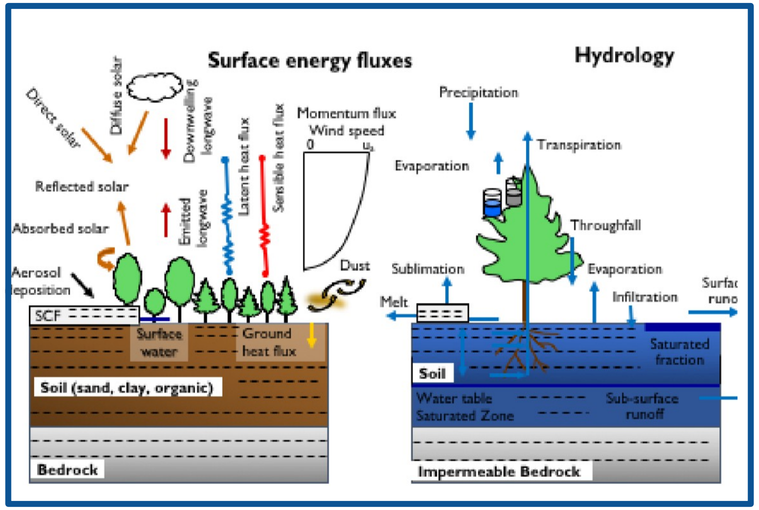
Conceptual models:
developed for hydrological applications (simulation of discharge, water balances in river catchments, water management).

Land surface models:
developed with the scope of simulating exchange processes between land surface and atmosphere as represented by atmospheric circulation models.

Hydrological models and land surface models







- entire river (sub-) basin seen as one unit
- impose many assumptions
- empirical equations describe the physics

- surface water fluxes and lateral fluxes
- transport in the river network
- groundwater
- human water abstraction

- Complete (physical) energy and water balances at the land surface
- vertical water fluxes
- no explicite river routing

Hydrological models and land surface models

Model	Type	Components	Anthropogenic	Reference
AWRA-L	LSM	SW, SM, GW	Partial	Viney et al. (2014)
CABLE	LSM	SW, SM	No	Kowalczyk et al. (2006)
CLM3.5	LSM	SW, SM	No	Oleson et al. (2007)
CLM4	LSM	SW, SM	No	Lawrence et al. (2011)
CLM5-CRUNCEP	LSM	SW, SM	No	Lawrence et al. (2019)
CLM5-GSWP3	LSM	SW, SM	No	Lawrence et al. (2019)
CLSM	LSM	SW, SM	No	Koster et al. (2000a)
HBV-SIMREG	GHM	SW, SM	No	Lindström et al. (1997)
HTESSEL	LSM	SW, SM	No	Balsamo et al. (2015)
LISFLOOD	GHM	SW, SM, GW	Yes	Van Der Knijff et al. (2010)
MESH	GHM/LSM Hybrid	SW, SM, GW	Yes	Pietroniro et al. (2007)
MGB	GHM	SW, SM, GW	Yes	Collischonn et al. (2007)
Noah	LSM	SW, SM	No	Ek et al. (2003)
Noah-MP	LSM	SW, SM	No	Niu et al. (2011)
ORCHIDEE	LSM	SW, SM, GW	Partial	Polcher et al. (2011)
ParFlow-CLM	LSM+G	SW, SM, GW	No	Maxwell et al. (2015)
PCR-GLOBWB	GHM	SW, SM, GW	Yes	Sutanudjaja et al. (2018)
SURFEX-TRIP	LSM	SW, SM	No	Decharme et al. (2013)
SWBM	GHM	SW, SM	No	Koster and Mahanama (2012)
VIC	LSM	SW, SM	Yes	Liang et al. (1996)
W3RA	GHM	SW, SM	No	Van Dijk (2010)
WGHM	GHM	SW, SM, GW	Yes	Müller Schmied et al. (2021)
WBM	GHM	SW, SM, GW	Yes	Tiaden et al. (1998)

Validation: Hydrological Models

human water consumption

natural vegetation

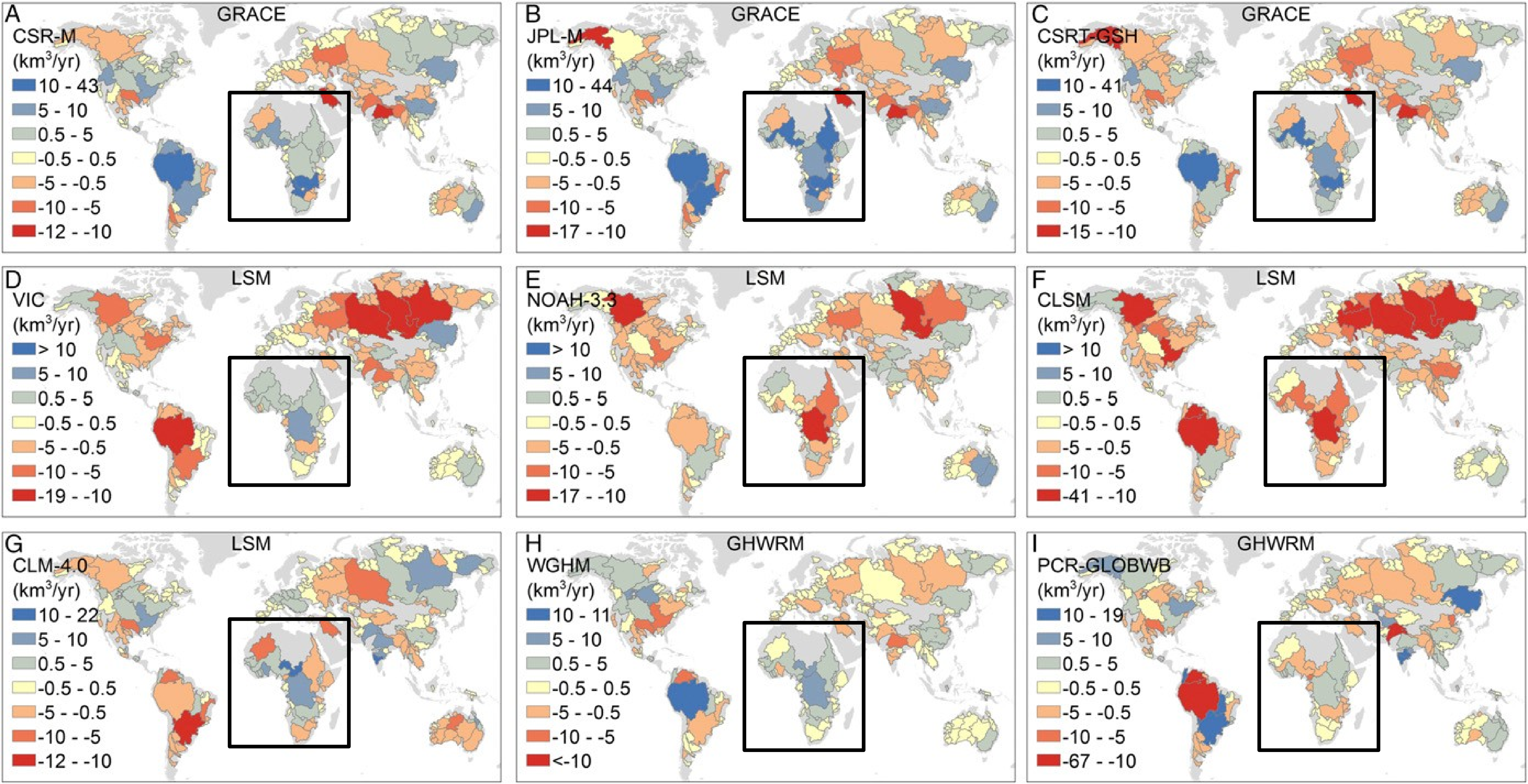
TWS = snow + (canopy) + surface water + soil moisture + groundwater

GRACE

land surface models (LSMs)

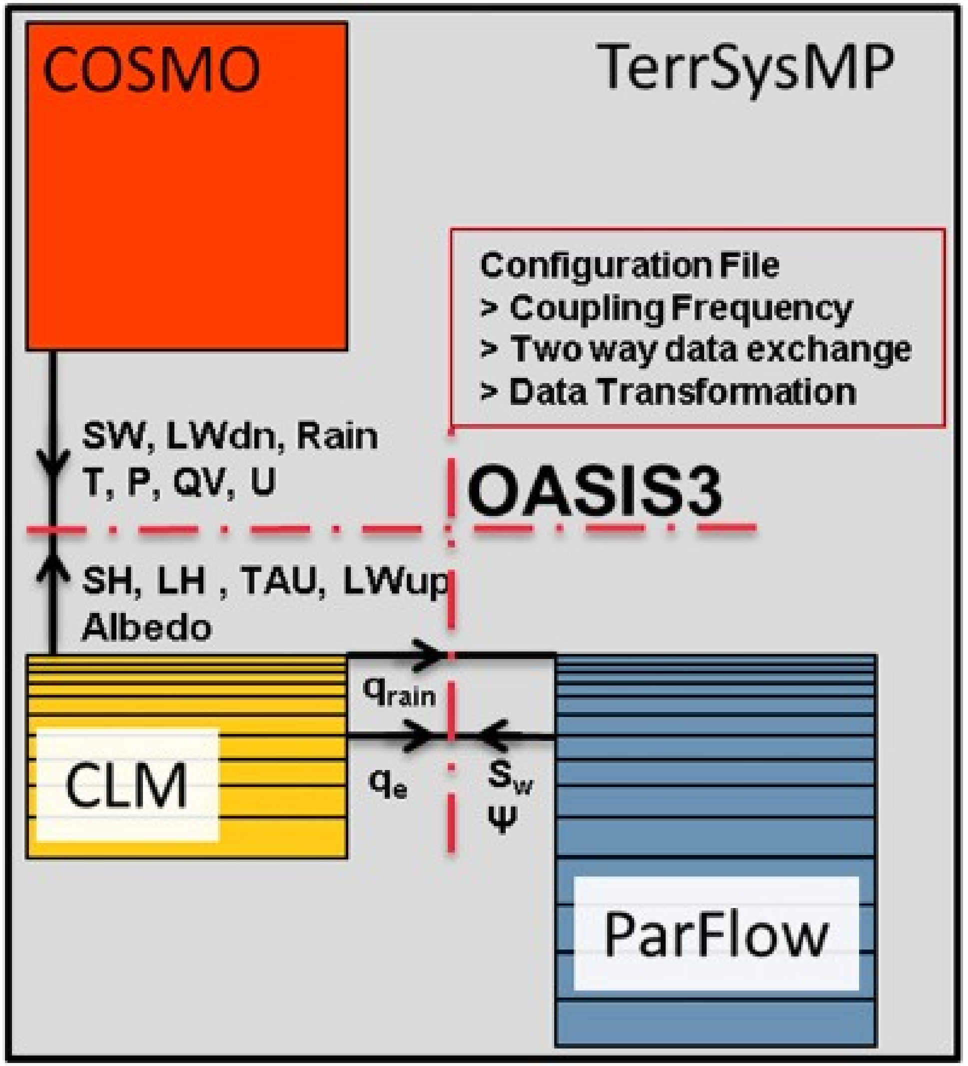
global hydrological and water resource models (GHWRMs)

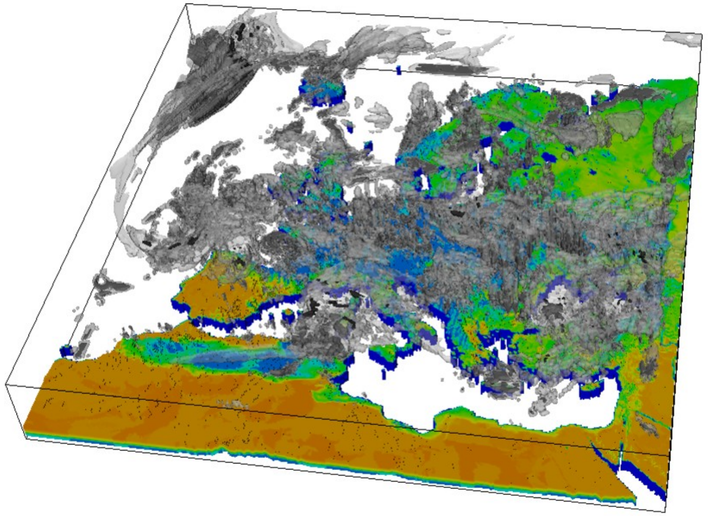
anthropogenic influences

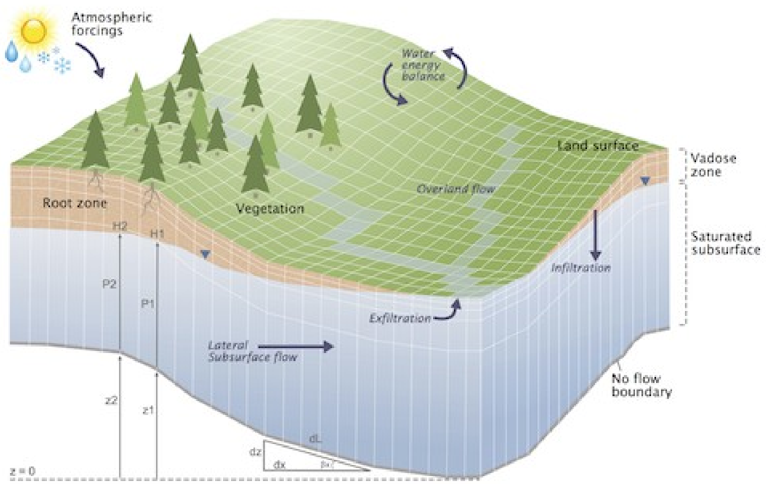


Scanlon et al. (2018) Global models underestimating large decadal declining and rising water storage trends relative to GRACE satellite data, ***Proceedings of the National Academy of Sciences.***

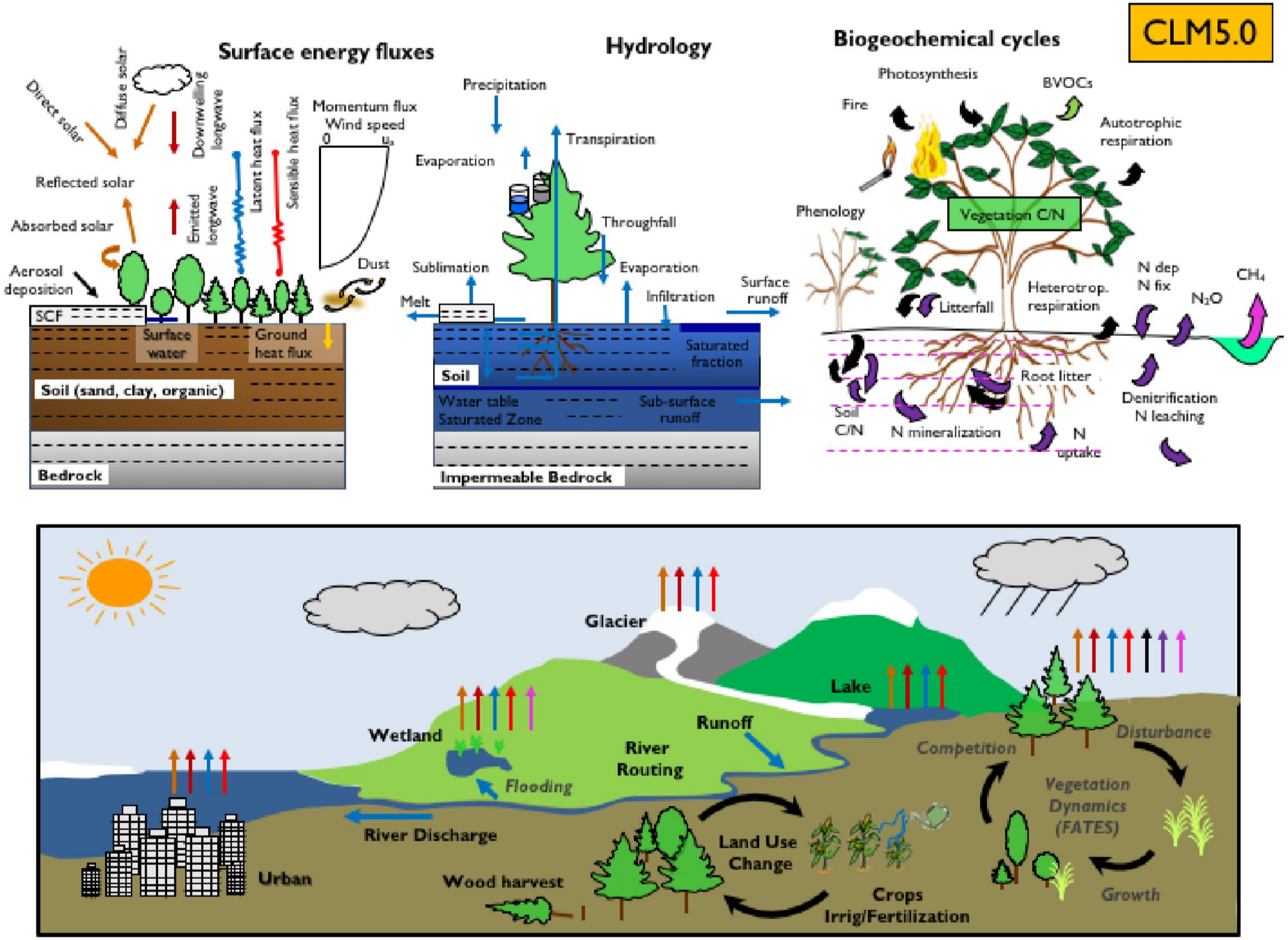
A regional Earth system model



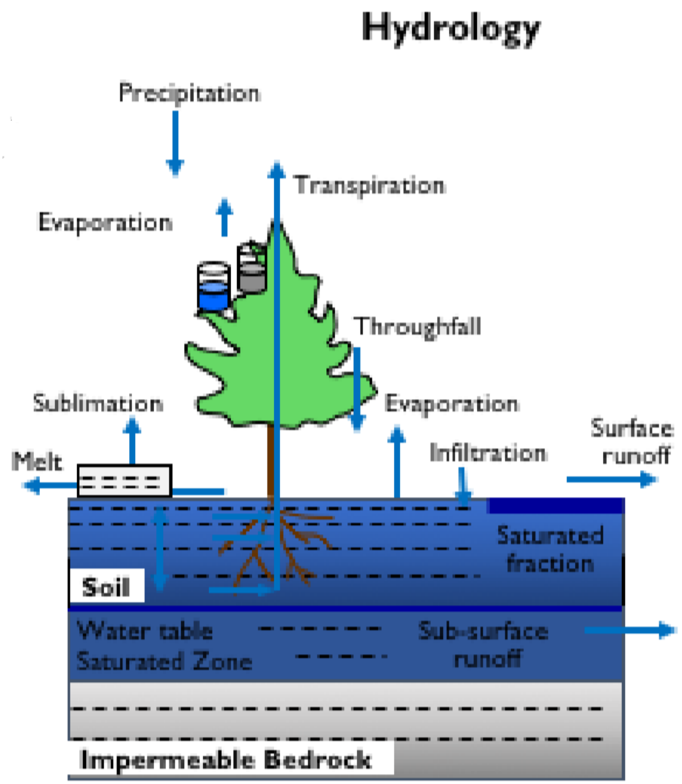




The Community Land Model CLM5



Das Community Land Model CLM5



Many, many equations….

7.3.2 Numerical Solution

With reference to Figure 7.2, the equation for conservation of mass (equation (7.41)) can be integrated over each layer as

$$\int_{-z_{h,\,i}}^{-z_{h,\,i-1}} \frac{\partial \theta}{\partial t}\, dz = -\int_{-z_{h,\,i}}^{-z_{h,\,i-1}} \frac{\partial q}{\partial z}\, dz - \int_{-z_{h,\,i}}^{-z_{h,\,i-1}} e\, dz. \tag{7.64}$$

Note that the integration limits are negative since z is defined as positive upward from the soil surface. This equation can be written as

$$\Delta z_i \frac{\partial \theta_{liq,\,i}}{\partial t} = -q_{i-1} + q_i - e_i \tag{7.65}$$

where q_i is the flux of water across interface $z_{h,\,i}$, q_{i-1} is the flux of water across interface $z_{h,\,i-1}$, and e_i is a layer-averaged soil moisture sink term (ET loss) defined as positive for flow out of the layer (mm s^{-1}). Taking the finite difference with time and evaluating the fluxes implicitly at time $n+1$ yields

$$\frac{\Delta z_i \Delta \theta_{liq,\,i}}{\Delta t} = -q_{i-1}^{n+1} + q_i^{n+1} - e_i \tag{7.66}$$

where $\Delta \theta_{liq,\,i} = \theta_{liq,\,i}^{n+1} - \theta_{liq,\,i}^{n}$ is the change in volumetric soil liquid water of layer i in time $\Delta t$and Δz_i is the thickness of layer i (mm).

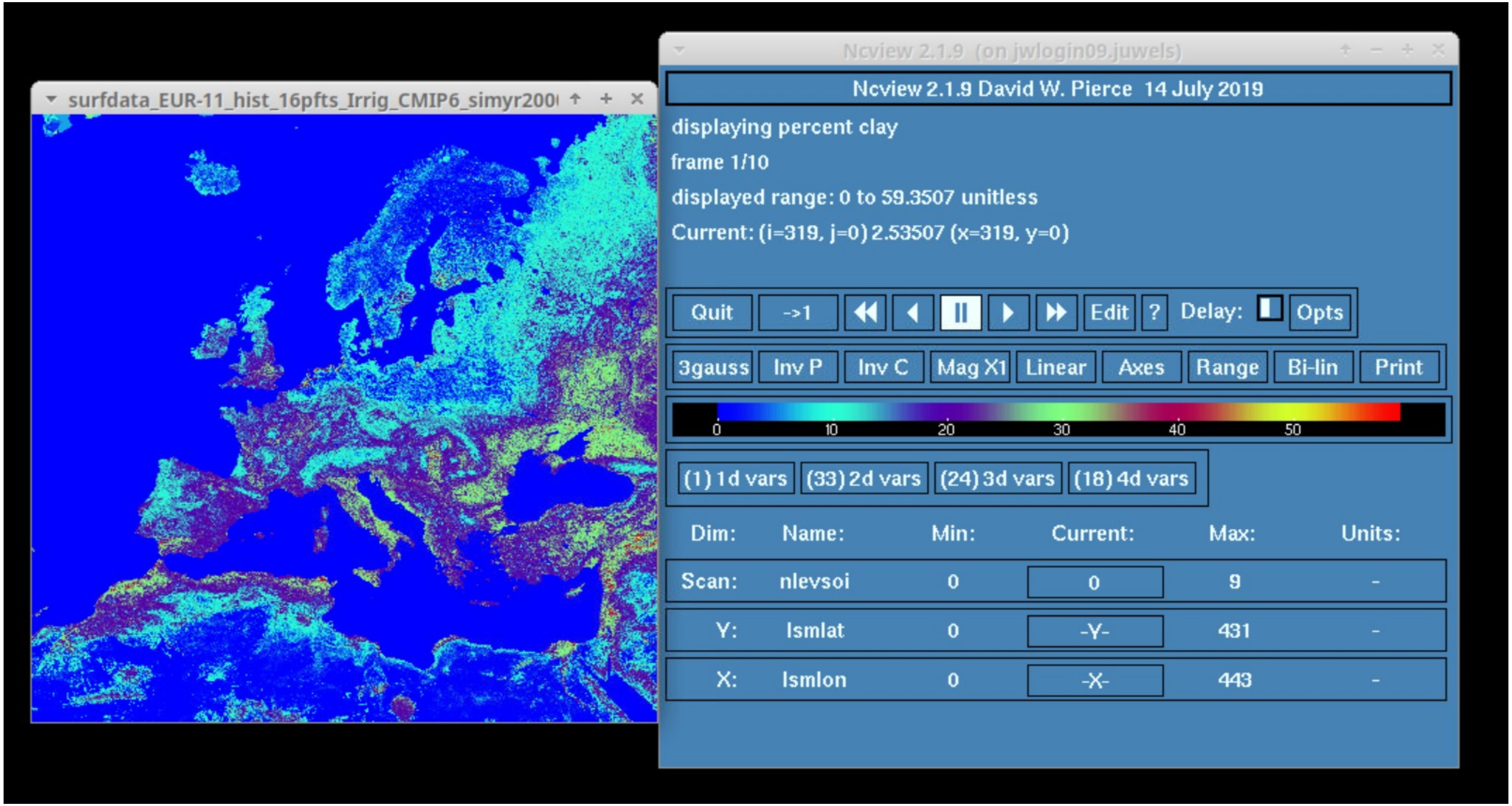
The water removed by transpiration in each layer e_i is a function of the total transpiration E_v^t (Chapter 5) and the effective root fraction $r_{e,\,i}$

$$e_i = r_{e,\,i} E_v^t. \tag{7.67}$$

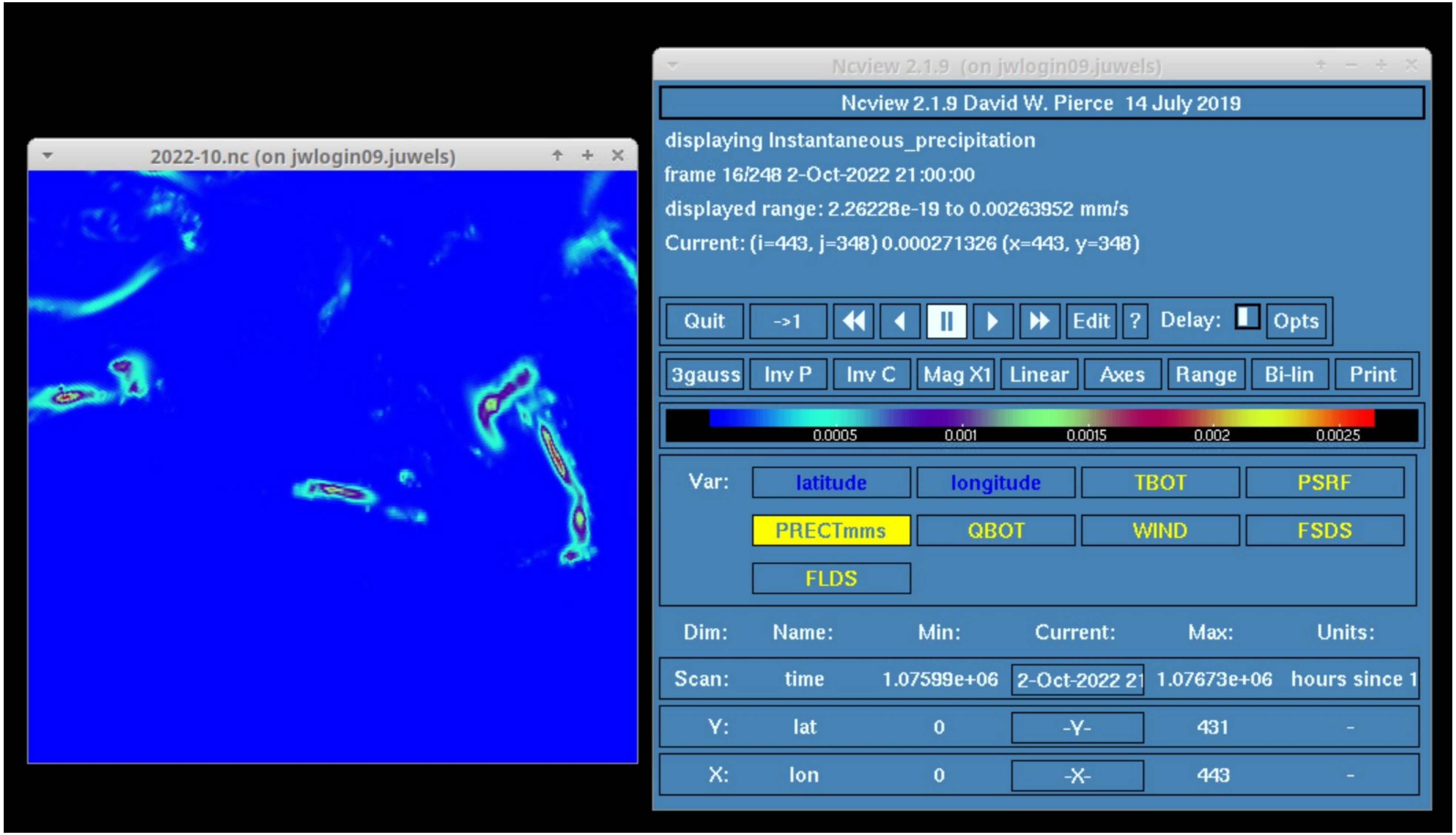
The code

```
do fc = 1,num_hydrologyc
   c = filter_hydrologyc(fc)

   do j = 1, nlevsoi
      h2osoi_liq(c,j) = h2osoi_liq(c,j) + dwat2(c,j)*dzmm(c,j)
   end do

   ! calculate qcharge for case jwt < nlevsoi
   if(jwt(c) < nlevsoi) then
      wh_zwt = 0._r8   !since wh_zwt = -sucsat - zq_zwt, where zq_zwt = -sucsat

      ! Recharge rate qcharge to groundwater (positive to aquifer)
      s_node = max(h2osoi_vol(c,jwt(c)+1)/watsat(c,jwt(c)+1), 0.01_r8)
      s1 = min(1._r8, s_node)

      !scs: this is the expression for unsaturated hk
      ka = imped(c,jwt(c)+1)*hksat(c,jwt(c)+1) &
           *s1**(2._r8*bsw(c,jwt(c)+1)+3._r8)

      !compute unsaturated hk, this shall be tested later, because it
      !is not bit for bit
      !call soil_water_retention_curve%soil_hk(hksat(c,jwt(c)+1), s1, bsw(c,jwt(c)+1), ka)
      !apply ice impedance
      !ka = imped(c,jwt(c)+1) * ka
      ! Recharge rate qcharge to groundwater (positive to aquifer)
      smp1 = max(smpmin(c), smp(c,max(1,jwt(c))))
      wh      = smp1 - zq(c,max(1,jwt(c)))

      !scs: original formulation
      if(jwt(c) == 0) then
         qcharge(c) = -ka * (wh_zwt-wh)  /((zwt(c)+1.e-3)*1000._r8)
      else
         !             qcharge(c) = -ka * (wh_zwt-wh)/((zwt(c)-z(c,jwt(c)))*1000._r8)
         !scs: 1/2, assuming flux is at zwt interface, saturation deeper than zwt
         qcharge(c) = -ka * (wh_zwt-wh)/((zwt(c)-z(c,jwt(c)))*1000._r8*2.0)
      endif

      ! To limit qcharge  (for the first several timesteps)
      qcharge(c) = max(-10.0_r8/dtime,qcharge(c))
      qcharge(c) = min( 10.0_r8/dtime,qcharge(c))
   else
      ! if water table is below soil column, compute qcharge from dwat2(11)
      qcharge(c) = dwat2(c,nlevsoi+1)*dzmm(c,nlevsoi+1)/dtime
   endif
end do

! compute the water deficit and reset negative liquid water content
!  Jinyun Tang
do fc = 1, num_hydrologyc
   c = filter_hydrologyc(fc)
   qflx_deficit(c) = 0._r8
```

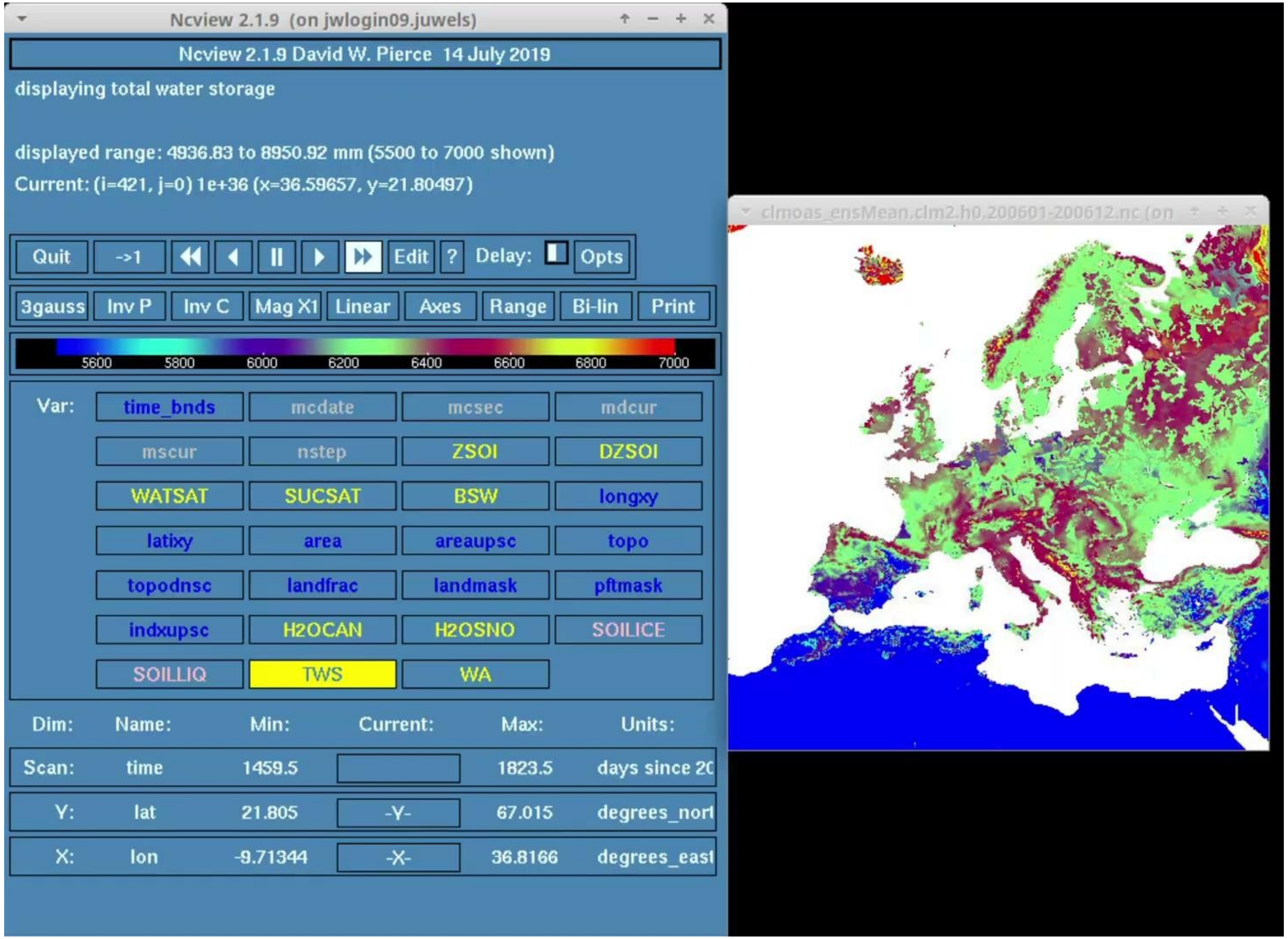
Surface data



Atmospheric forcings



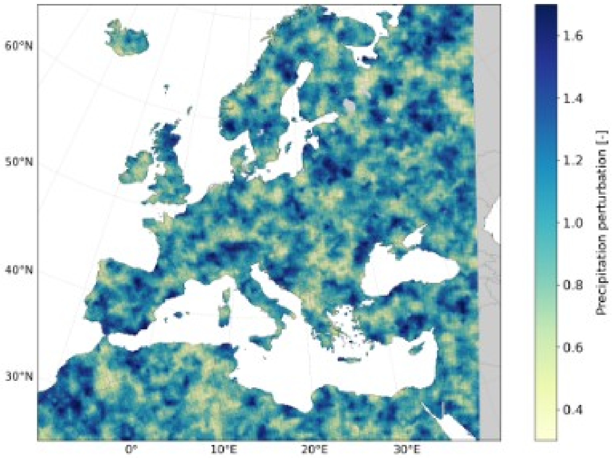
Water storage changes over Europe



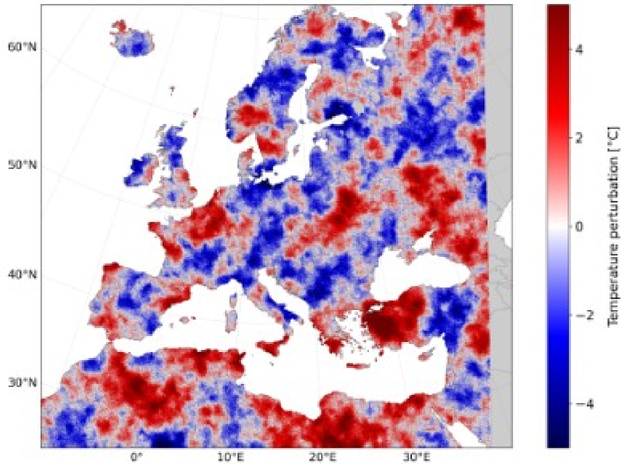
Error representation of hydrological models

Perturbation of
- initial conditions
- atmospheric forcings
- soil properties
- model states

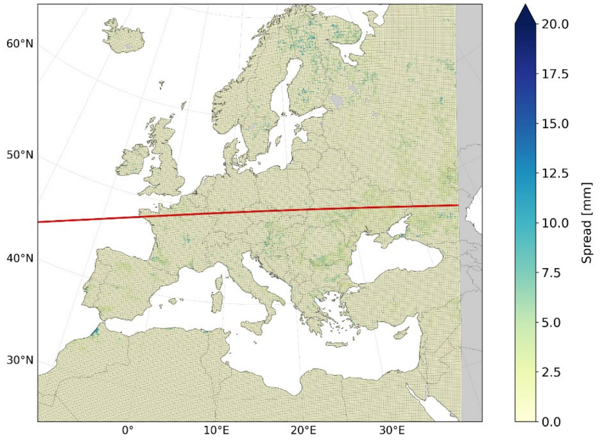
Precipitation perturbation



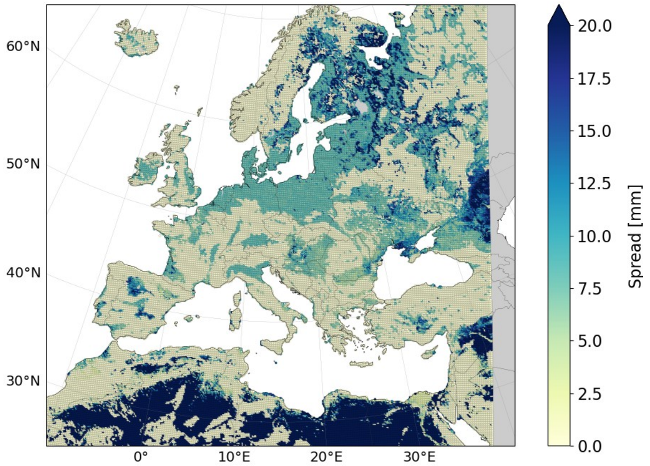
Temperature perturbation



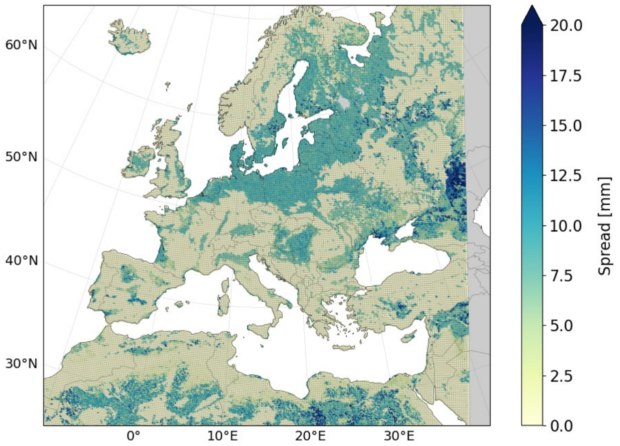
Spread of soil liquid water in deepest soil la



Atmospheric perturbation



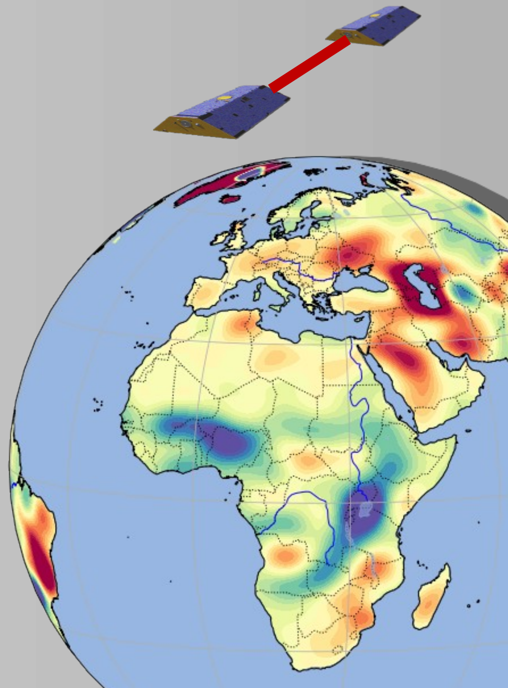
Standard texture perturbation



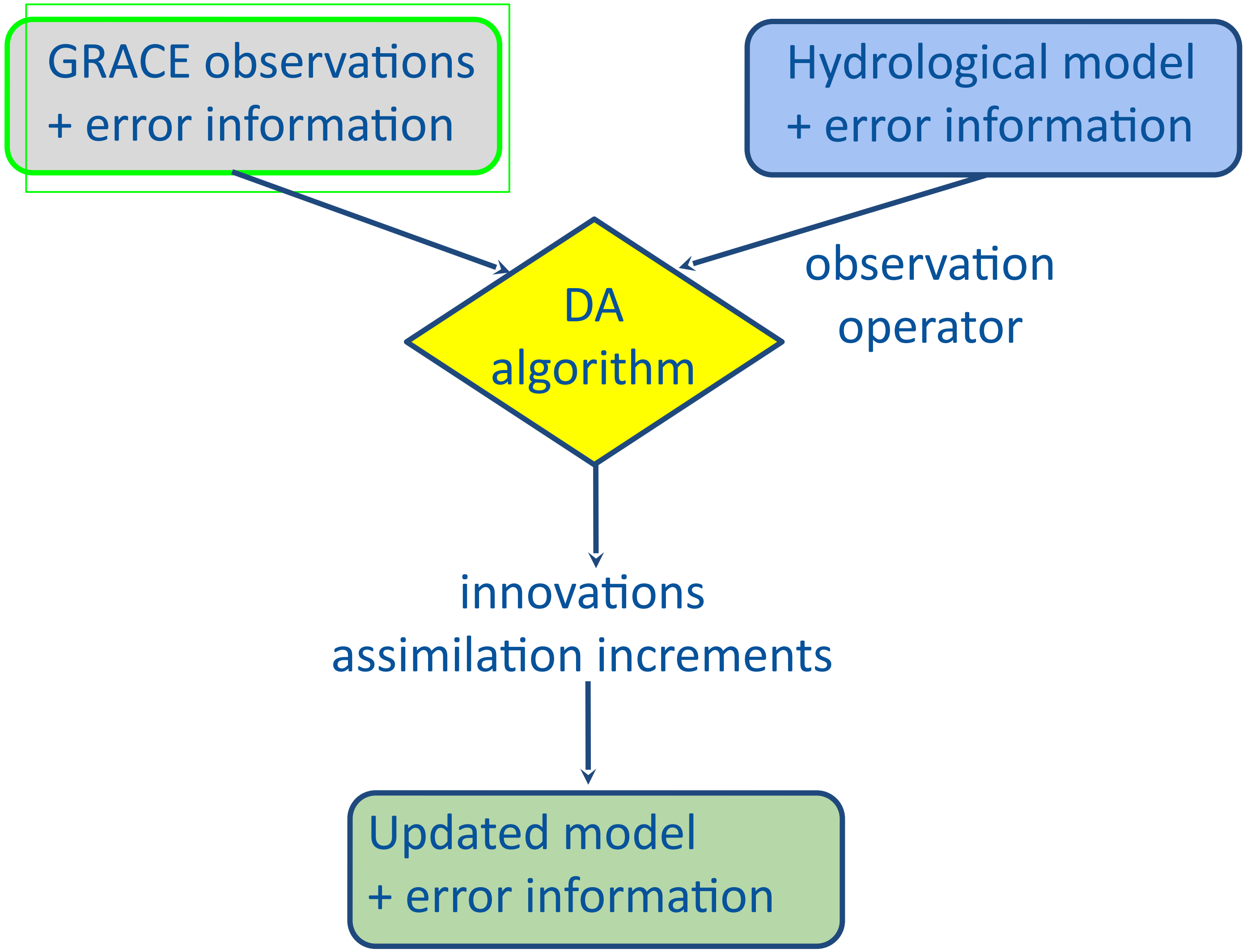
Advanced texture perturbation

Recap hydrology and land surface models

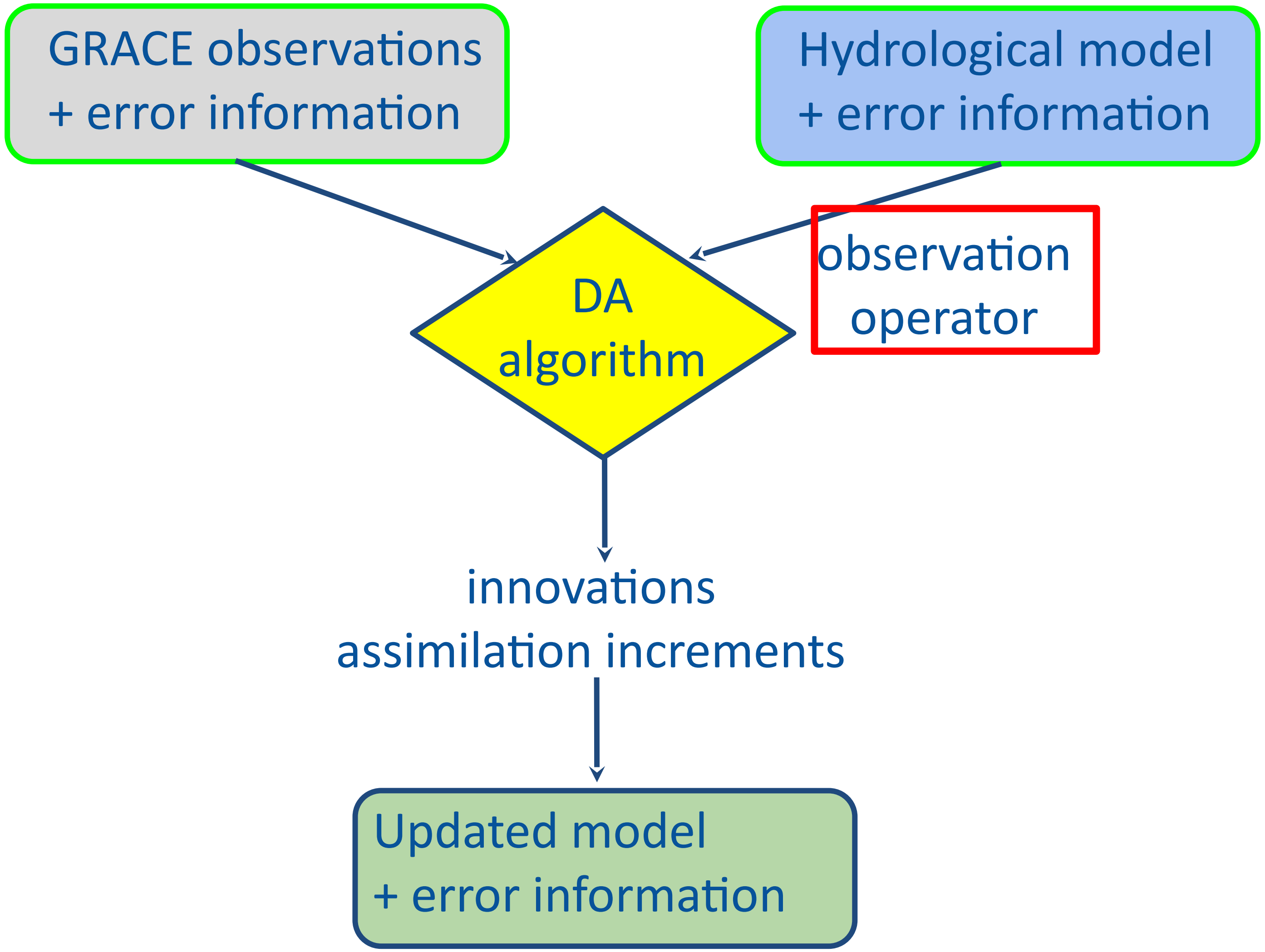
- Models represent different water storages and are available at various spatial and temporal resolutions.
- Models often do not account for / underestimate human impacts.
- Representing the uncertainty of modeled water storages is challenging.



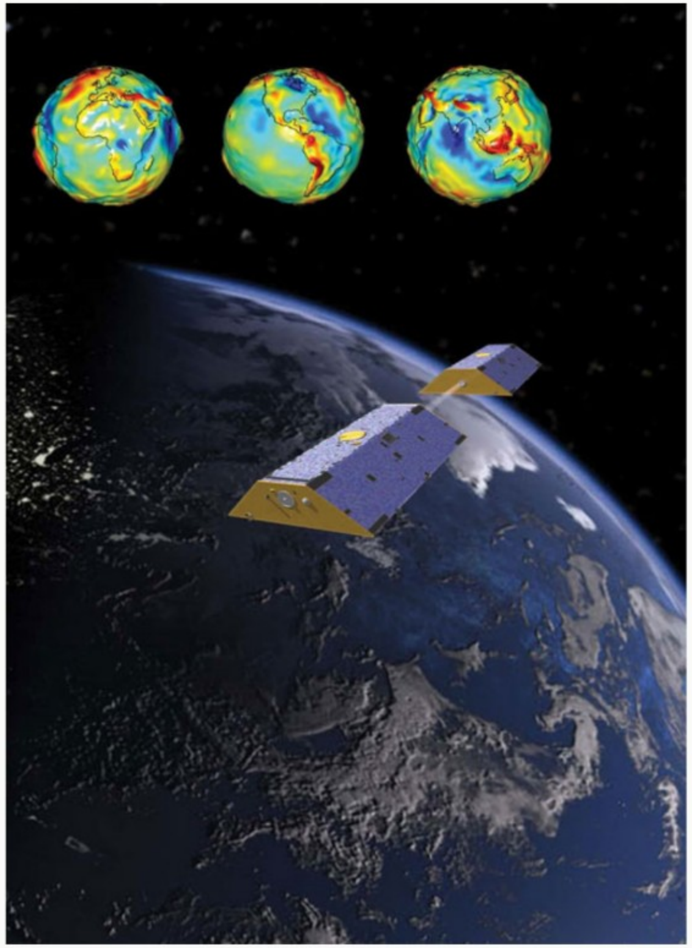
Data assimilation frameworks

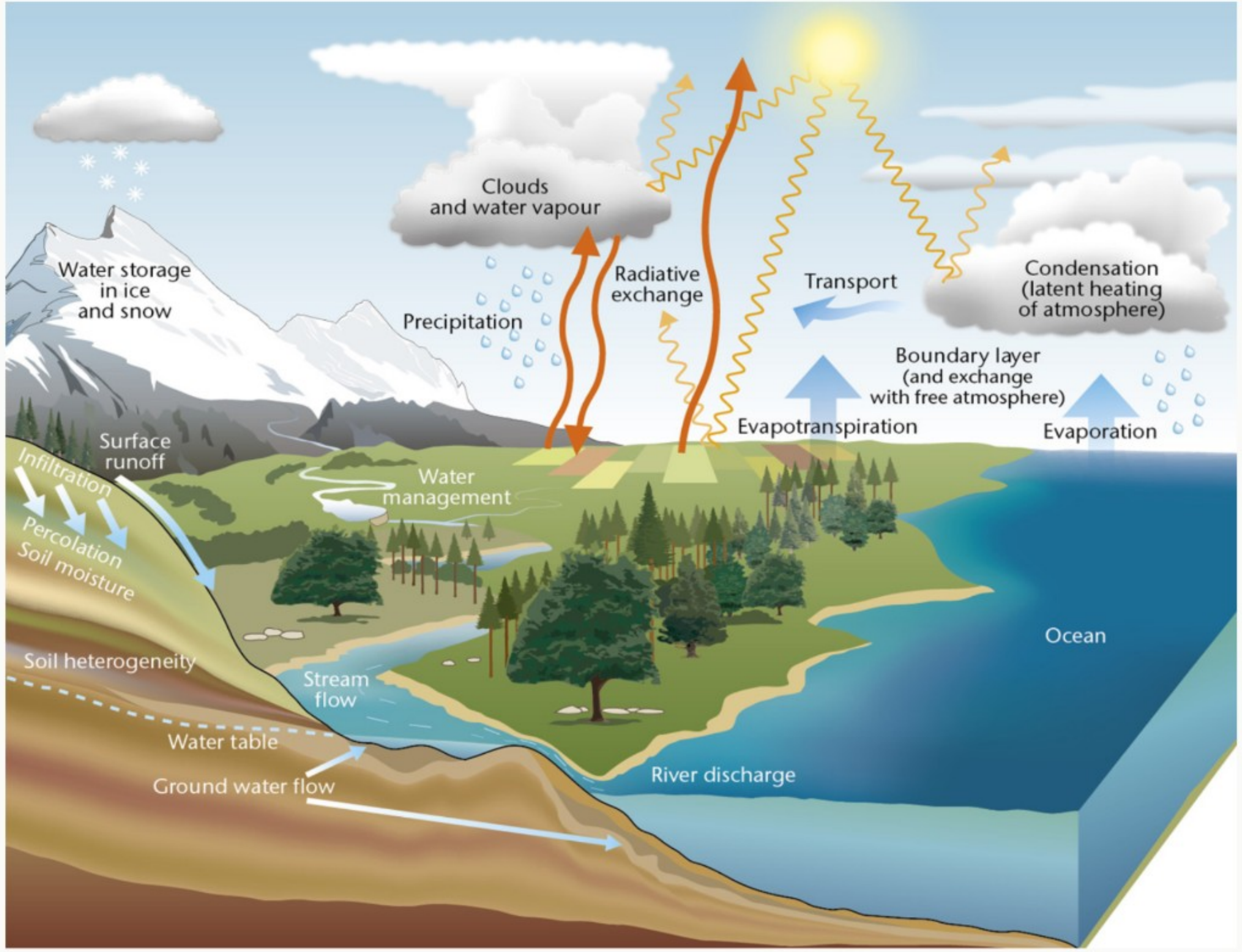


Data assimilation frameworks



Hydrological models versus GRACE observations





snow

ice

surface waters

wetlands

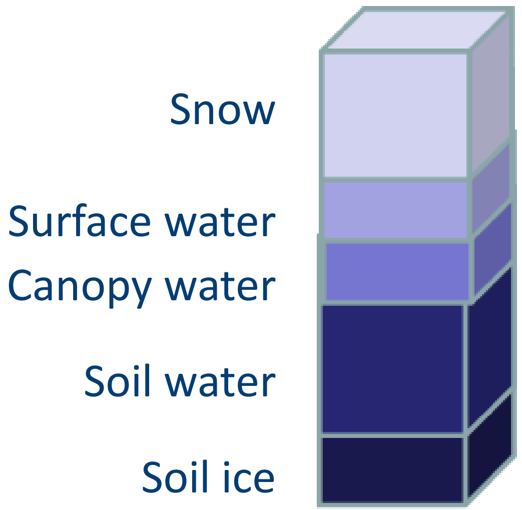
canopy

soil water

soil ice

groundwater

State vector



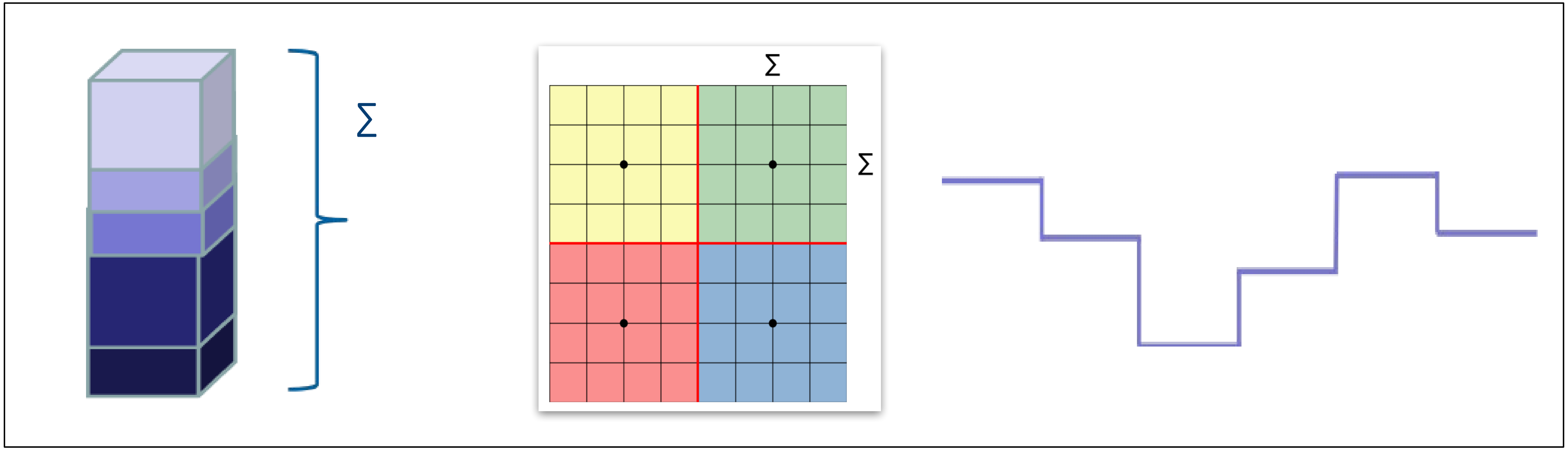
Different options:

- All compartments that contribute to TWS
- Neglect small compartments, e.g. canopy water
- Sum up compartments, e.g. soil water and soil ice in each layer
- Sum up different soil layers
- Put directly TWS into the state vector

Observation operator

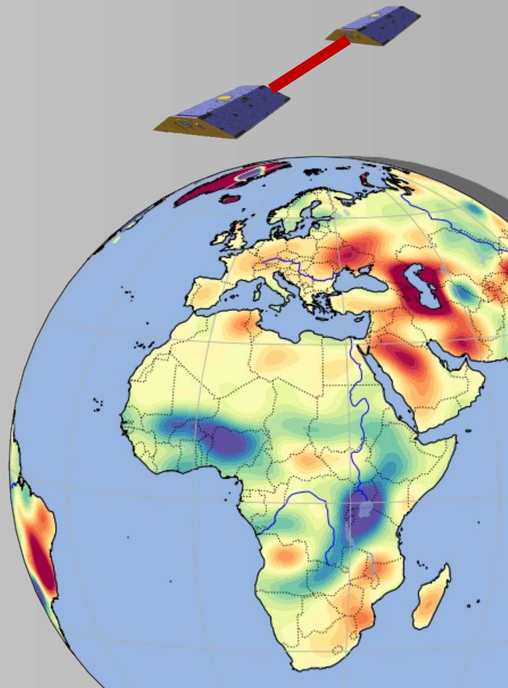
Relate the model output to the observations

- sum up modeled storage compartments vertically
- aggregate modeled storage compartments to the GRACE grid
- compute monthly means of TWS

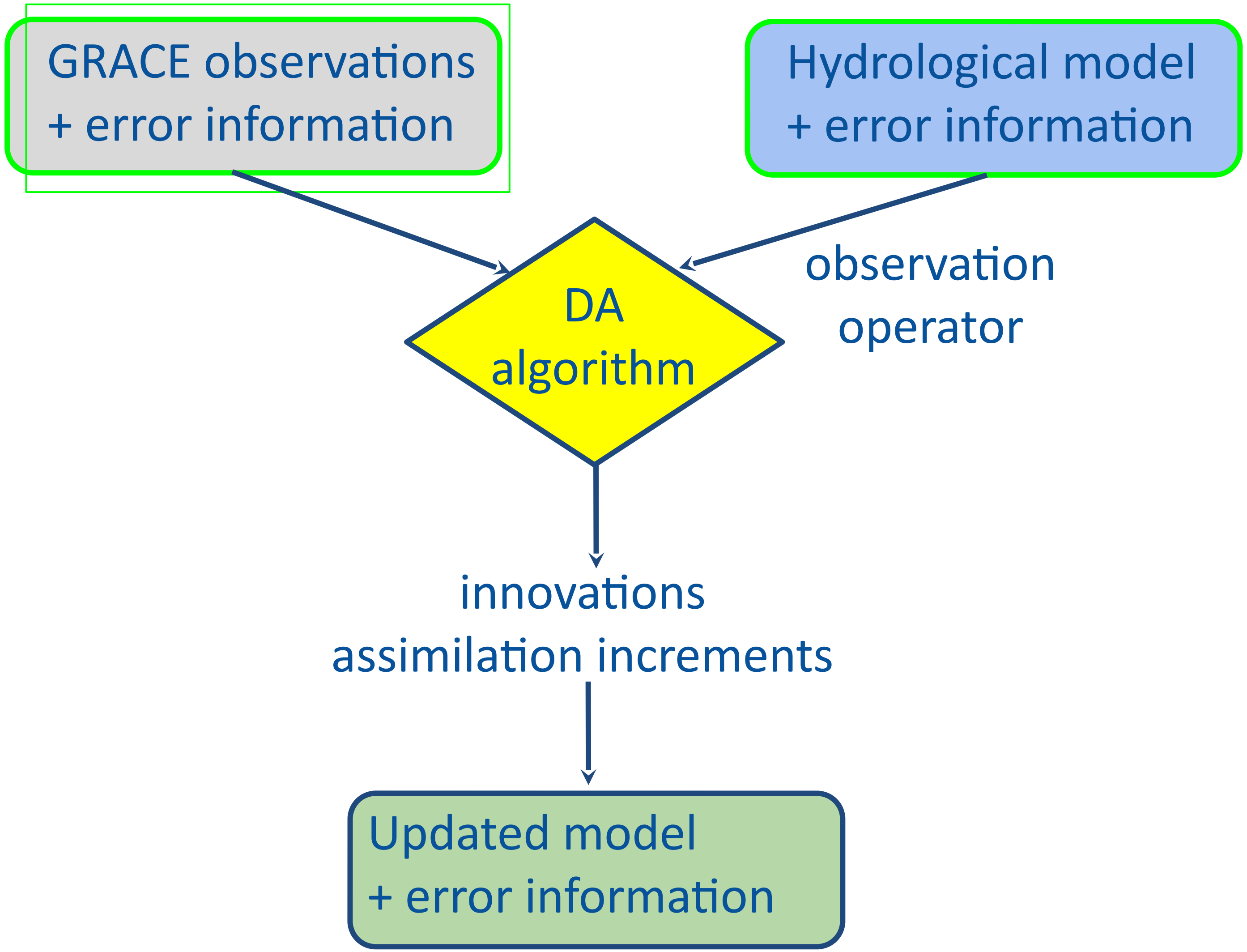


Recap state vector and observation operator

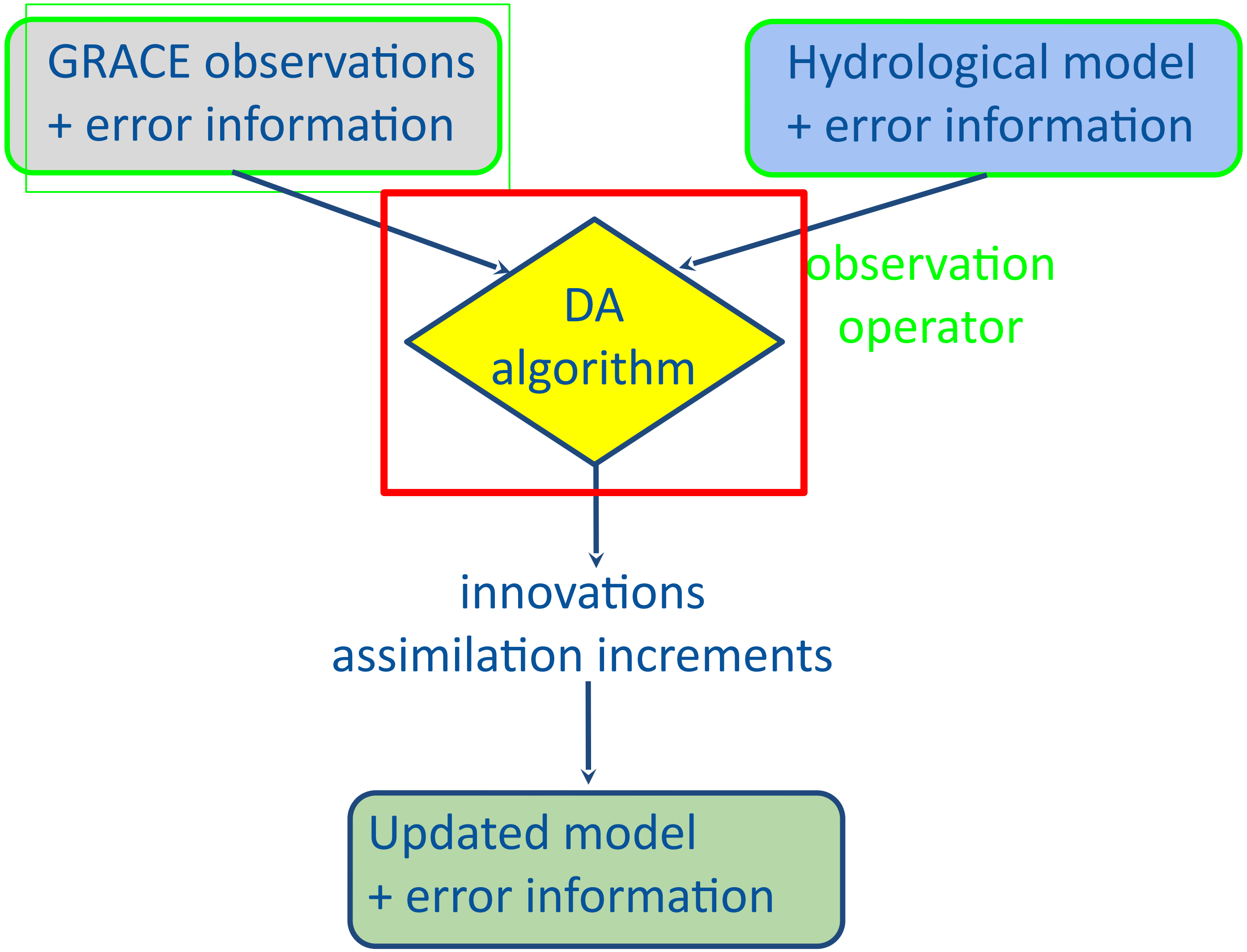
- We obtain updates for all entries of the state vector.
- Computation of the observation operator can be challenging in the case of highly-parallelized model code.



Set up of a data assimilation framework



Set up of a data assimilation framework



We start in 1960 - Development of the Kalman Filter

A New Approach to Linear Filtering and Prediction Problems[1]

R. E. KALMAN
Research Institute for Advanced Study,[2]
Baltimore, Md.

The classical filtering and prediction problem is re-examined using the Bode-Shannon representation of random processes and the "state transition" method of analysis of dynamic systems. New results are:

(1) The formulation and methods of solution of the problem apply without modification to stationary and nonstationary statistics and to growing-memory and infinite-memory filters.

(2) A nonlinear difference (or differential) equation is derived for the covariance matrix of the optimal estimation error. From the solution of this equation the coefficients of the difference (or differential) equation of the optimal linear filter are obtained without further calculations.

(3) The filtering problem is shown to be the dual of the noise-free regulator problem.

The new method developed here is applied to two well-known problems, confirming and extending earlier results.

The discussion is largely self-contained and proceeds from first principles; basic concepts of the theory of random processes are reviewed in the Appendix.

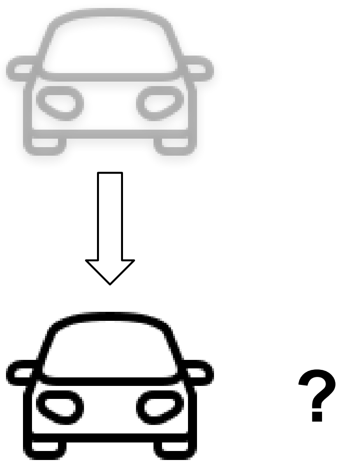
Journal of Basic Engineering, March 1960

The Kalman Filter - Early applications

- **Apollo Space Program** (1960s): Position and velocity of the lunar module using onboard sensors
- **Missile guidance and radar tracking** (1960s - 1970s): military applications, positions and velocities under noisy conditions
- **Control systems** (1970s): aircraft autopilot systems, navigation systems
- **GPS navigations** (1980s - Present): core component!
- **Atmospheric dynamics** (1970s): application to low order atmospheric models, state estimation incorporating sparse meteorological observations
- **River flow and hydrological modeling** (1970s): streamflow forecasting, flood prediccitons
- **Soil moisture estimation** (1980s): assimilation of rainfall and evapotranspiration data

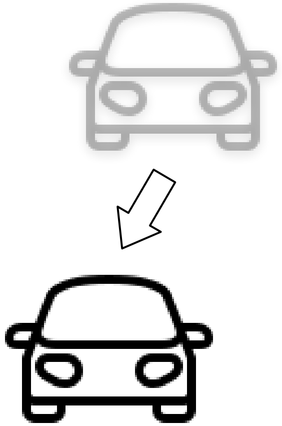
> AN IMPORTANT class of theoretical and practical problems in communication and control is of a statistical nature. Such problems are: (i) Prediction of random signals; (ii) separation of random signals from random noise; (iii) detection of signals of known form (pulses, sinusoids) in the presence of random noise.
>
> Kalman, 1960

Toy example



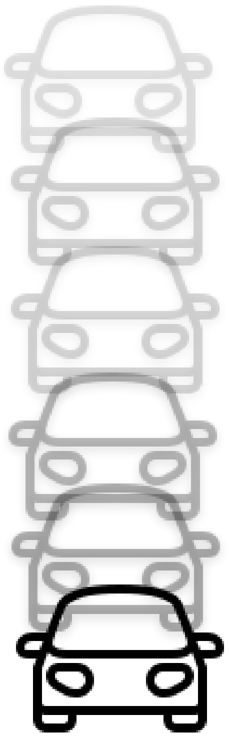
- Car drives, what is the position at time t+1 if we know the position at time t?

Toy example



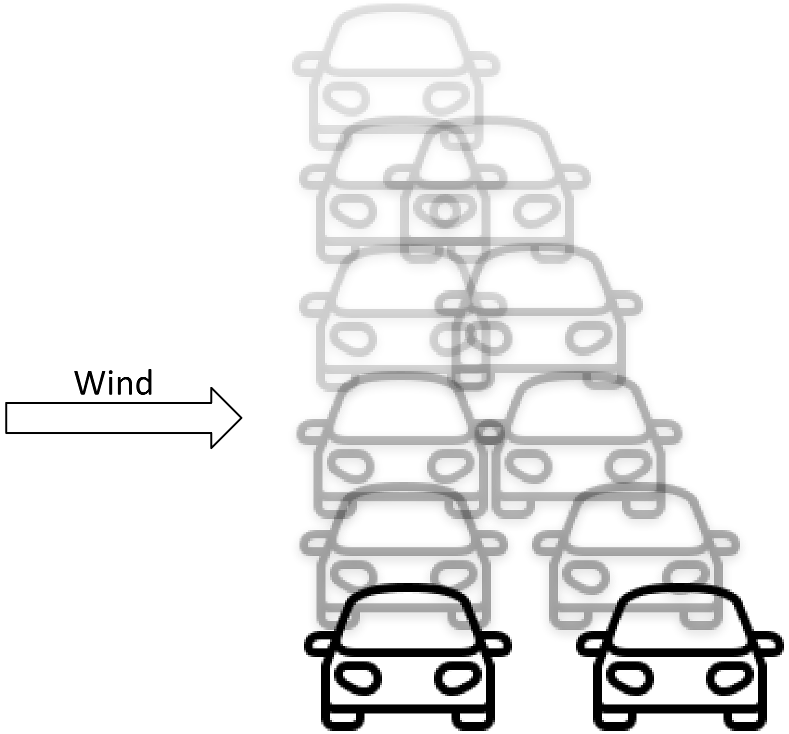
- Car drives, what is the position at time t+1 if we know the position at time t?
- system equation (model)
  - Steering change → orientation
  - Deceleration/acceleration → velocity

Toy example



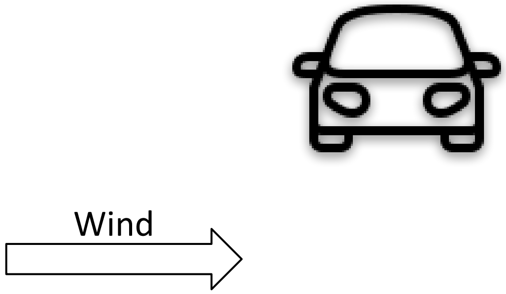
- Car drives, what is the position at time t+1 if we know the position at time t?
- system equation (model)
  - Steering change → orientation
  - Deceleration/acceleration → velocity

Toy example



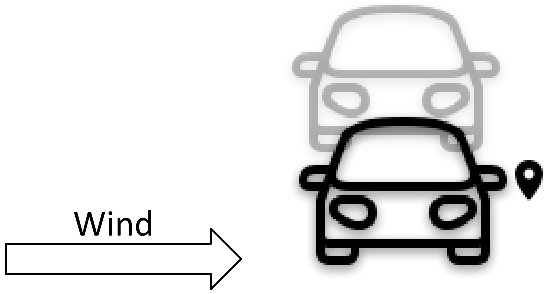
- Car drives, what is the position at time t+1 if we know the position at time t?
- system equation (model)
  - Steering change → orientation
  - Deceleration/acceleration → velocity
- Problems: Unmodeled processes (e.g. wind, friction,...), systematics, imperfect input data

Toy example



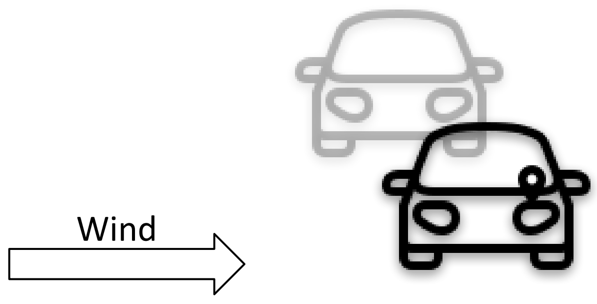
- Car drives, what is the position at time t+1 if we know the position at time t?
- system equation (model)
  - Steering change → orientation
  - Deceleration/acceleration → velocity
- Problems: Unmodeled processes (e.g. wind, friction,...), systematics, imperfect input data

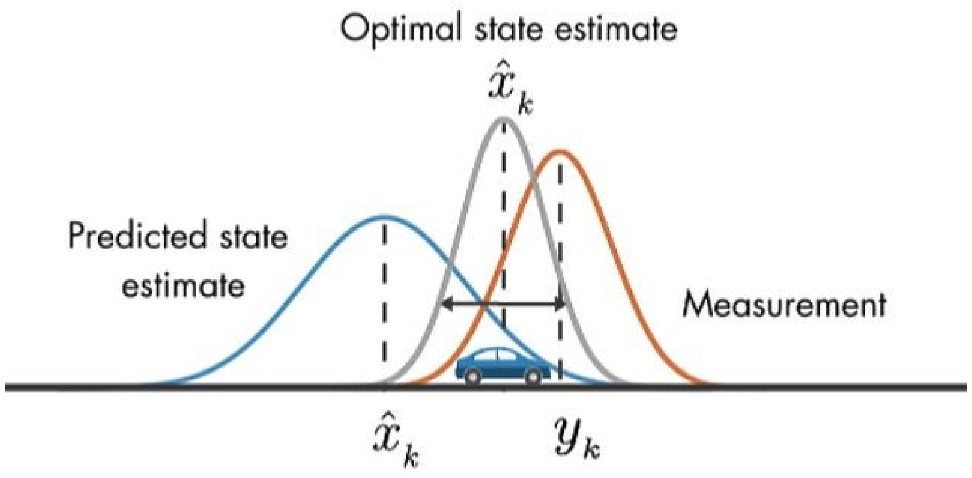
Toy example



- Car drives, what is the position at time t+1 if we know the position at time t?
- system equation (model)
  - Steering change → orientation
  - Deceleration/acceleration → velocity
- Problems: Unmodeled processes (e.g. wind, friction,...), systematics, imperfect input data

Toy example





- Car drives, what is the position at time t+1 if we know the position at time t?
- system equation (model)
  - Steering change → orientation
  - Deceleration/acceleration → velocity
- Problems: Unmodeled processes (e.g. wind, friction,...), systematics, imperfect input data
- Update model estimate based on observations

Kalman Filter Basics - Bayes Theorem

$\mathbf{x}^T = (x_1^T, ..., x_N^T)$ ... model state vector: contains all uncertain quantities we wish to estimate

$\mathbf{y}$ ... observation vector

Bayes theorem

wanted!

$$\boxed{p\left(\mathbf{x}|\mathbf{y}\right)} \propto p\left(\mathbf{y}|\mathbf{x}\right) p\left(\mathbf{x}\right)$$

posterior PDF of the model states depending on the observations

observation or likelihood PDF given the model states

model forecast or prior PDF

PDF ... probability density function, Gaussian

Kalman Filter Basics - Model state and observations

$\mathbf{x}_k^f$ … forecasted model state at time k

$\mathbf{q}_k$ … model error assuming white Gaussian noise with zero mean

$\mathbf{y}_k$ … observations at time k

ϵ_k … observation error vector

$\mathbf{x}_k^a$ … analyzed model state at time k

$$\mathbf{x}_k^f = \mathcal{M}\mathbf{x}_{k-1}^f + \mathbf{q}_k$$

… forecasted model state based on the previous time step using **model operator** $\mathcal{M}$

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k^f + \epsilon_k$$

… linking model state to observations with the **observation operator** $\mathbf{H}_k$

Kalman Filter Basics - linear case

$$\mathbf{x}_k^f = \mathbf{M}\mathbf{x}_{k-1}^f + \mathbf{q}_k$$ ... forecasted model state

$$\mathbf{C}_{xx}^f(k) = \mathbf{M}^T\mathbf{C}_{xx}^f(k-1)\mathbf{M} + \mathbf{Q}_k$$ ... predicted model error covariance matrix at time step k

Cost function:

$$J(\mathbf{x}_k^a) = \left(\mathbf{x}_k^f - \mathbf{x}_k^a\right)^T \mathbf{C}_{xx}^f(k)^{-1}\left(\mathbf{x}_k^f - \mathbf{x}_k^a\right) + (\mathbf{y}_k - \mathbf{H}_k\mathbf{x}_k^a)^T\,\mathbf{C}_{yy}(k)^{-1}\,(\mathbf{y}_k - \mathbf{H}_k\mathbf{x}_k^a)$$

Kalman Filter Basics - fundamental equations

Cost function:

$$J(\mathbf{x}_k^a) = \left(\mathbf{x}_k^f - \mathbf{x}_k^a\right)^T \mathbf{C}_{xx}^f(k)^{-1} \left(\mathbf{x}_k^f - \mathbf{x}_k^a\right) + (\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k^a)^T \, \mathbf{C}_{yy}(k)^{-1} \, (\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k^a)$$

$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k \mathbf{d}_k$$ … analyzed model state

$$\mathbf{d}_k = \mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k^f$$ … innovations

$$\mathbf{K}_k = \mathbf{C}_{xx}^f(k) \mathbf{H}_k^T \left(\mathbf{H}_k \mathbf{C}_{xx}^f(k) \mathbf{H}_k^T + \mathbf{C}_{yy}(k)\right)^{-1}$$ … Kalman Gain Matrix

$$\mathbf{C}_{xx}^a(k) = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \, \mathbf{C}_{xx}^f(k)$$ … covariance matrix of the analyzed model state

The Extended Kalman Filter

In reality, we often have to deal with **nonlinear models**.

- The model forecast equations are linearized using **Taylor expansion**. This is computationally expensive.
- Analysis steps are identical to those of the linear Kalman filter. But the analysis equations represent only an **approximation to the optimal estimate**.
- Neglected higher order terms can lead to an unrealistic representation of the model error and to unbounded error growth leading to **instabilities** in the filter algorithm.

↓

Development of the **Ensemble Kalman Filter**:

- Handling strongly nonlinear systems more accurately.
- Reducing the high computational cost of the traditional Kalman filter.
- ➢ Approximation of the model error covariance matrix by a finite set of ensemble members, avoiding the need to store and update a massive covariance matrix.

We go back to 1994 - Development of the EnKF

JOURNAL OF GEOPHYSICAL RESEARCH
Oceans
AN AGU JOURNAL

Regular Section

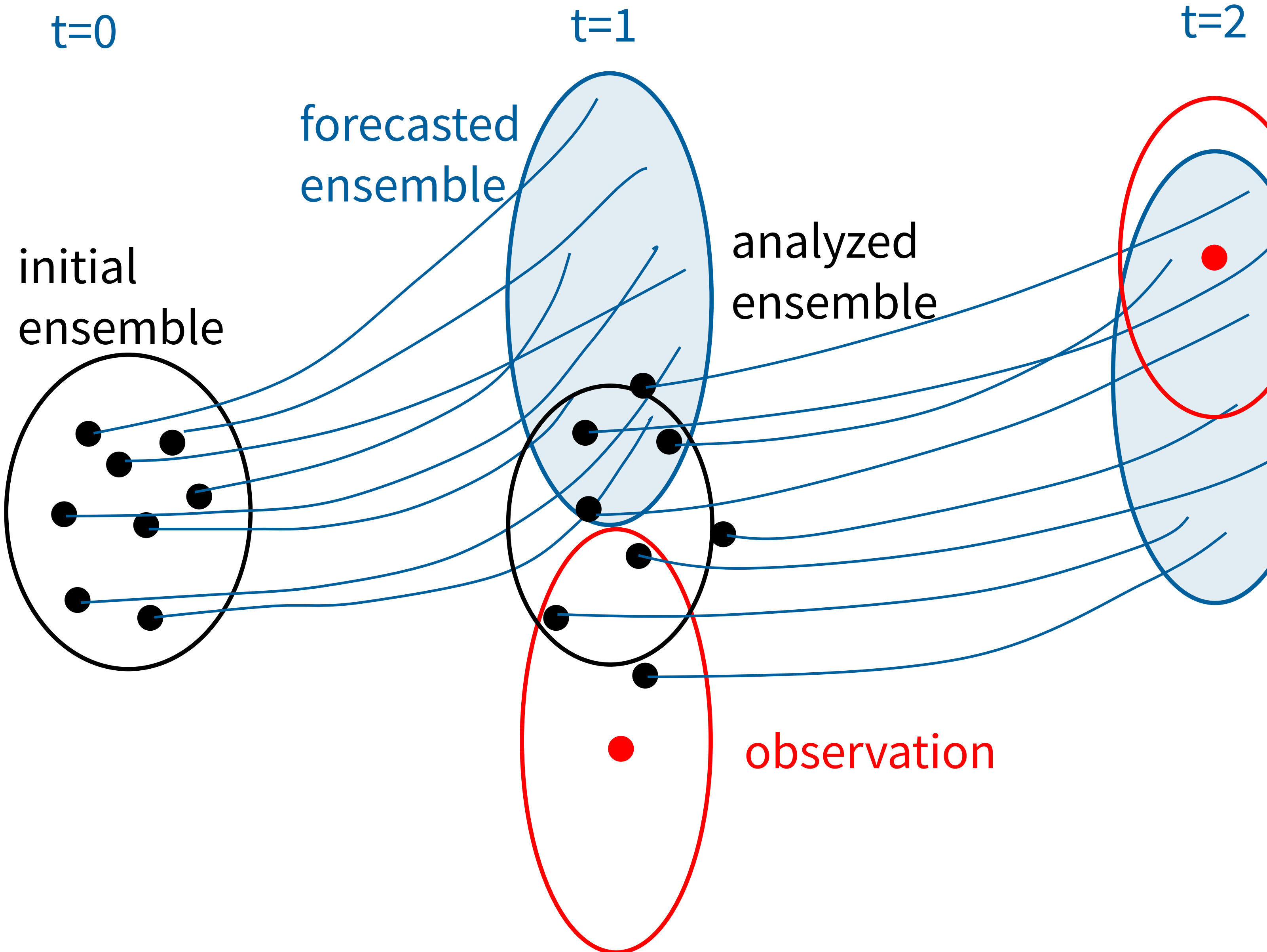
Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics

Geir Evensen

First published: 15 May 1994 | https://doi.org/10.1029/94JC00572 | Citations: 3,662

- Developed for non-linear ocean models, e.g., improved representation of the Gulf Stream behavior
- Later applied to meteorology, hydrology, and petroleum engineering
- Became a standard tool in flood forecasting, groundwater modeling, and land surface modeling

Ensemble Kalman Filter



The Ensemble Kalman Filter

$$\tilde{\mathbf{C}}^f_{xx}(k) = \frac{1}{N_{e-1}} \sum_{i=1}^{N_e} \left(\mathbf{x}_k^{f(i)} - \overline{\mathbf{x}}_k^f\right)\left(\mathbf{x}_k^{f(i)} - \overline{\mathbf{x}}_k^f\right)^T$$

... model error covariance matrix

$$\mathbf{X}^a = \mathbf{X}^f + \mathbf{KD} = \mathbf{X}^f + \underbrace{\tilde{\mathbf{C}}^f_{xx}\mathbf{H}^T\left(\mathbf{H}\tilde{\mathbf{C}}^f_{xx}\mathbf{H}^T + \mathbf{C}_{yy}\right)^{-1}}_{\text{Kalman Gain Matrix}} \underbrace{\left(\mathbf{Y} - \mathbf{H}\mathbf{X}^f\right)}_{\text{innovations}}$$

... analysis equation

Ensemble of model states

Square root formulation for efficient implementation:

$$\mathbf{X}'^f = \mathbf{X}^f - \overline{\mathbf{X}^f}$$

... ensemble perturbation matrix

$$\tilde{\mathbf{C}}^f_{xx} = \frac{\mathbf{X}'^f\left(\mathbf{X}'^f\right)^T}{N_e - 1}$$

... model error covariance matrix

$$\mathbf{X}^a = \mathbf{X}^f + \frac{1}{N_e - 1}\mathbf{X}'^f\left(\mathbf{X}'^f\right)^T\mathbf{H}^T\left(\frac{1}{N_e - 1}\mathbf{H}\mathbf{X}'^f\left(\mathbf{X}'^f\right)^T\mathbf{H}^T + \mathbf{C}_{yy}\right)^{-1}\left(\mathbf{Y} - \mathbf{H}\mathbf{X}^f\right)$$

The Ensemble Kalman Filter: Square root formulation

$$\mathbf{X}^a = \mathbf{X}^f + \frac{1}{N_e - 1}\mathbf{X}'^f \left(\mathbf{X}'^f\right)^T \mathbf{H}^T \left(\underbrace{\frac{1}{N_e - 1}\underbrace{\mathbf{H}\mathbf{X}'^f}_{\mathbf{S}} \underbrace{\left(\mathbf{X}'^f\right)^T \mathbf{H}^T}_{\mathbf{S}^T} + \mathbf{C}_{yy}}_{\mathbf{F}}\right)^{-1} \left(\mathbf{Y} - \mathbf{H}\mathbf{X}^f\right)$$

innovation covariance matrix

$$\mathbf{X}^a = \mathbf{X}^f + \frac{1}{N_e - 1}\mathbf{X}'^f \mathbf{S}^T \mathbf{F}^{-1} \left(\mathbf{Y} - \mathbf{H}\mathbf{X}^f\right)$$

... analysis step in square root formulation

$$\mathbf{X}'^a \left(\mathbf{X}'^a\right)^T = \mathbf{X}'^f \left(\mathbf{T}_{EnKF}\mathbf{T}_{EnKF}^T\right) \left(\mathbf{X}'^f\right)^T$$

... analyzed ensemble perturbations

$$\mathbf{T}_{EnKF}\mathbf{T}_{EnKF}^T = \mathbb{1} - \mathbf{S}^T \mathbf{F}^{-1} \mathbf{S}$$

transform matrix

Variants of the EnKF

- Improve stability, reduce sampling noise, efficiently update ensemble perturbations
- Avoid the need for perturbed observations
- Different computations of the transform matrix $\mathbf{T}$

- Ensemble Transform Kalman Filter (ETKF)
  - ensemble perturbations are transformed using a deterministic transformation matrix
- Singular Evolutive Interpolated Kalman Filter (SEIK)
  - Analysis step is performed in the ensemble error subspace
  - More efficient computation
  - More robust for small ensemble sized
- Error Subspace Transform Kalman Filter (ESTKF)
  - favorable combination of ETKF and SEIK
  - improved consistency and stability

- **Do not use the EnKF for GRACE DA, prefer the ESTKF!**

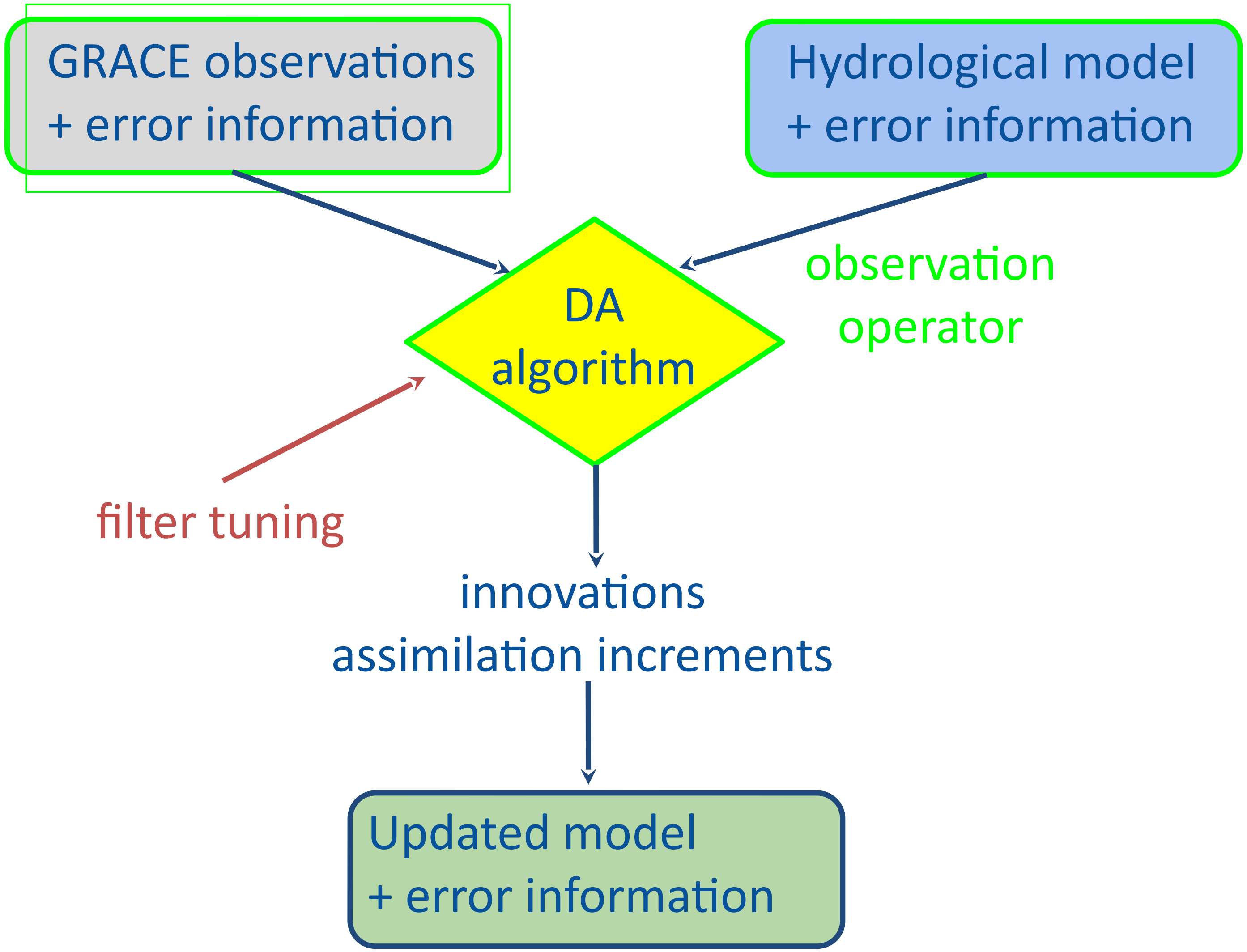
How to implement?

PDAF Parallel
Data Assimilation
Framework

- Implementation of the analysis step
  - Implementation for ESTKF
  - Implementation for LESTKF
  - Implementation for ETKF
  - Implementation for LETKF
  - Implementation for SEIK
  - Implementation for LSEIK
  - Implementation for SEEK
  - Implementation for EnKF
  - Implementation for LEnKF
  - Implementation for NETF
  - Implementation for LNETF
  - Implementation for PF
  - Implementation for 3D-Var
  - Implementation for 3D Ensemble Var
  - Implementation for Hybrid 3D-Var

- Provides fully implemented, parallelized, and optimized ensemble-based algorithms
- Online and offline mode available
- Provides a number of different filters and tuning options
- Provides the so-called “OMI” interface for simultaneous assimilation of different observation types

Set up of a data assimilation framework



Tuning of the Filter Algorithm: Inflation

- Counteract excessive variance reduction caused by spurious correlations in the updates
- Avoid filter divergence in operational ensemble DA systems with small ensemble sizes
- Multiplicative inflation factor:

$$\mathbf{x}_i^f = r\left(\mathbf{x}_i^f - \overline{\mathbf{x}}^f\right) + \overline{\mathbf{x}}^f$$

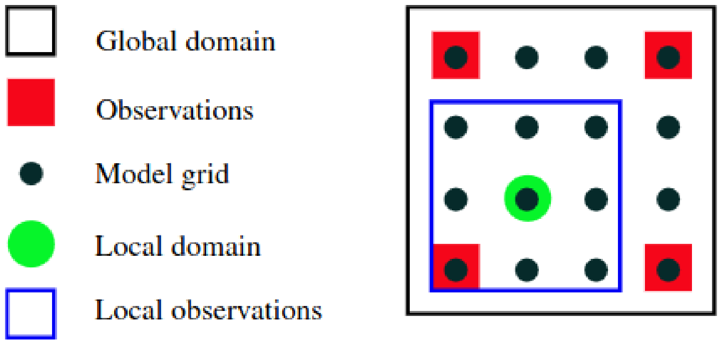
- Adaptive inflation factors are spatially and temporally variable

Tuning of the Filter Algorithm: Localization

- Mitigate spurious long range correlations
- Restrict influence of observations on nearby grid cells
- Improve performance with limited ensemble size

Domain localization

- analysis step is applied to model subdomains



Kirchgessner et al. (2012)

Covariance localization

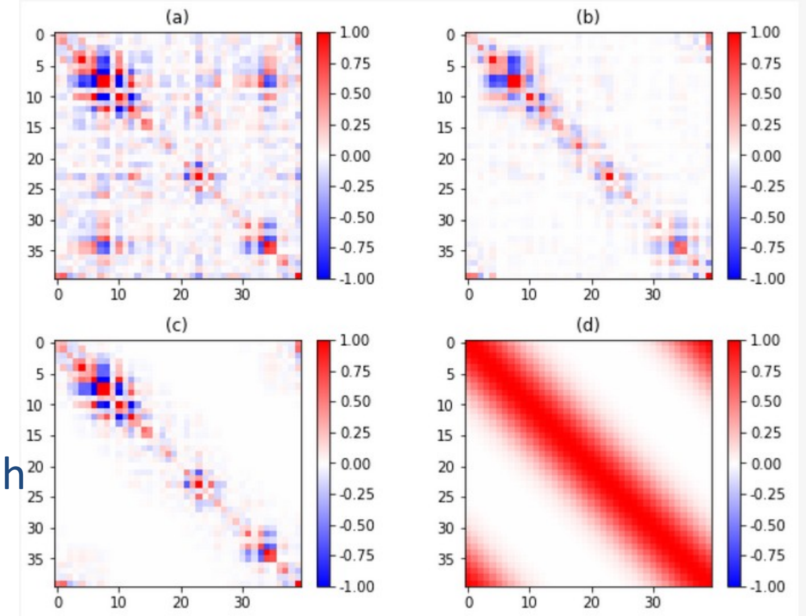
- a weight matrix is applied to the model covariance matrix

20 ensemble members

100 ensemble members

20 ensemble members with localization

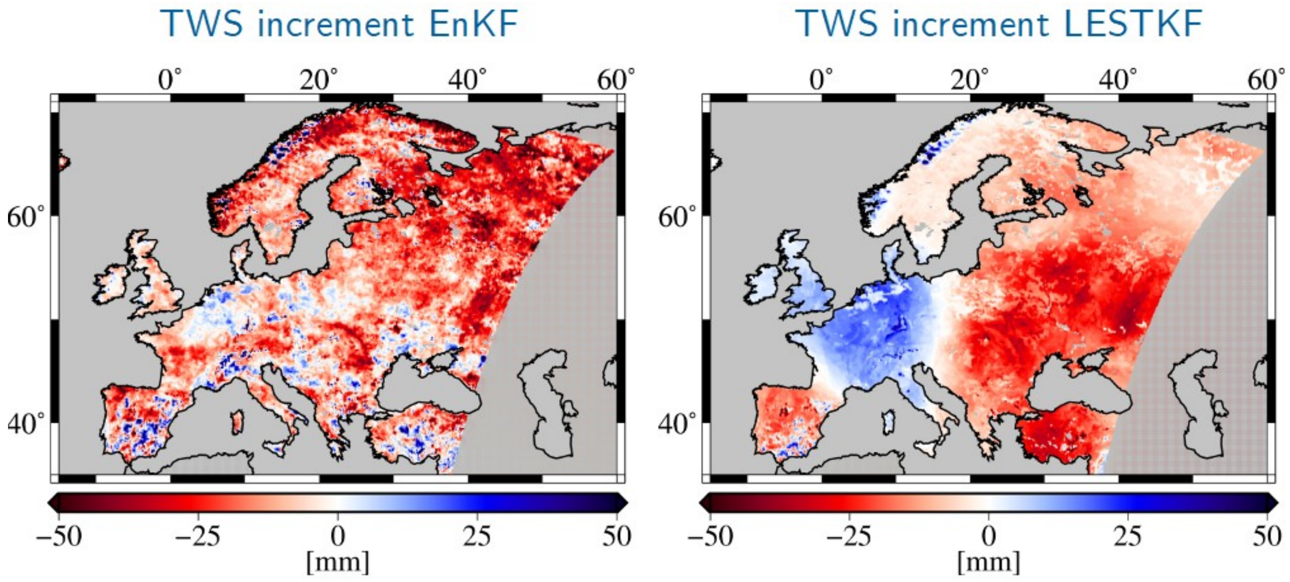
localization matrix



Xing et al. (2021)

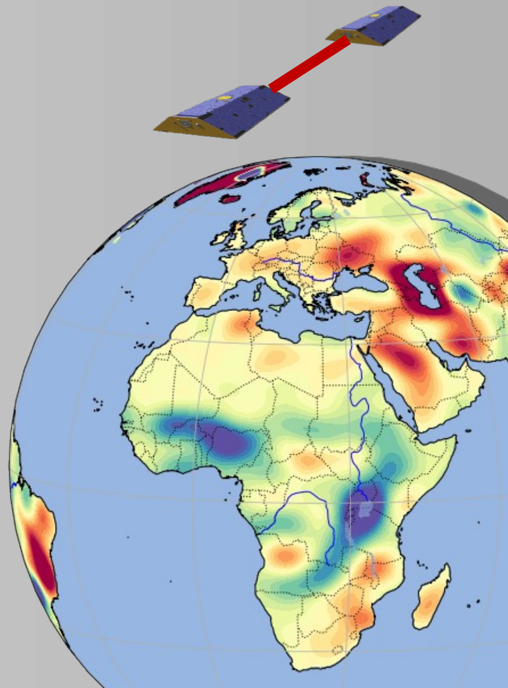
Tuning of the Filter Algorithm: Localization

- Mitigate spurious long range correlations
- Restrict influence of observations on nearby grid cells
- Improve performance with limited ensemble size

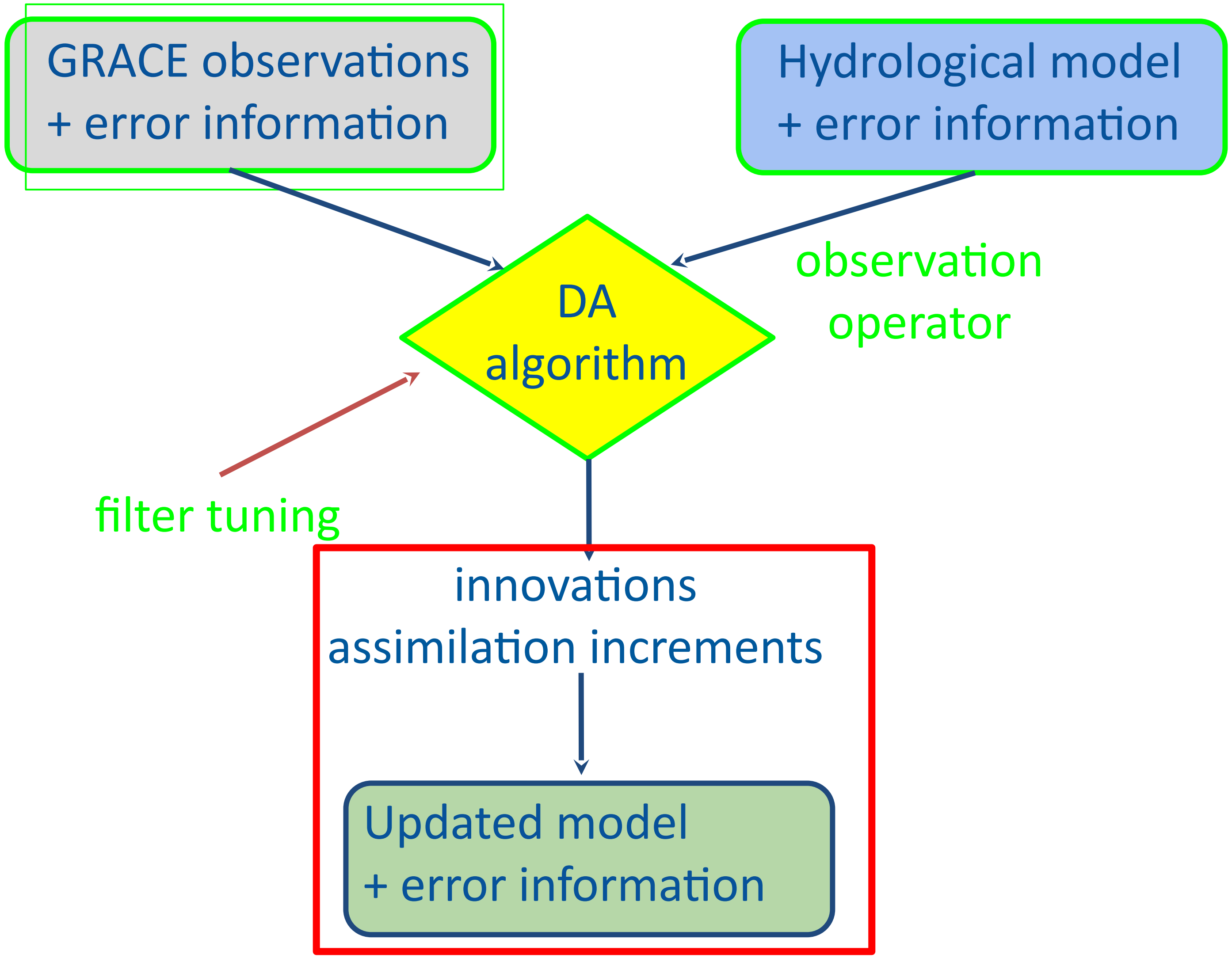


Recap data assimilation algorithm

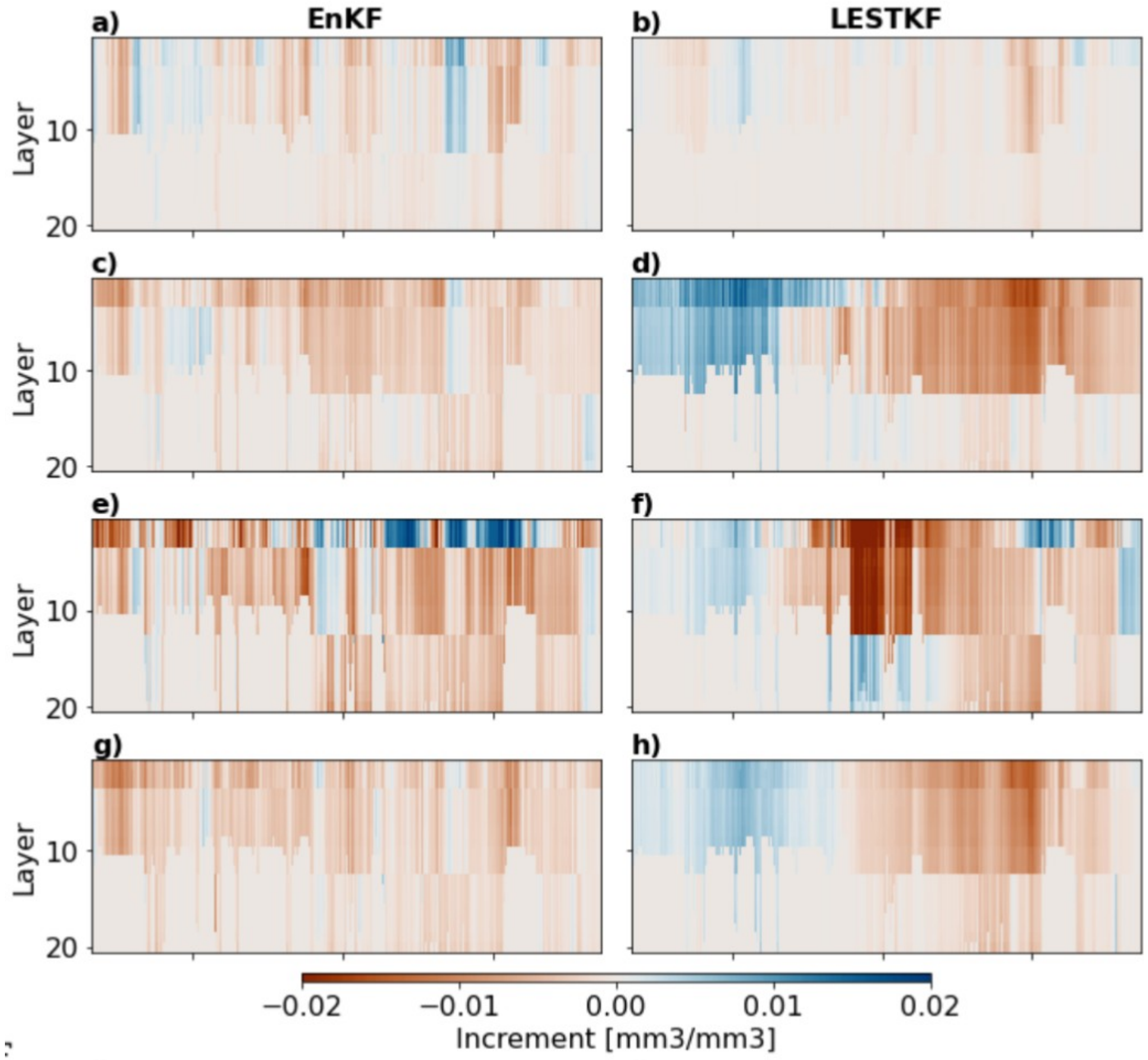
- Most frameworks use ensemble Kalman Filter based algorithms for GRACE/-FO data assimilation.
- Efficient toolboxes with implementations of various filters are available.
- Do not use the EnKF, but rather the ESTKF.
- Localization is very beneficial for GRACE/-FO data assimilation.



Set up of a data assimilation framework



Assimilation increments depend on the settings of the DA framework

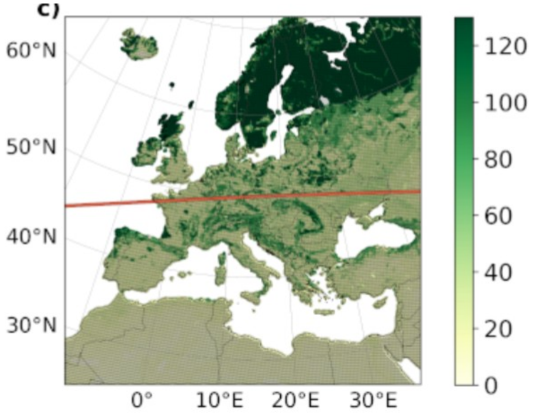


atmospheric perturbation

standard texture perturbation

advanced perturbation of soil properties, case 1

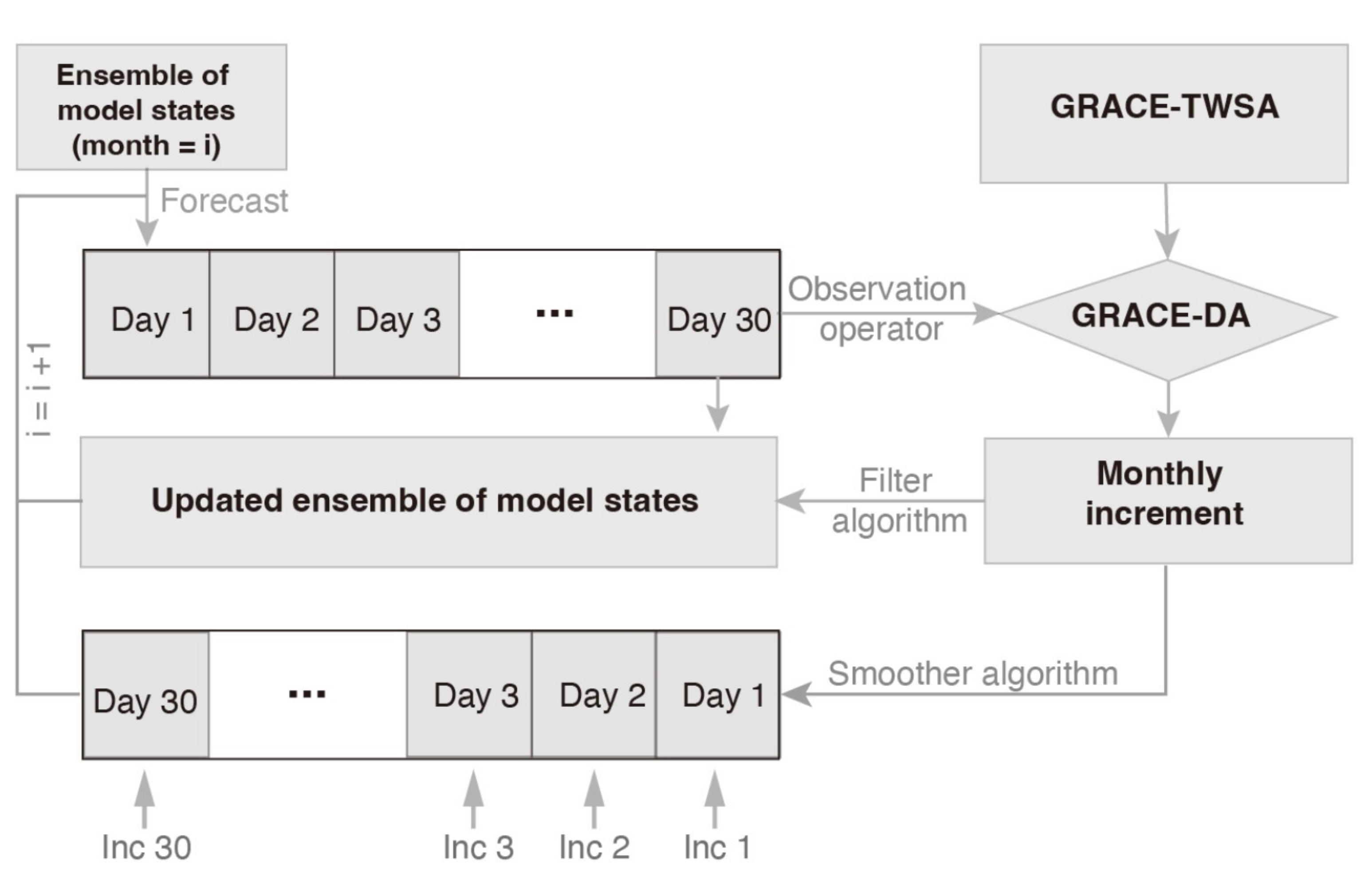
advanced perturbation of soil properties, case 2

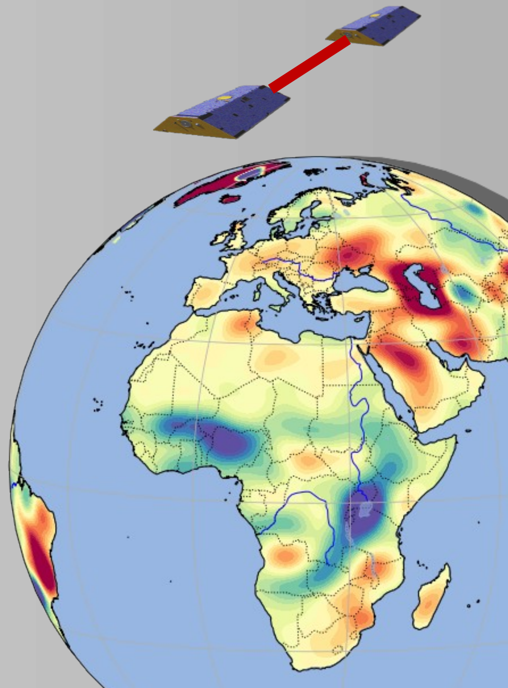


Ensure consistency of the updated model states

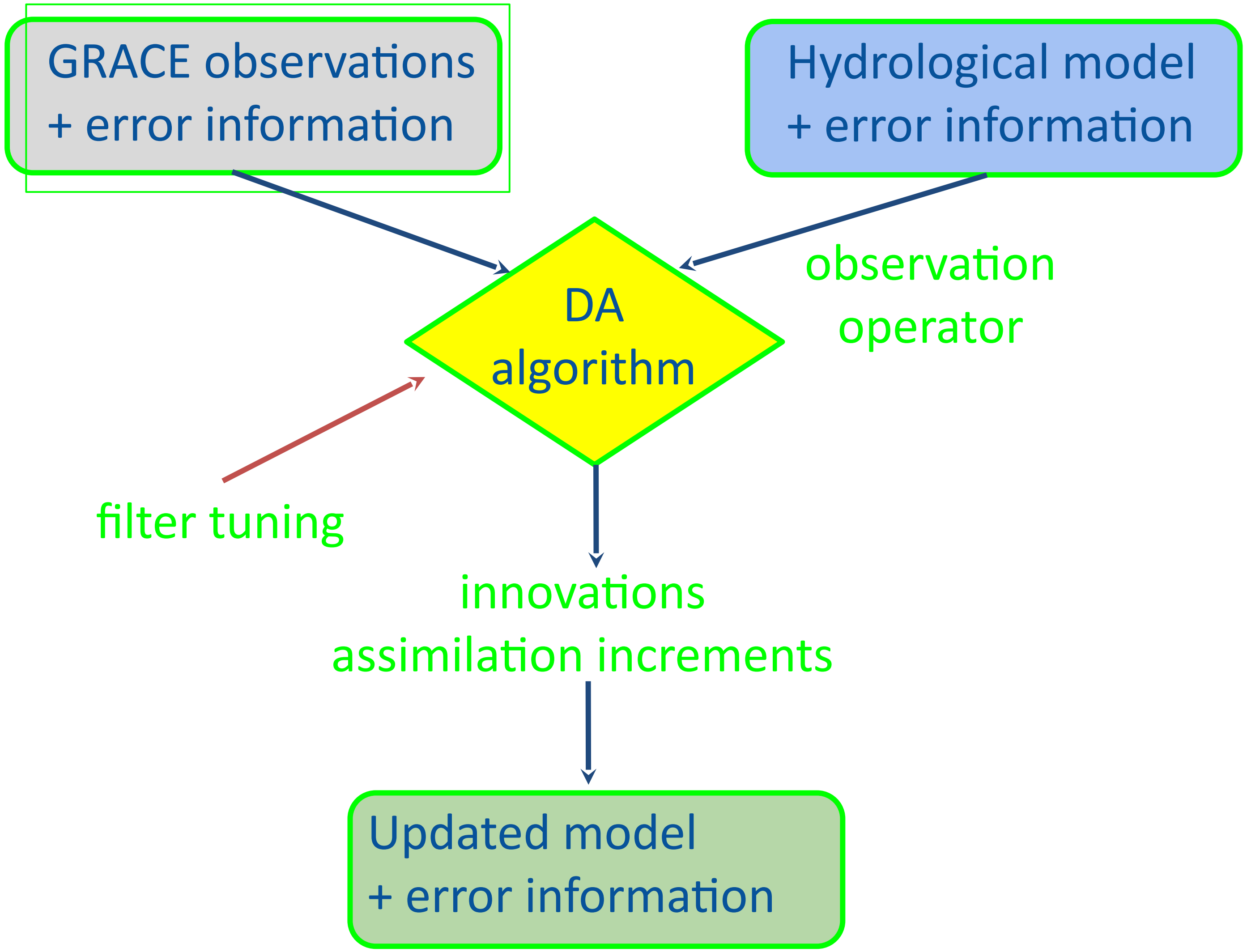
- Update depending variables, e.g. snow depth.
- Apply constraints regarding maximum and minimum values.
- Apply constraints regarding maximum increments.

Recap GRACE/-FO DA frameworks



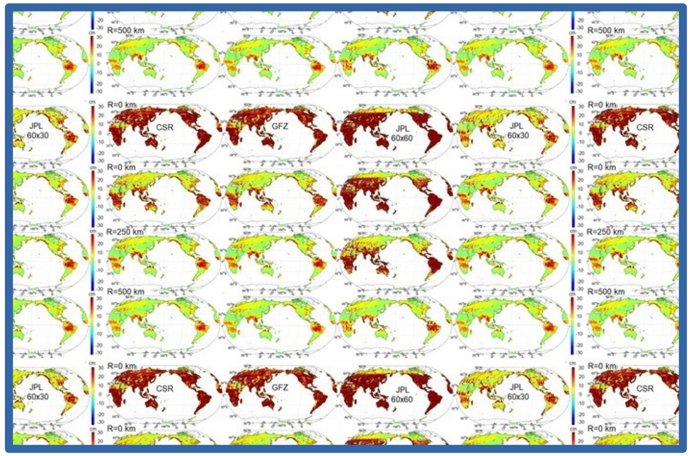


Set up of a data assimilation framework



GRACE DA choices

GRACE product and observation error

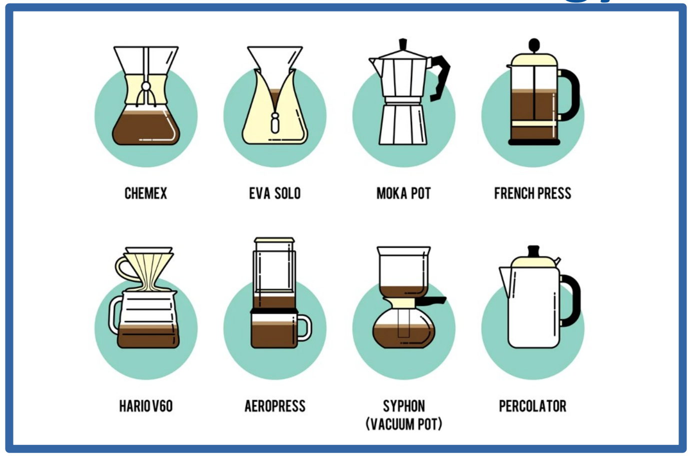


- Spherical harmonics (SH)
- Mascons
- Gridded level 3 product
- Line-of-sight gravity difference (LGD)

Geophysical corrections

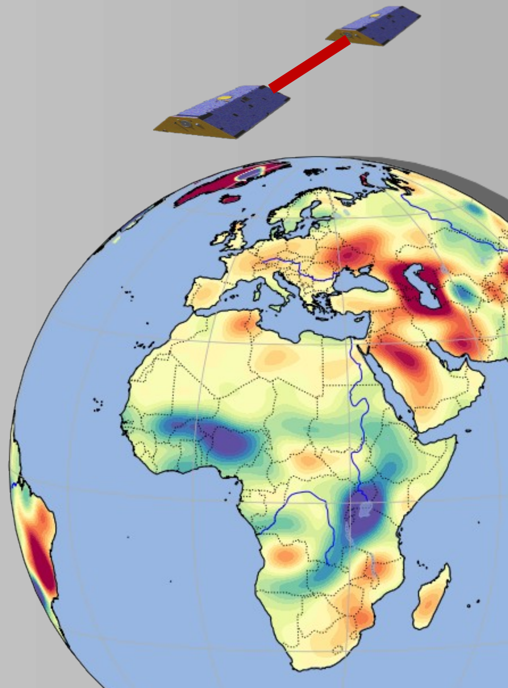
- Glacial isostatic adjustment (GIA)
- Lakes / reservoirs
- Earthquakes
- Glaciers

Assimilation strategy

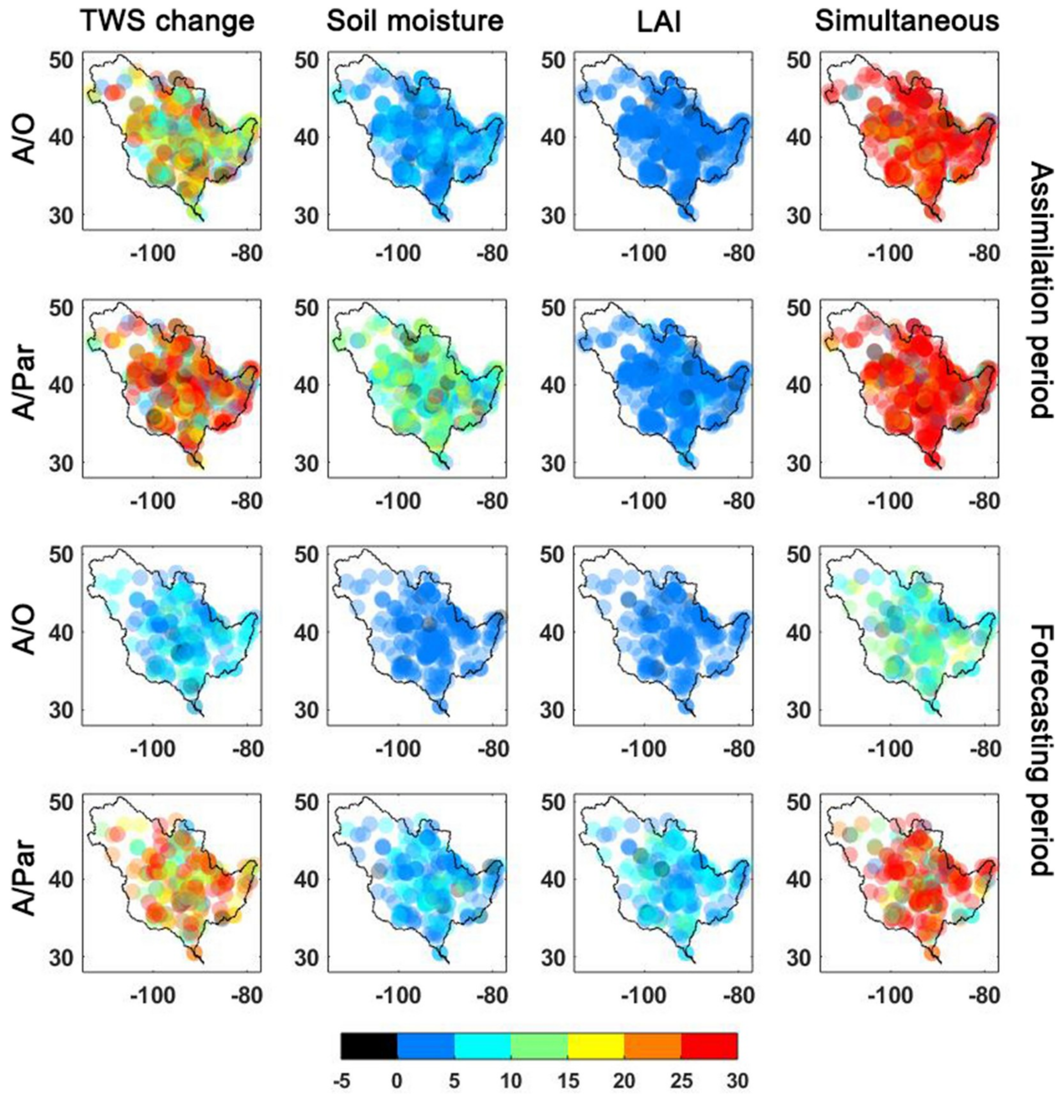


- Observation operator
- DA algorithm
- Application of increments

Research applications and service applications



Multi-mission DA frameworks



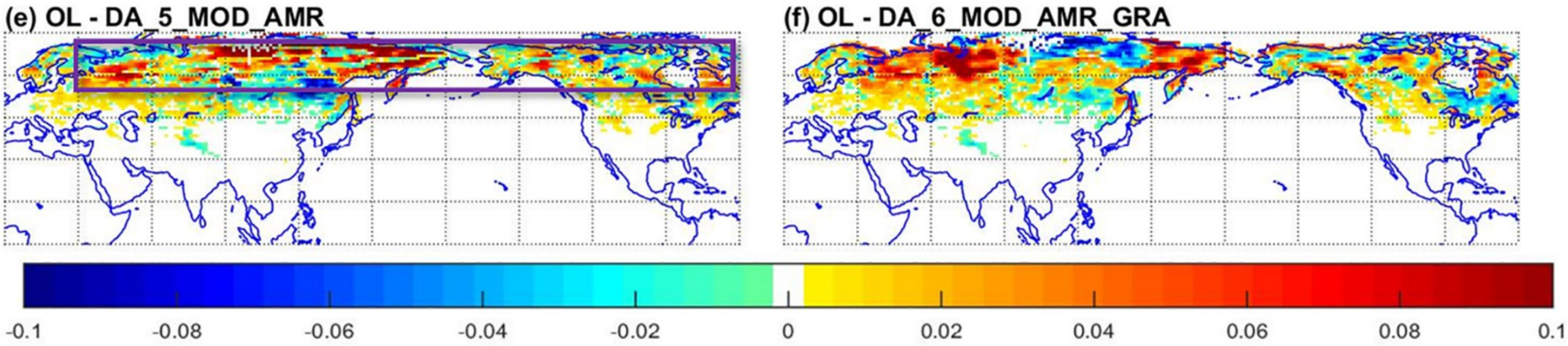
Assimilation of individual variables and simultaneous DA in the Mississippi river basin

- RMSD reduction with respect to groundwater wells

Multi-mission DA helps to better constrain individual water storages and fluxes.

Khaki et al. (2020), **Multi-mission satellite remote sensing data for improving land hydrological models via data assimilation**, Scientific reports

Improved snow estimates



RMSD differences between OL and DA of estimated snow depth as compared to the Canadian Meteorological Centre Daily Snow Depth Analysis Data

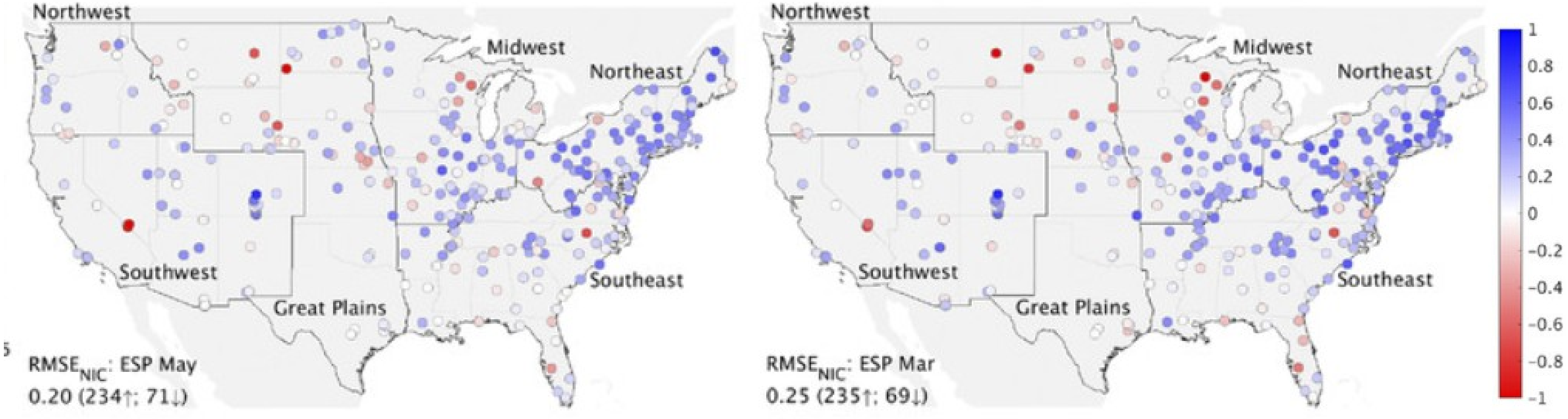
- warm colors indicated improvements due to DA

Zhao et al. (2018), **Multi-sensor land data assimilation: Toward a robust global soil moisture and snow estimation**, Remote Sensing of Environment

GRACE DA helps to better estimate seasonal snow changes at high temporal and spatial resolution.

- Significantly improves the skill of numerical weather prediction
- Contributes to water resources management

Groundwater forecast



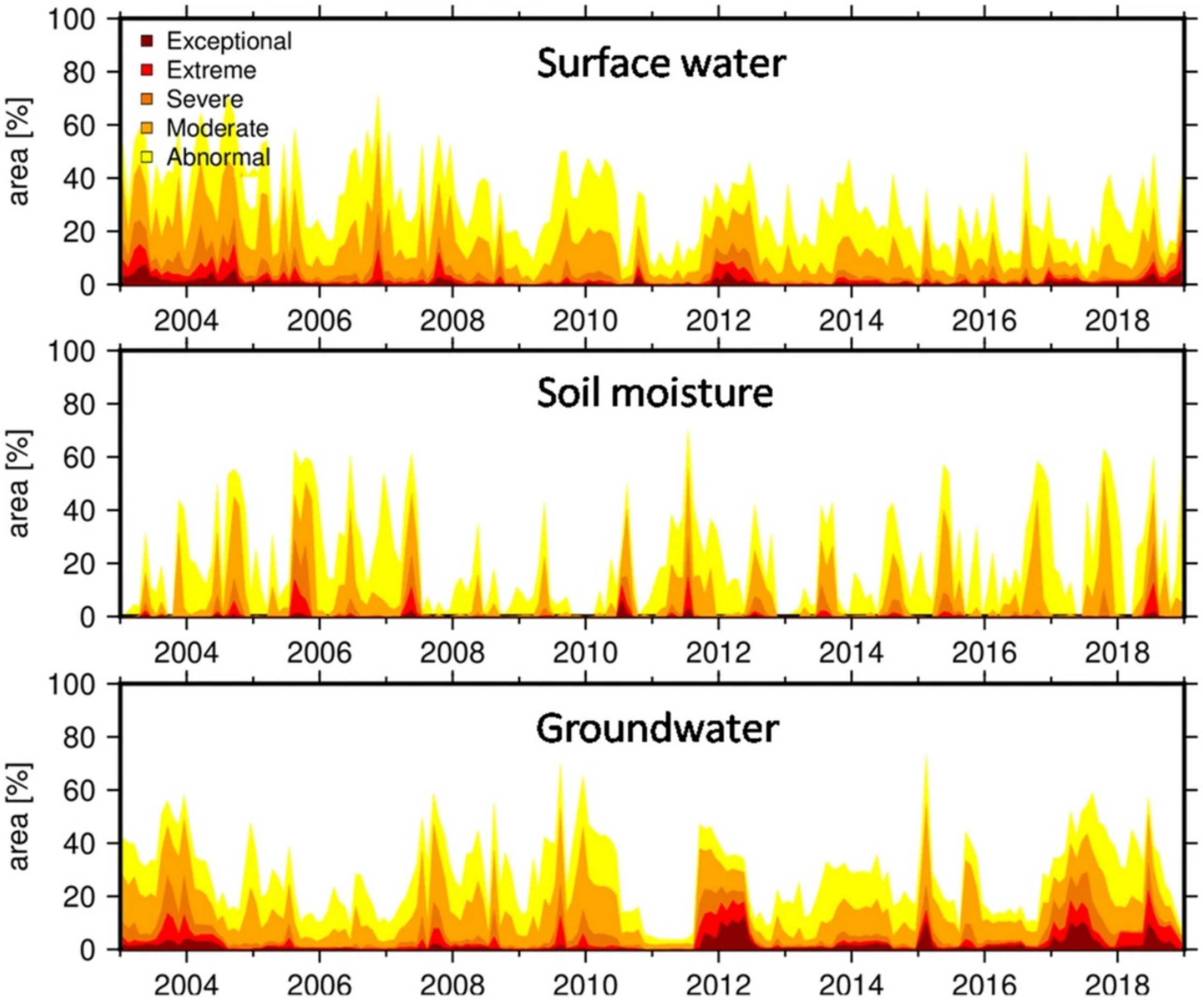
Getirana et al. (2020), **GRACE Improves Seasonal Groundwater Forecast Initialization over the United States**, J. Hydrometeor.

GRACE DA impact on RMSD of seasonal groundwater forecast (three month hindcasts) with respect to groundwater well observations.

GRACE DA improves groundwater estimates and prediction.

- Valuable tool for early warning systems and water resources management.

Improved representation of extreme events



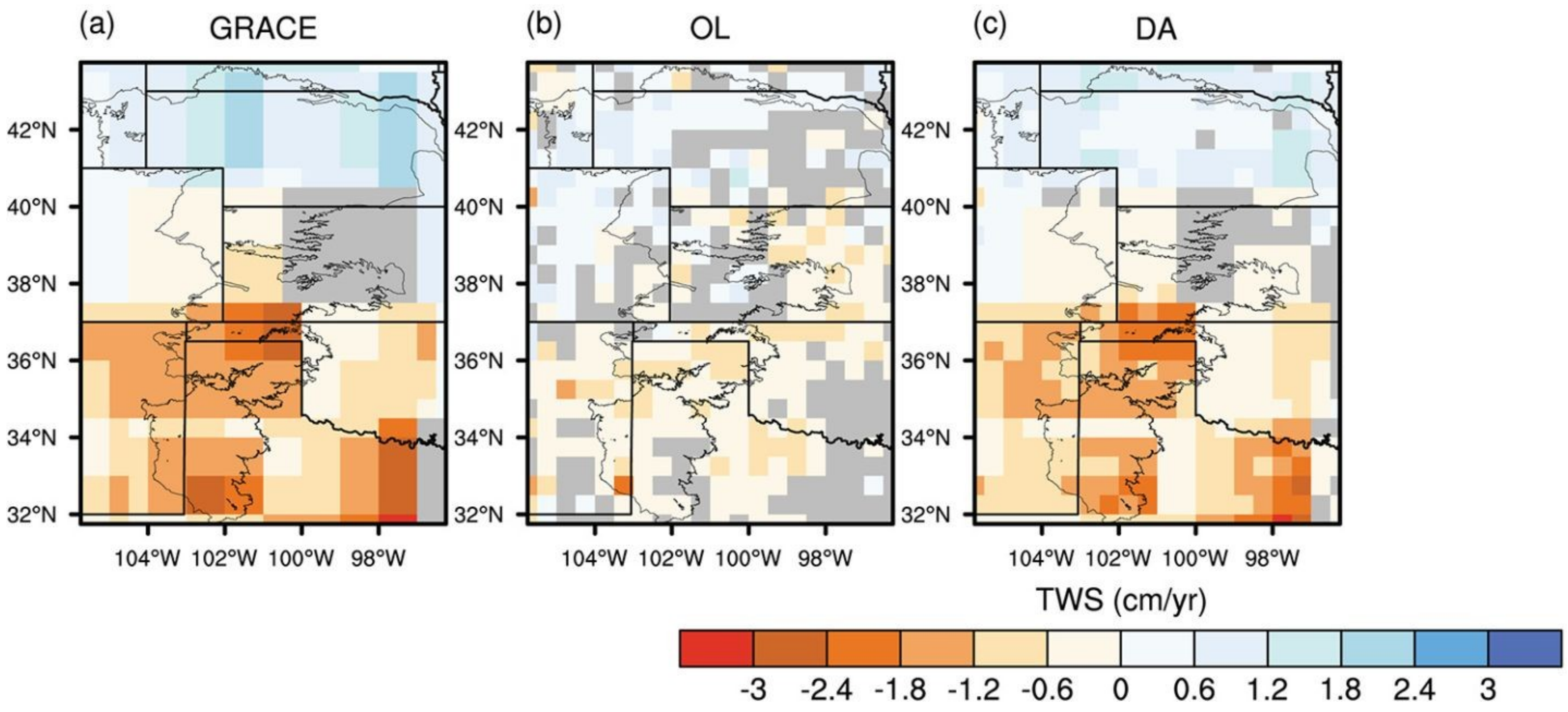
Dibi-Anoh et al. (2022), **Hydrometeorological Extreme Events in West Africa: Droughts**, Surveys in Geophysics

Area of the Niger basin (%) that is affected by drought computed based on a GRACE-assimilating hydrological model

GRACE DA helps to understand the evolution of droughts.

- Also employed in operational drought monitors.

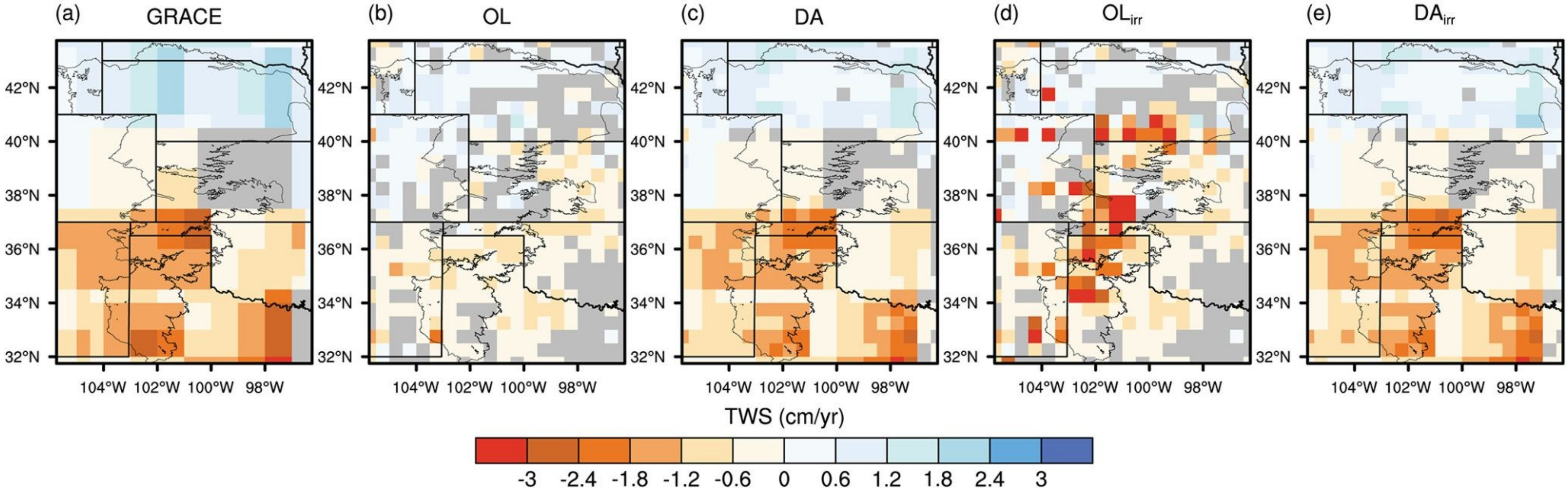
Human impact analysis



Nie et al. (2019), **Assimilating GRACE Into a Land Surface Model in the Presence of an Irrigation-Induced Groundwater Trend**, Water Resources Research.

GRACE DA can help to identify shortcomings of current models. However, neglected human impacts are not necessarily corrected in a meaningful way.

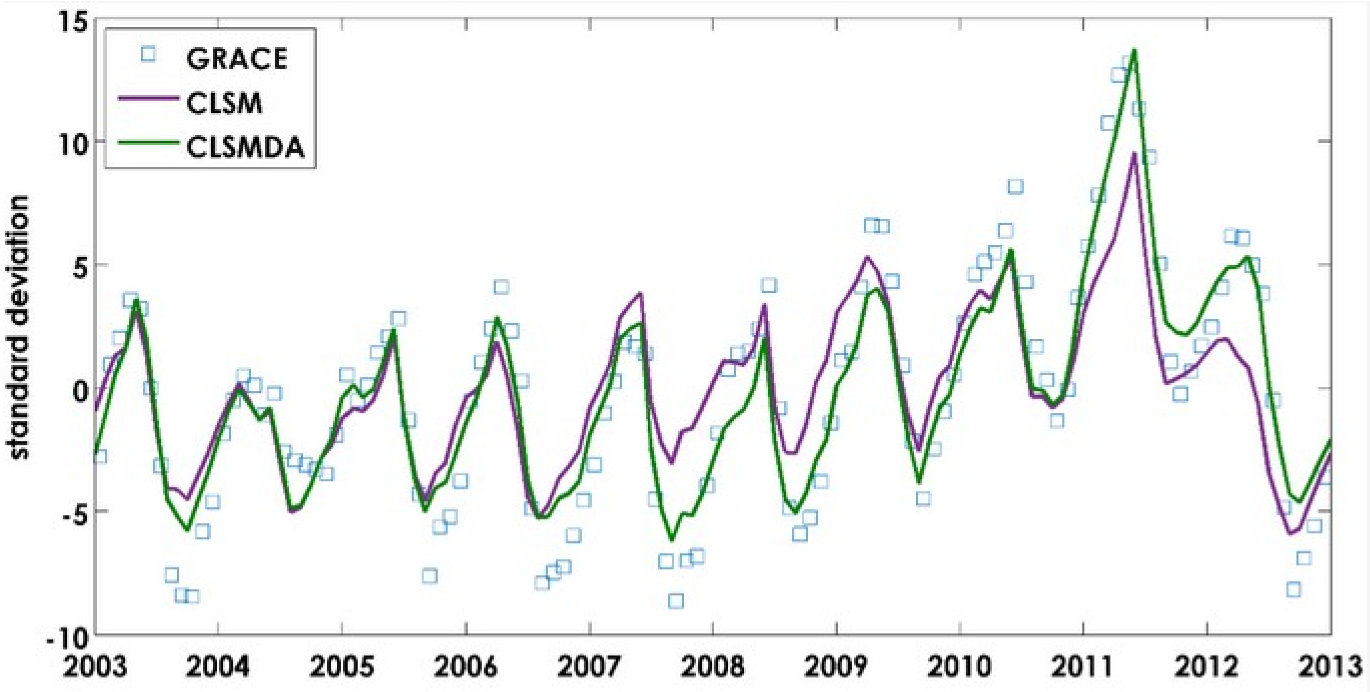
Human impact analysis

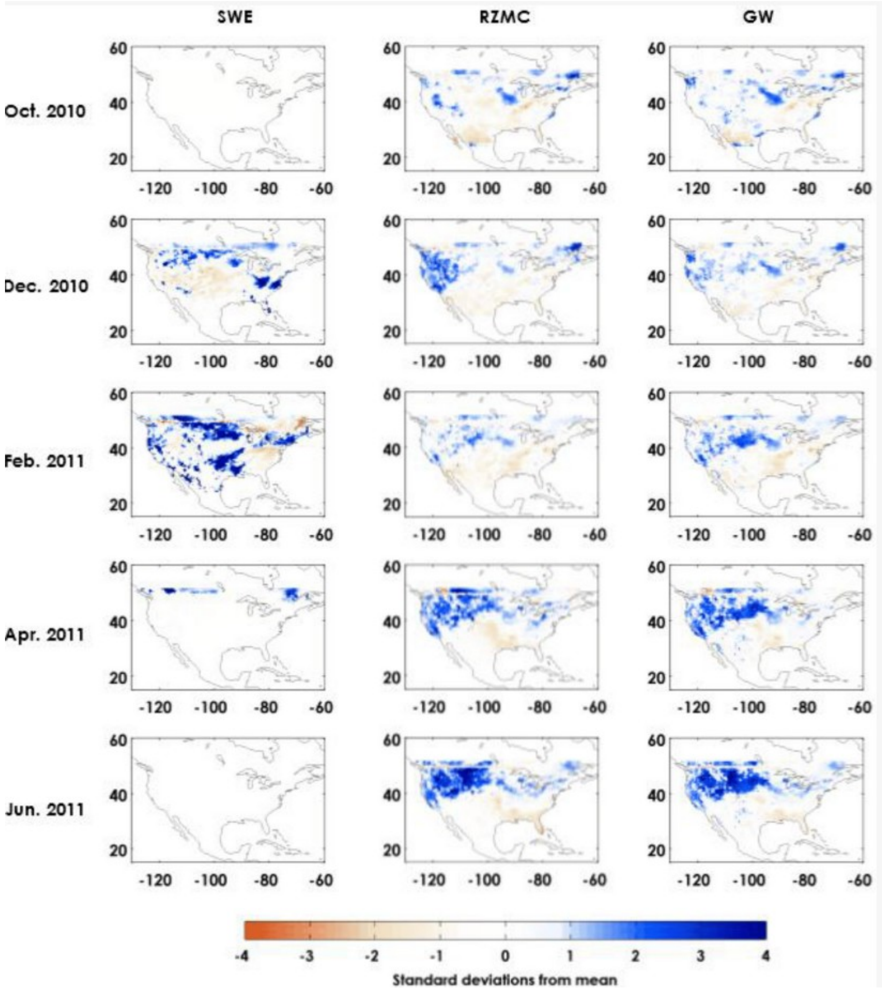


Nie et al. (2019), **Assimilating GRACE Into a Land Surface Model in the Presence of an Irrigation-Induced Groundwater Trend**, Water Resources Research.

GRACE DA can help to identify shortcomings of current models. However, neglected human impacts are not necessarily corrected in a meaningful way.

Assessment of regional flood potential

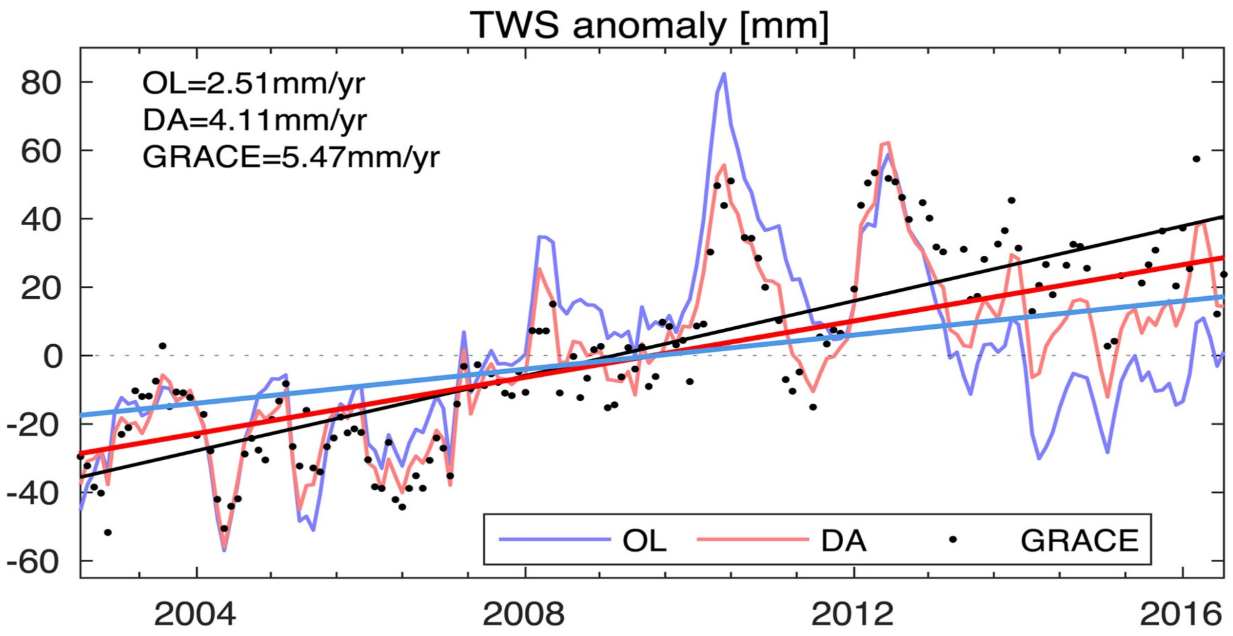




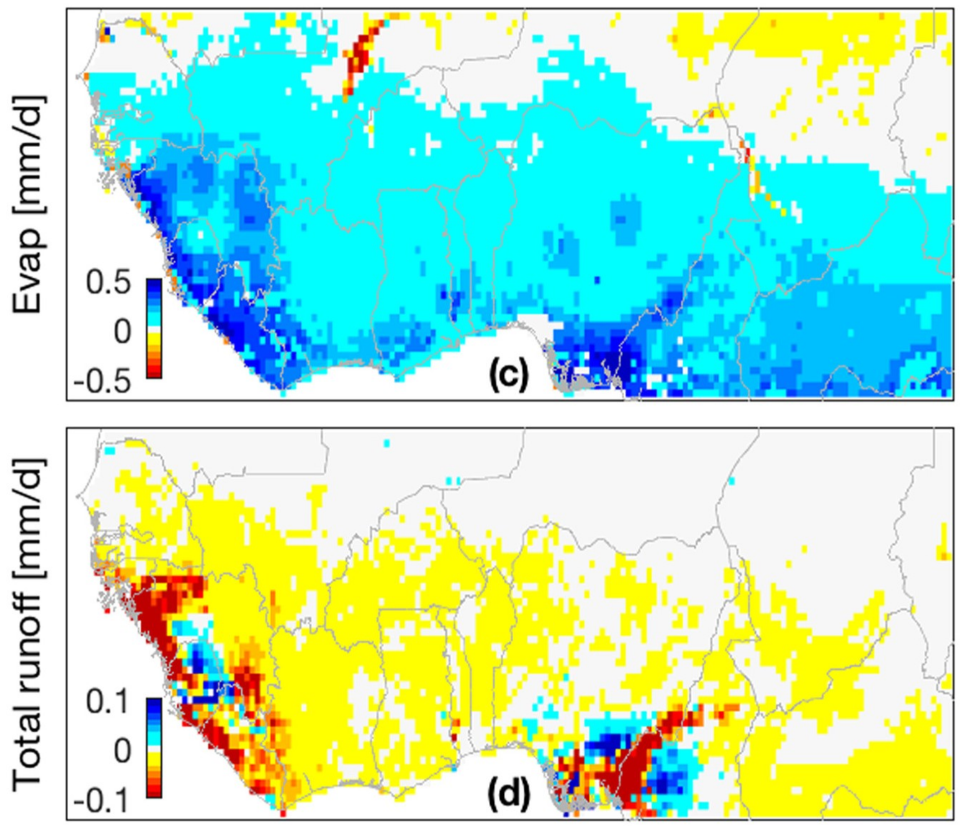
Reager et al. (2019), **Assimilation of GRACE Terrestrial Water Storage Observations into a Land Surface Model for the Assessment of Regional Flood Potential**, remote sensing.

GRACE DA improves the understanding of contributors to regional flood events and the predictability of such events.

Streamflow forecast



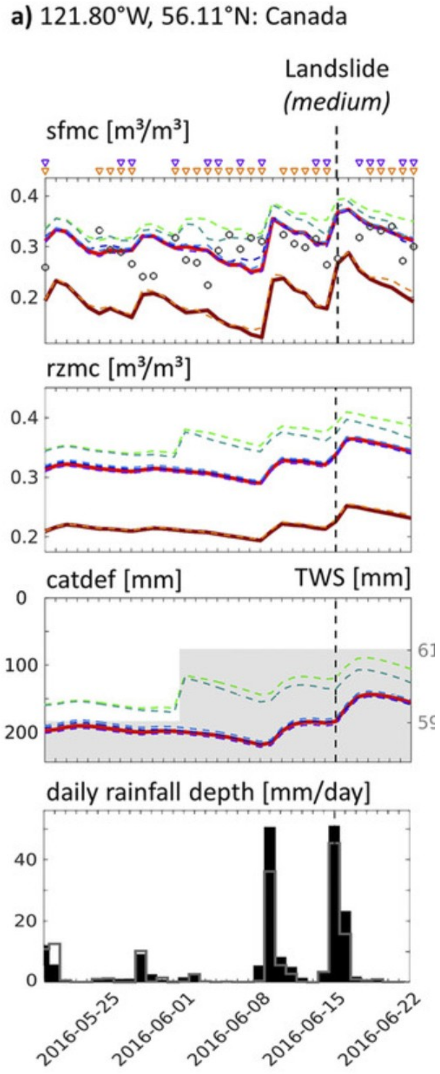
Getirana et al. (2020), **Satellite Gravimetry Improves Seasonal Streamflow Forecast Initialization in Africa**, Water Resources Research.



Differences DA-OL

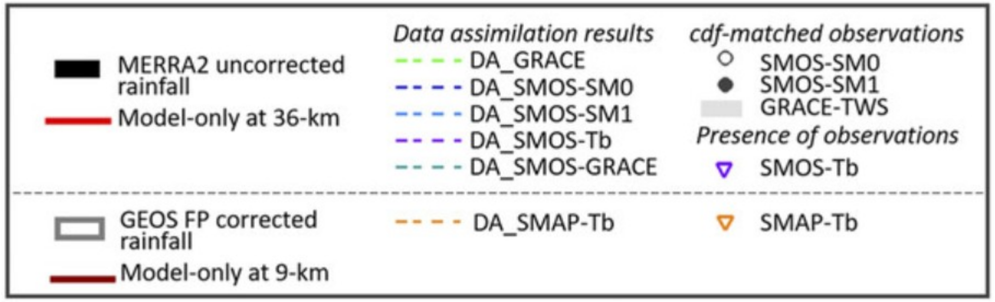
Streamflow forecasts are enhanced by the long memory of groundwater and deep soil moisture, two main TWS components updated by GRACE-DA.

Landslide prediction



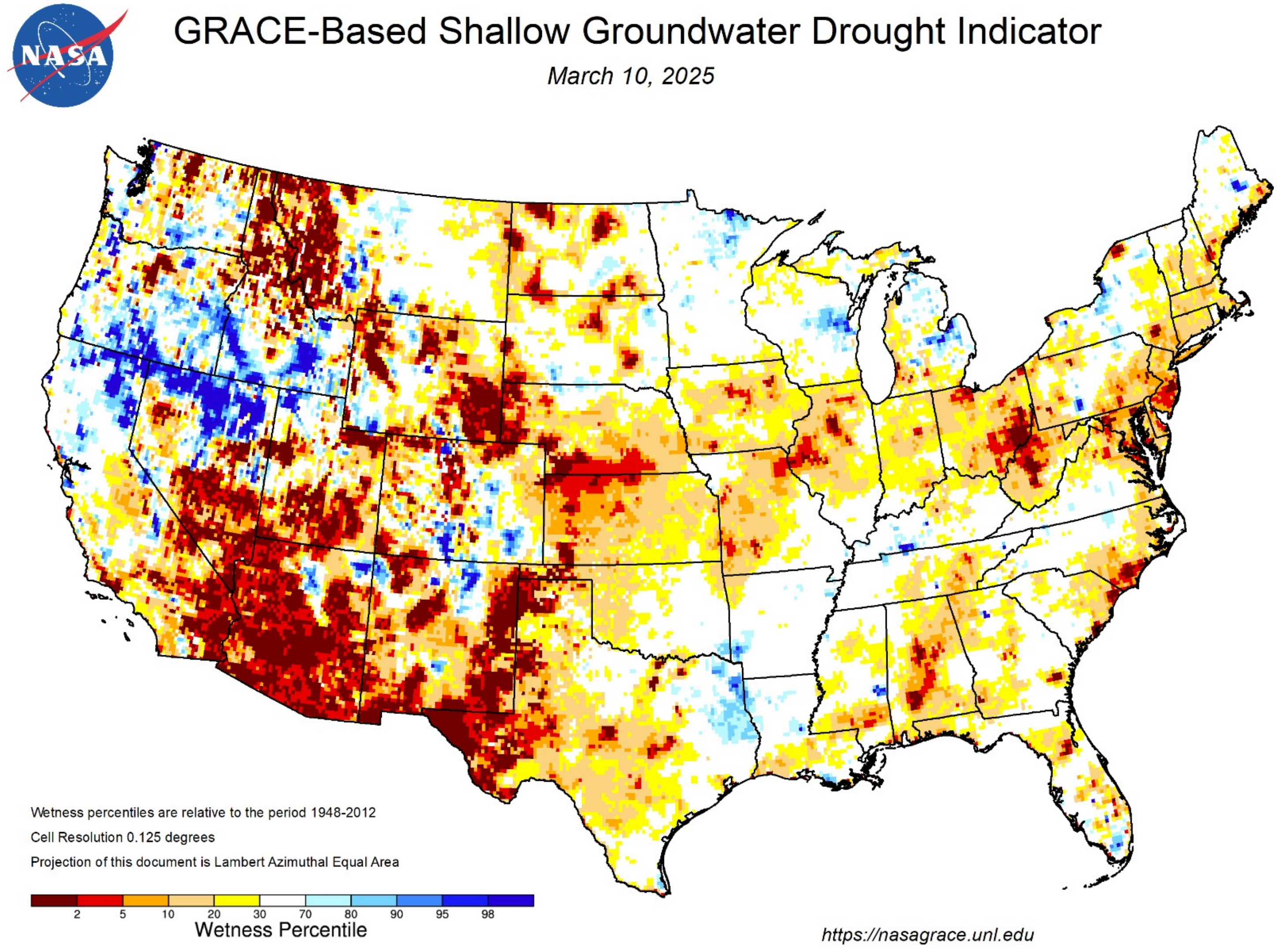
Felsberg et al. (2021), **Global Soil Water Estimates as Landslide Predictor: The Effectiveness of SMOS, SMAP, and GRACE Observations, Land Surface Simulations, and Data Assimilation**, J. Hydrometeor.

sfmc … surface soil moisture content
rzmc … root zone soil moisture content
catdef … catchment deficit

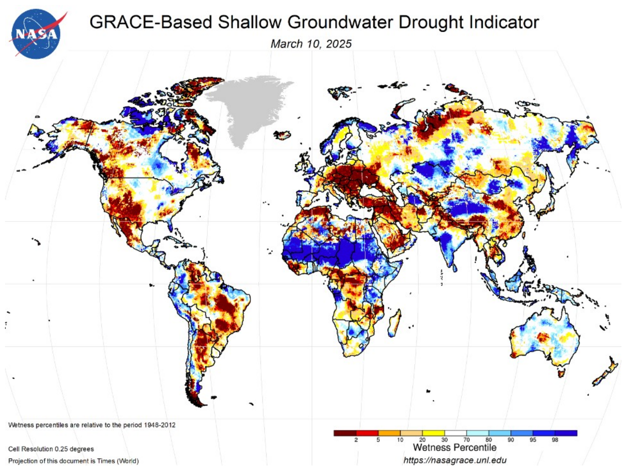


Global landslide modeling can benefit from GRACE data assimilation under certain conditions.

NASA's GRACE-Based Drought Indicators

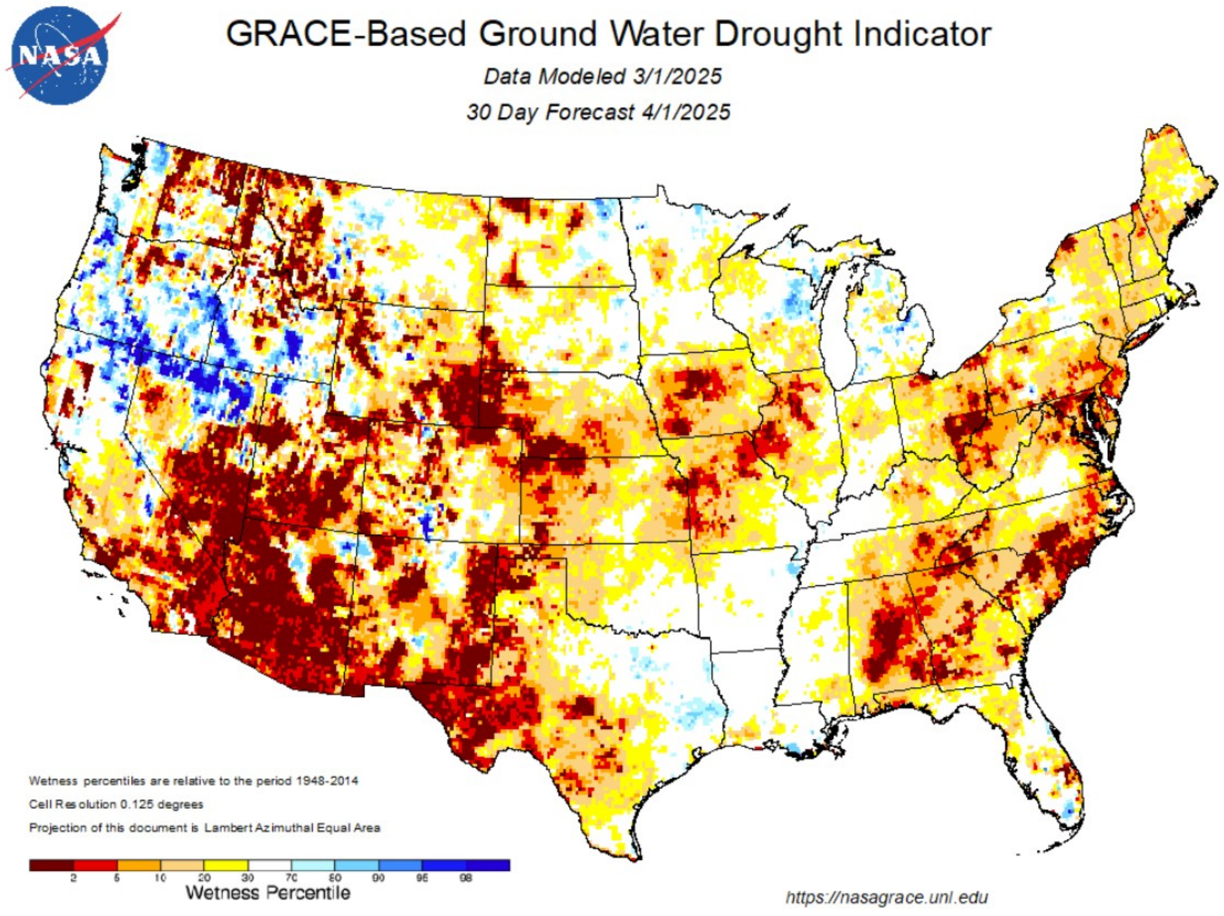


https://nasagrace.unl.edu/



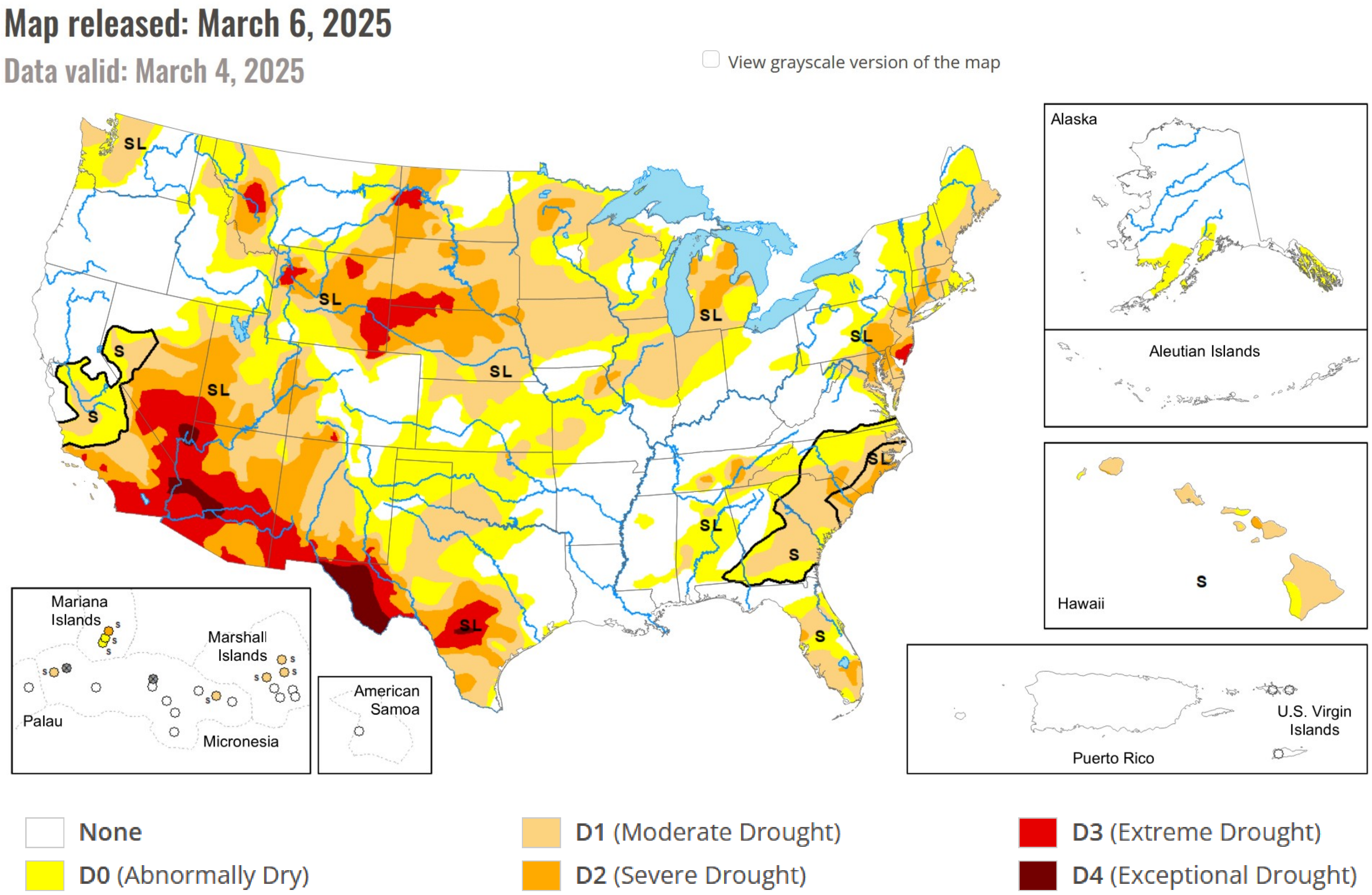
NASA's GRACE-Based Drought Indicators

Groundwater and Soil Moisture Forecasts Initialized from GRACE-FO Data Assimilation



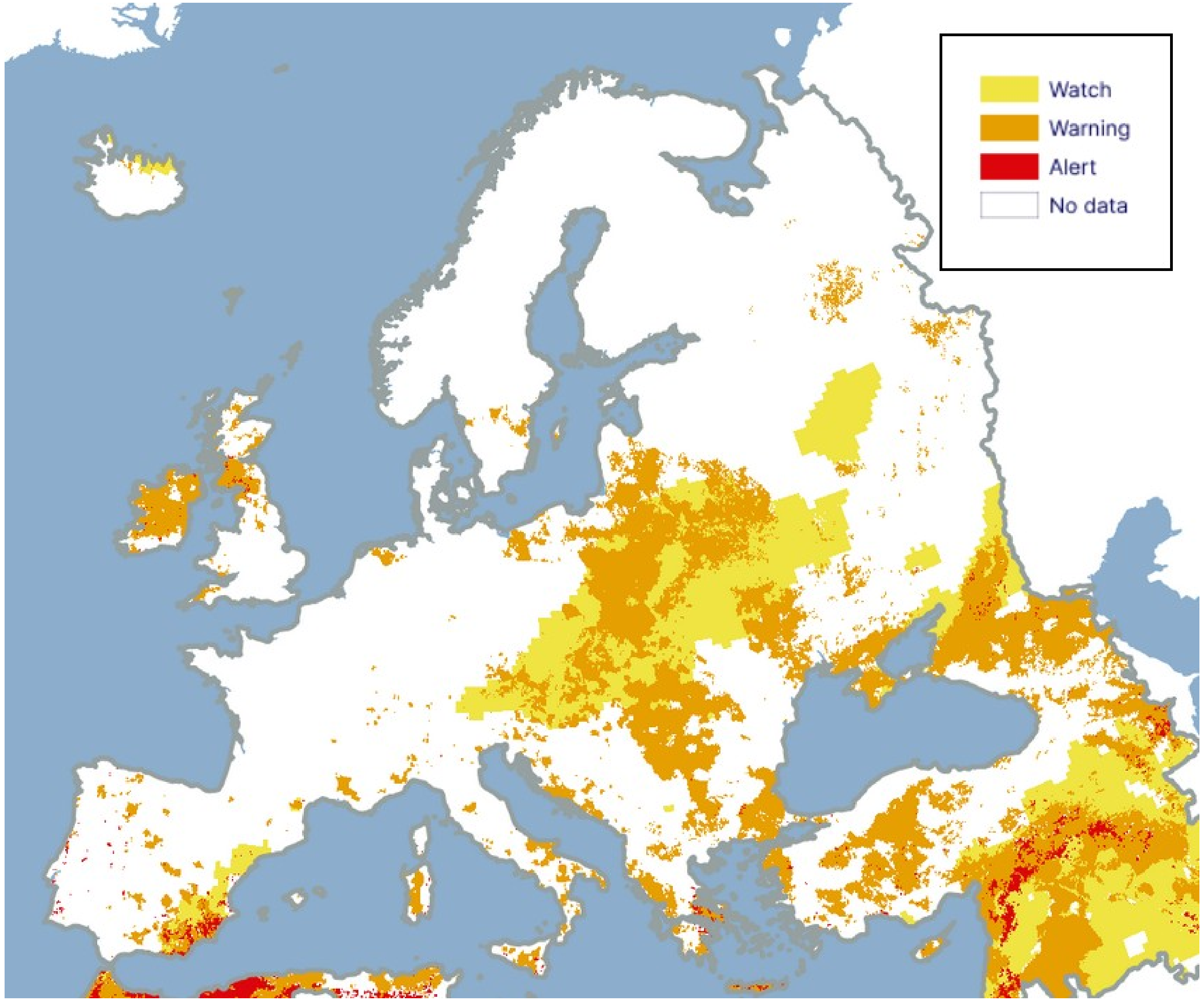
https://nasagrace.unl.edu/

U.S. Drought monitor (USDM)



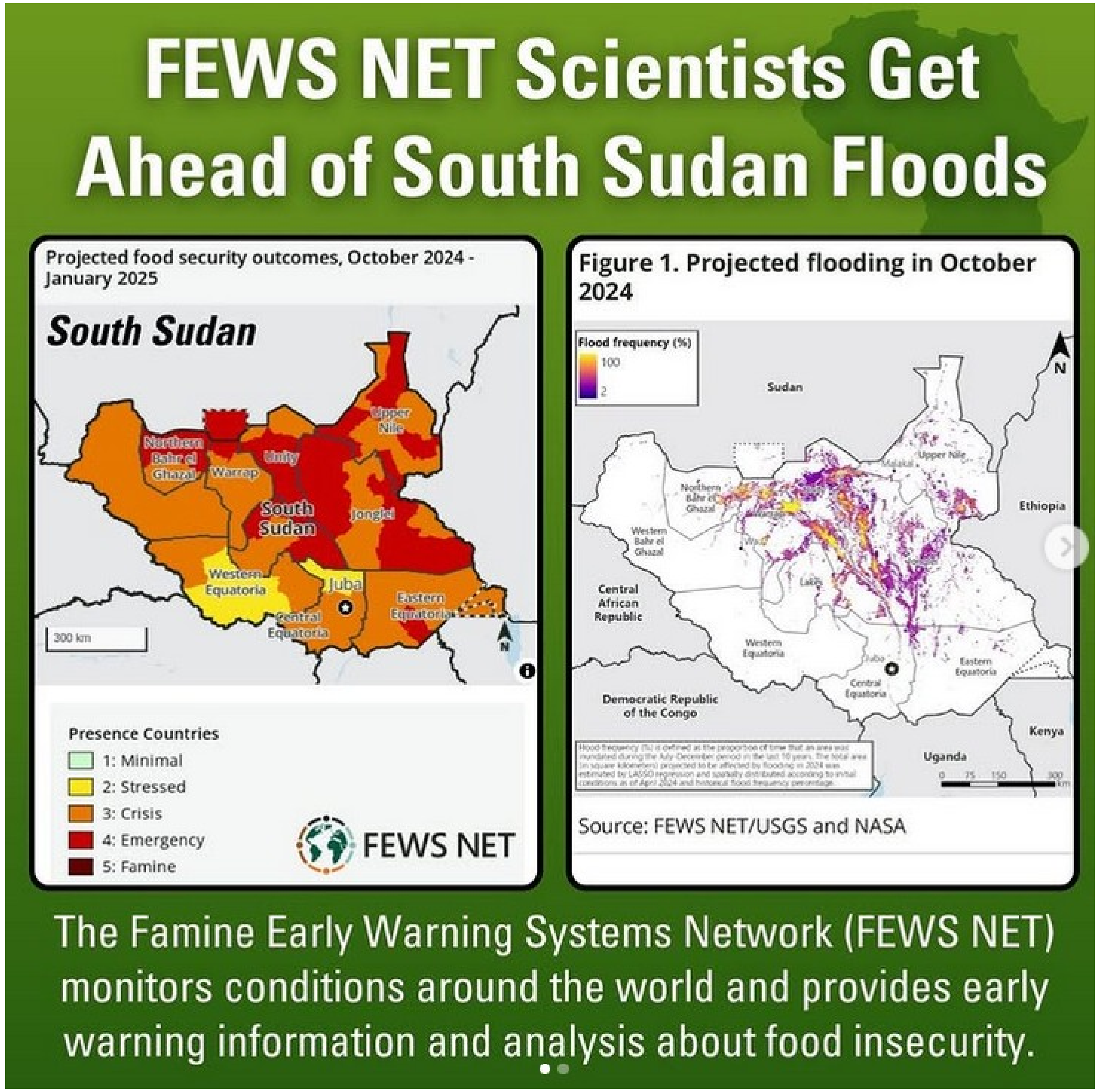
https://droughtmonitor.unl.edu/

European Drought Observatory (EDO)



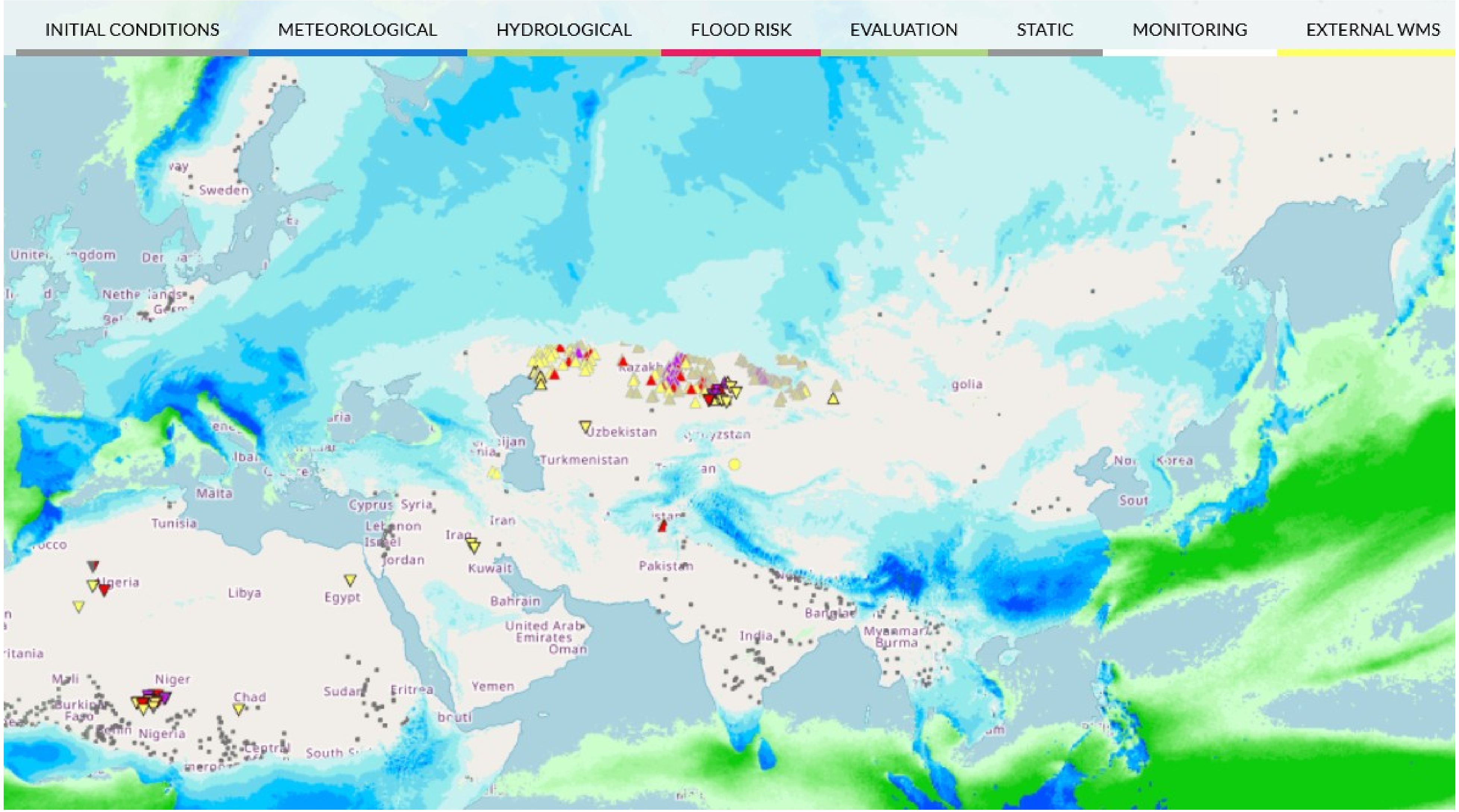
https://drought.emergency.copernicus.eu/

Famine Early Warning Systems Network (FEWS NET)



Global Flood Awareness System

https://global-flood.emergency.copernicus.eu/glofas-forecasting/



GRACE DA not directly included :(

WMO Hydrological Services

Some projects explore use of GRACE in data assimilation frameworks for hydrological services, such as:

- agriculture
- disaster risk reduction
- energy
- industry
- environment

Future directions

- Working towards a standard procedure for assimilating GRACE data.
- Exploring alternative GRACE products such as line-of-sight gravity differences for near real time applications.
- Assimilating GRACE data into coupled Earth system models.
- Upcoming new gravity missions will lead to substantial improvements in the spatial and temporal resolution of GRACE TWSA and be a great asset for assimilation frameworks.