Data assimilation in hydrological sciences

Fabio.Castelli@unifi.it

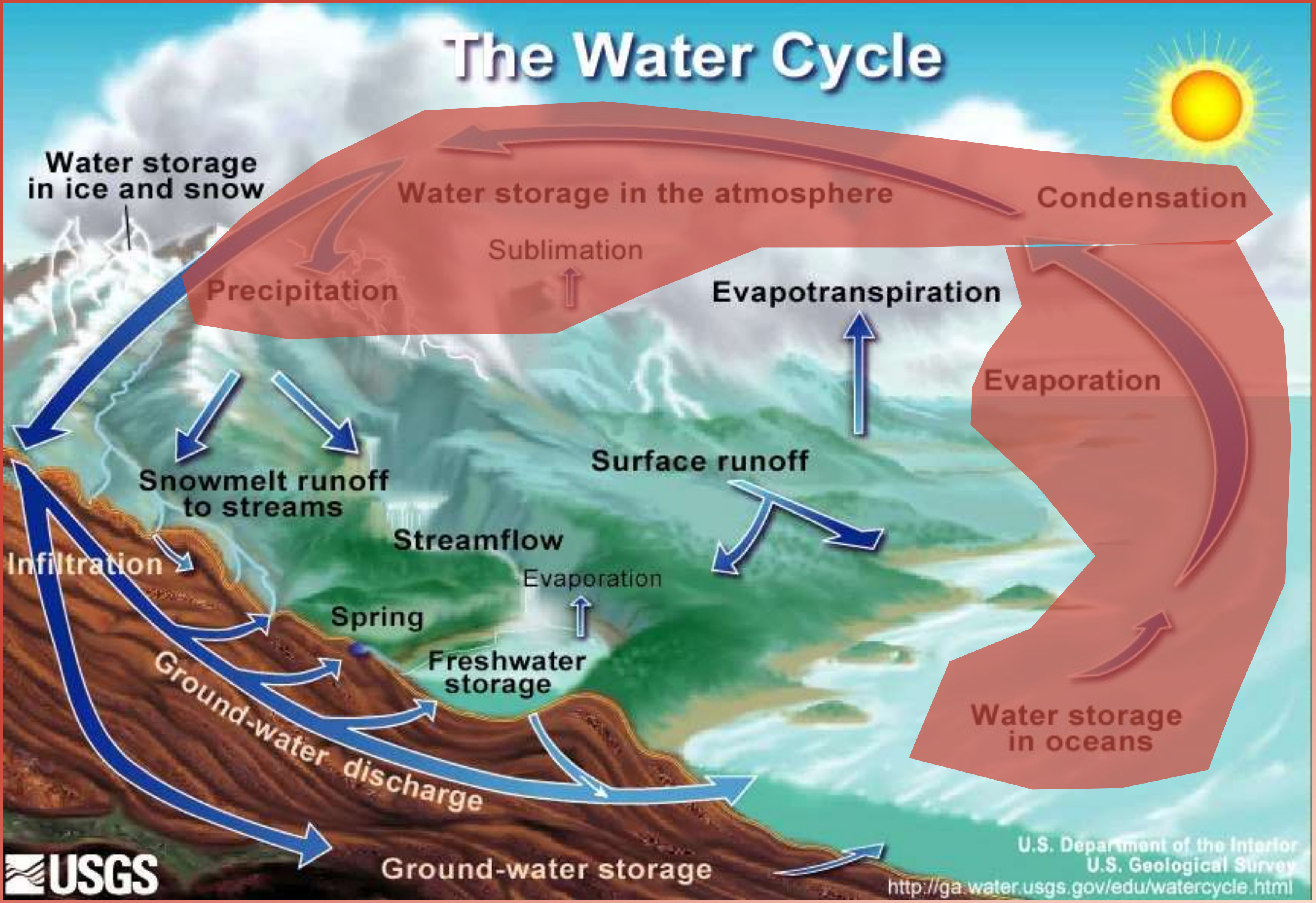
Outline

- Introduction: what peculiarities for D.A. in hydrology?

1. Improving remote sensing of land surface: dynamic (Kalman-based) filtering
2. State estimation with dynamic (Kalman-based) filtering and sparse observations
3. Geophysical inversion as a causal identification and quantification tool
4. Geophysical inversion and state estimation together

- Discussion

Hydrology: what peculiarities for D.A.?



Hydrology: what peculiarities for D.A.?

Some ‘classical’ applications of hydrological sciences

- Engineering design (water structures and infrastructures)
- Water resources management
- Natural hazards and extreme events (floods, droughts, …)
- Land-atmosphere interactions (lower flux boundary condition for weather and climate modelling)
- Water-related environmental quality

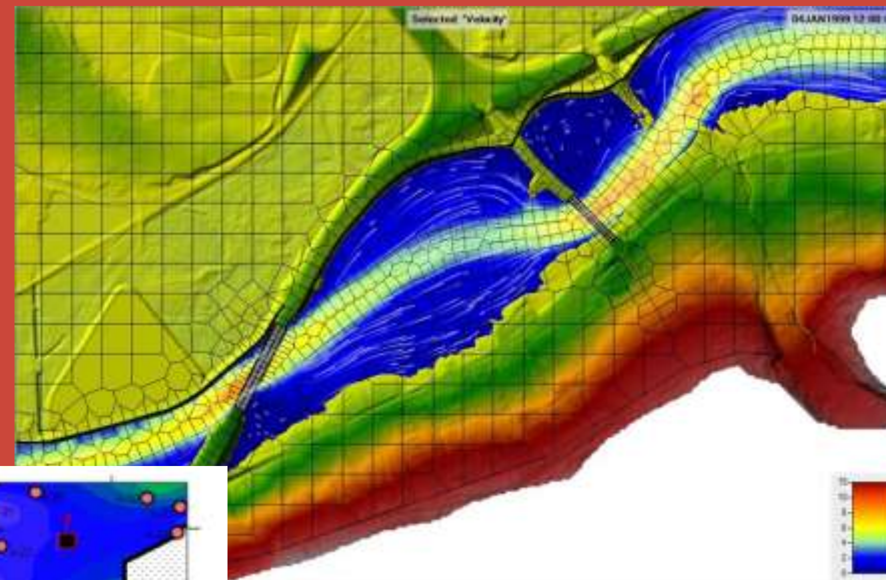
Cross-dciplinary themes

- Eco-hydrology
- Socio-hydrology
- Hydrogeochemistry
- Water security

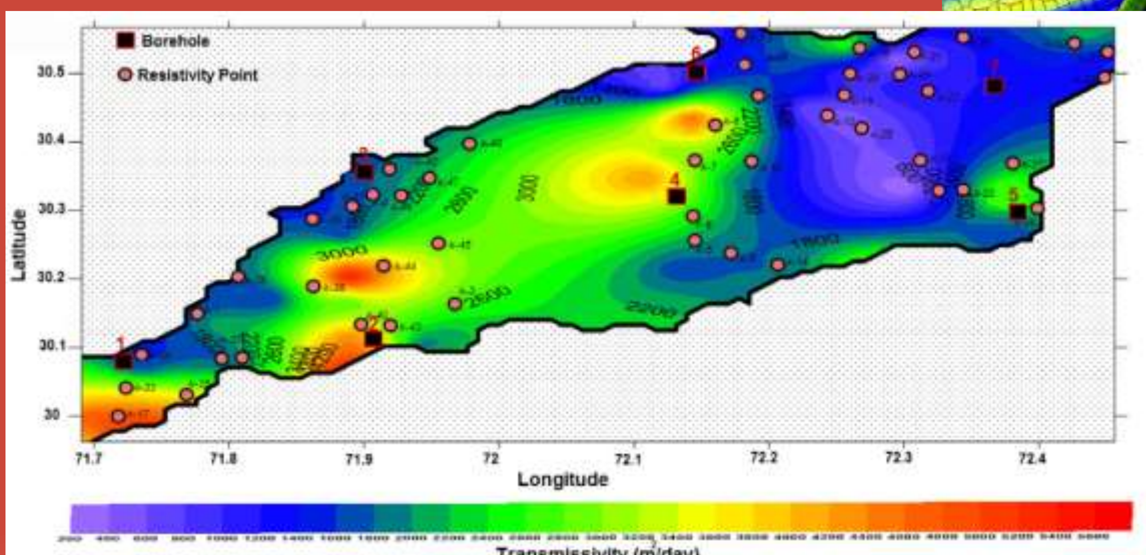
Climate change physics and impacts

Hydrology: what peculiarities for D.A.?

Most of the complexity resides not in the flow itself, but in the environment that contains/constraints the flow, and in the forcing.



HEC-RAS 6.0 River Flow Simulation System



Determination of aquifer parameters using geoelectrical sounding and pumping test data in Khanewal District, Pakistan

Gulraiz Akhter and M. Hasan

From the journal Open Geosciences

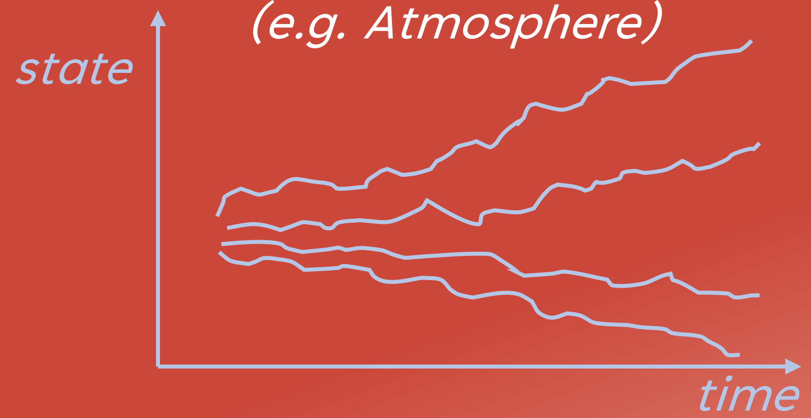
https://doi.org/10.1515/geo-2016-0071

Hydrology: what peculiarities for D.A.?

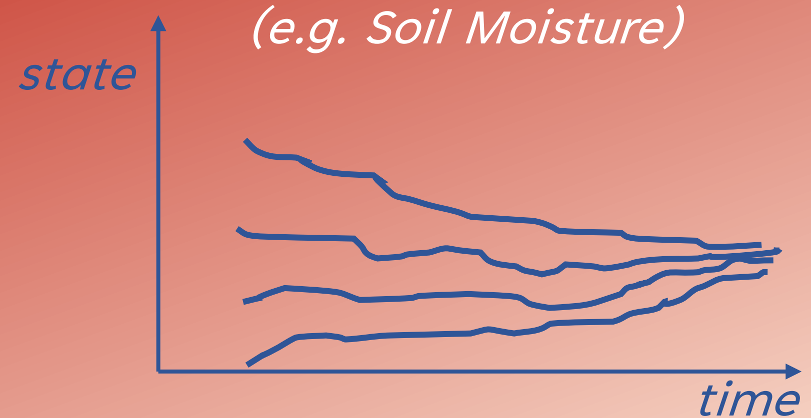
The most important governing equations in hydrology are <u>convergent</u> in nature (quick loss of memory of initial conditions)

If the model is ‘wrong’ (poor structure, bad parameters) it will rapidly converge back to the ‘wrong’ solution, after state update with D.A.

*A **divergent** system*
(e.g. Atmosphere)



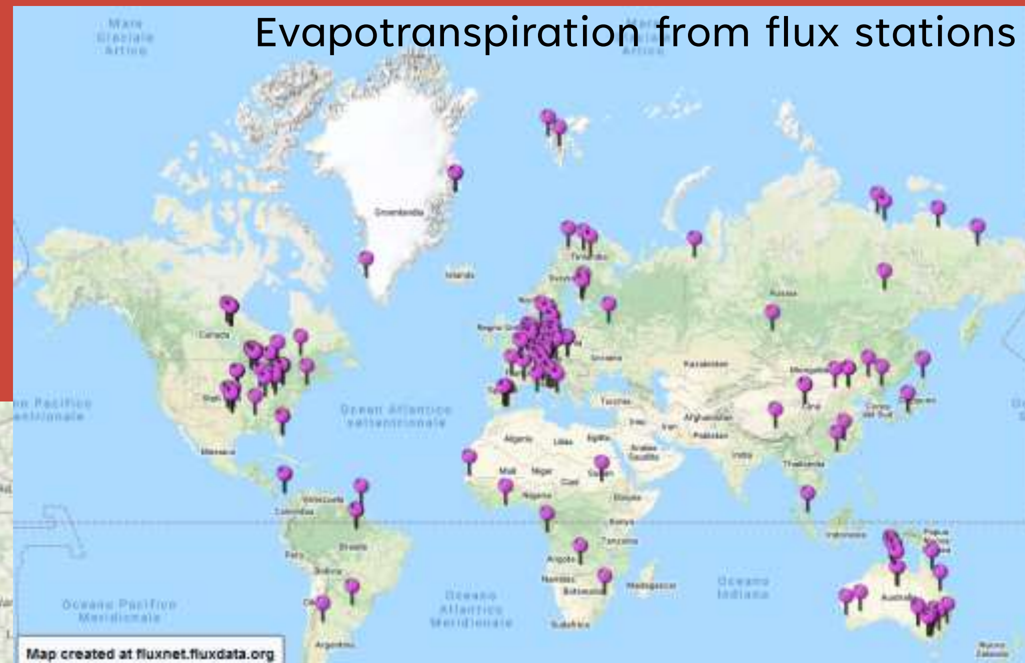
*A **convergent** system*
(e.g. Soil Moisture)



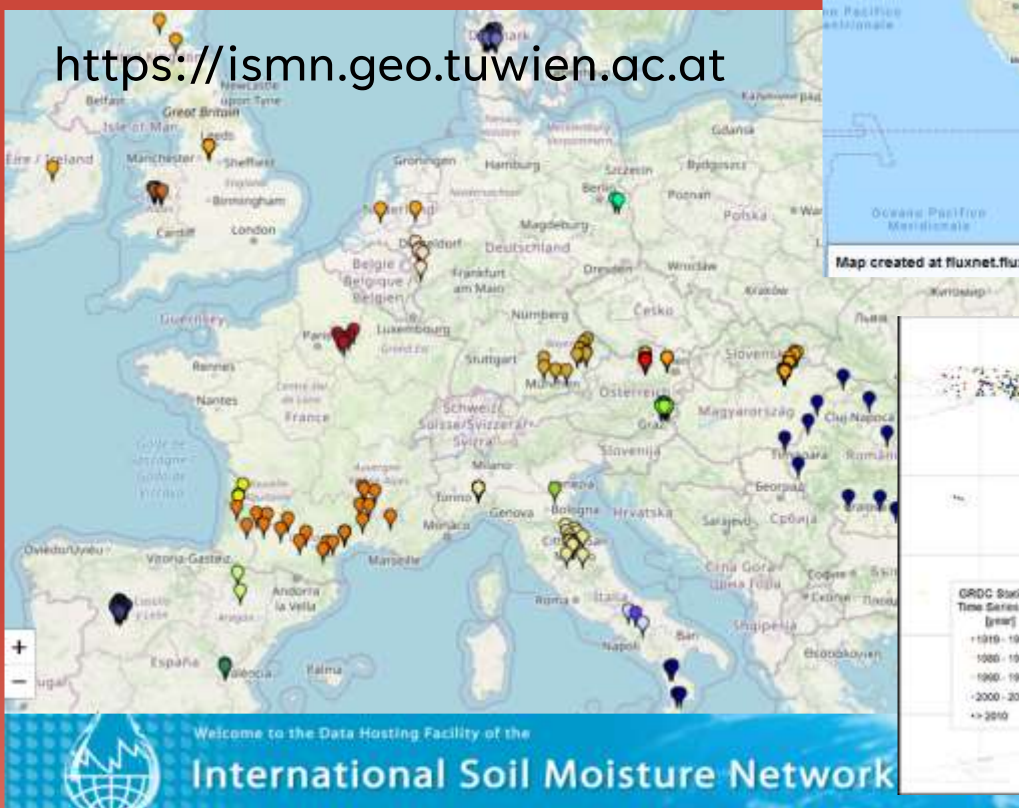
Hydrology: what peculiarities for D.A.?

Some key variables (e.g. soil moisture, evapotranspiration, river discharge, land surface temperature) are routinely measured at the ground with a very poor spatial coverage

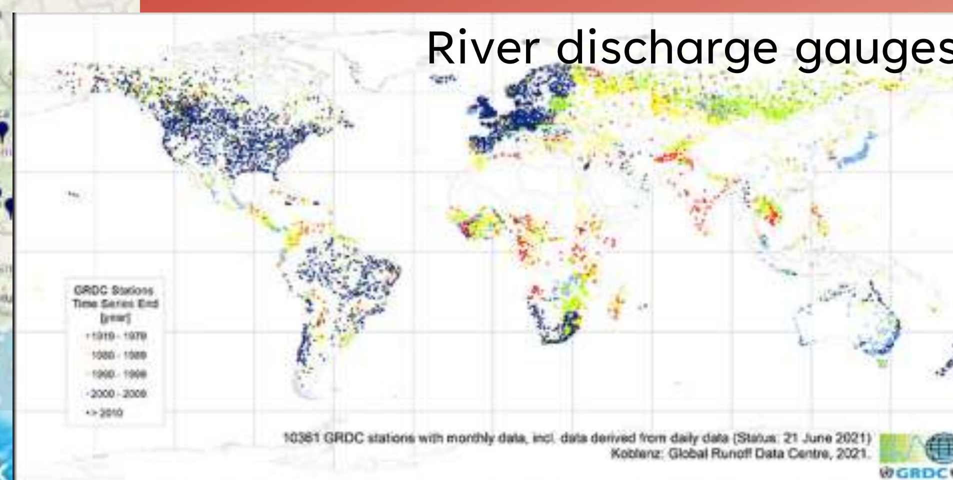
Evapotranspiration from flux stations



https://ismn.geo.tuwien.ac.at

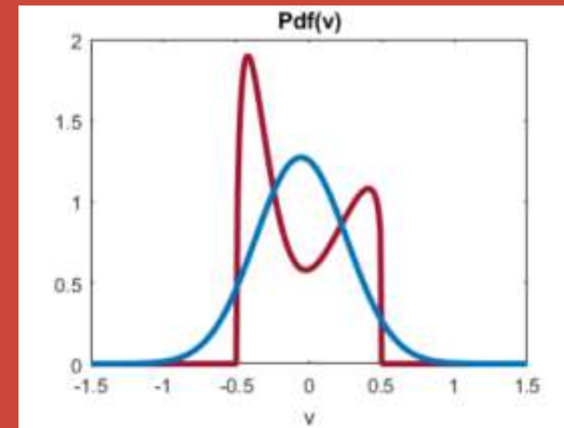


River discharge gauges

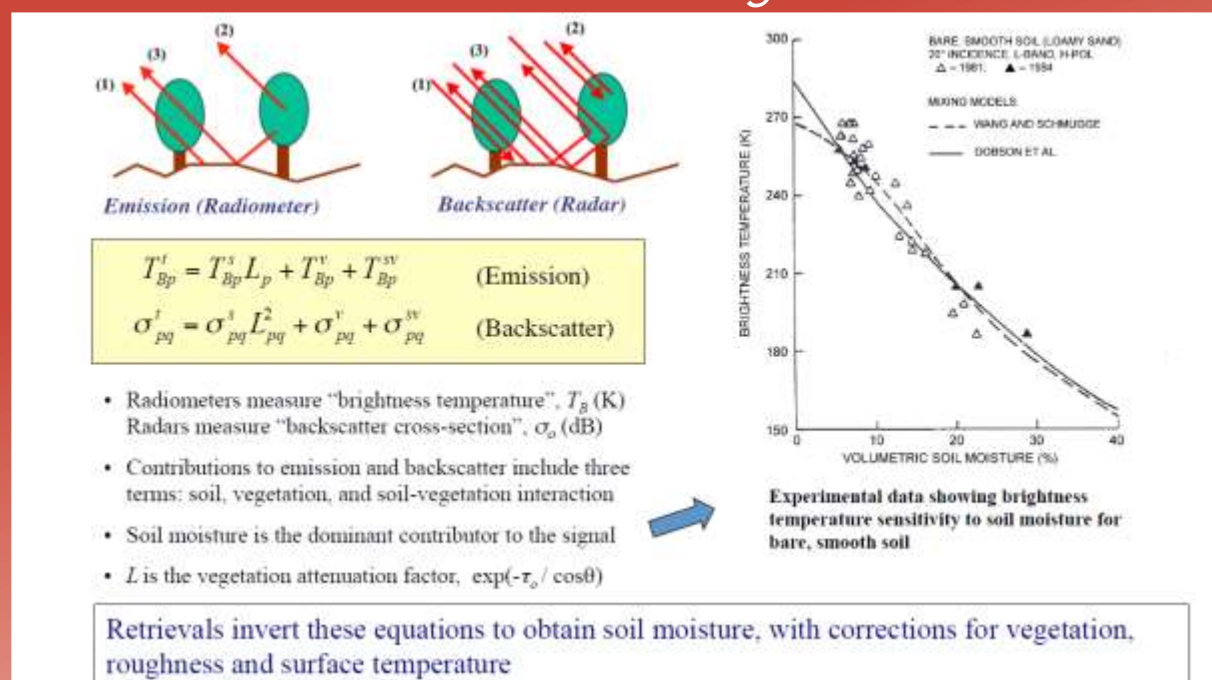


Hydrology: what peculiarities for D.A.?

New and long-flying satellite missions provide seamless estimates of those variables that are poorly gauged at the ground, but they are themselves the output from D.A. systems with <u>highly variable, non gaussian, errors structure</u>.



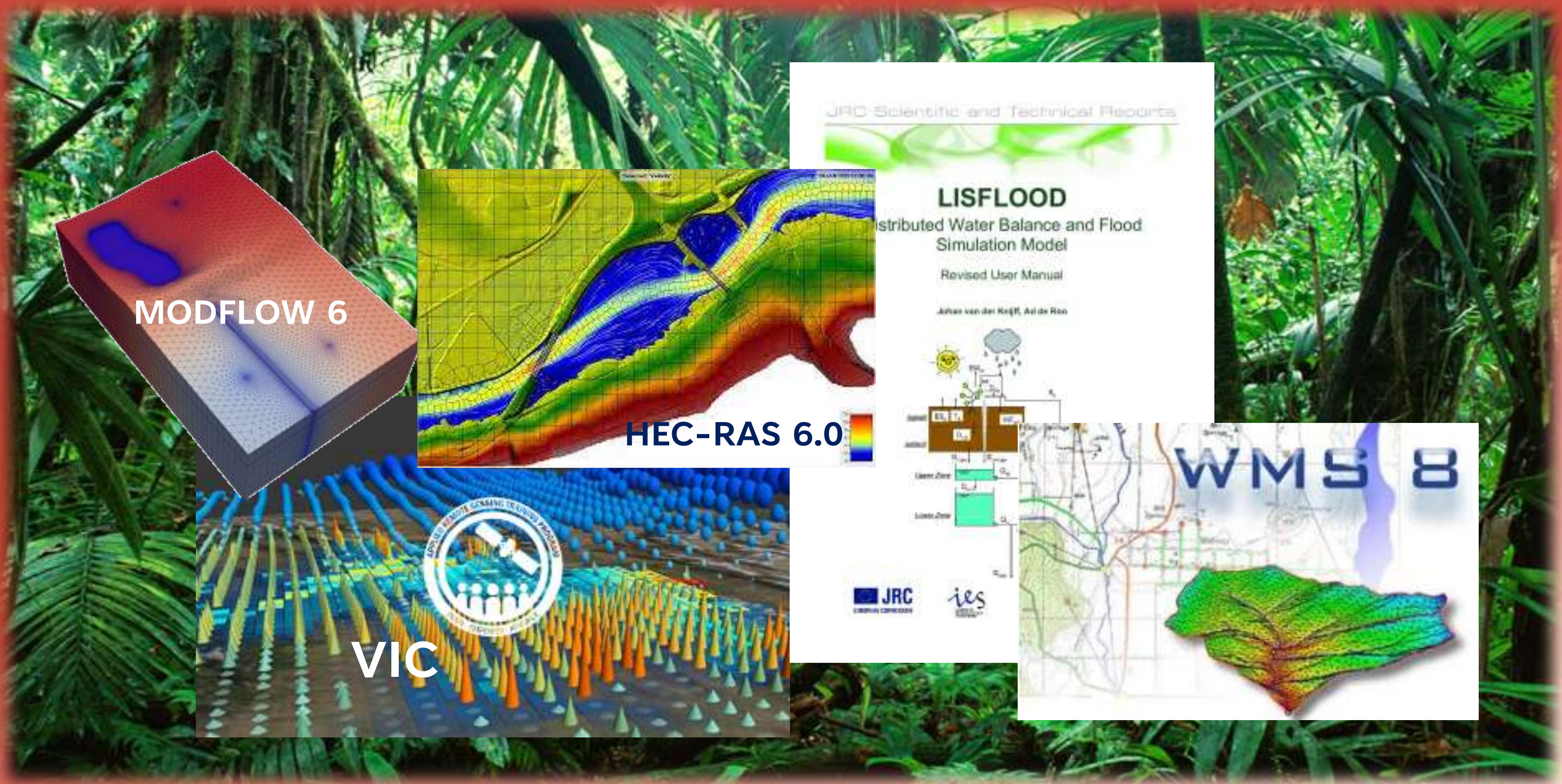
Microwave remote sensing of soil moisture

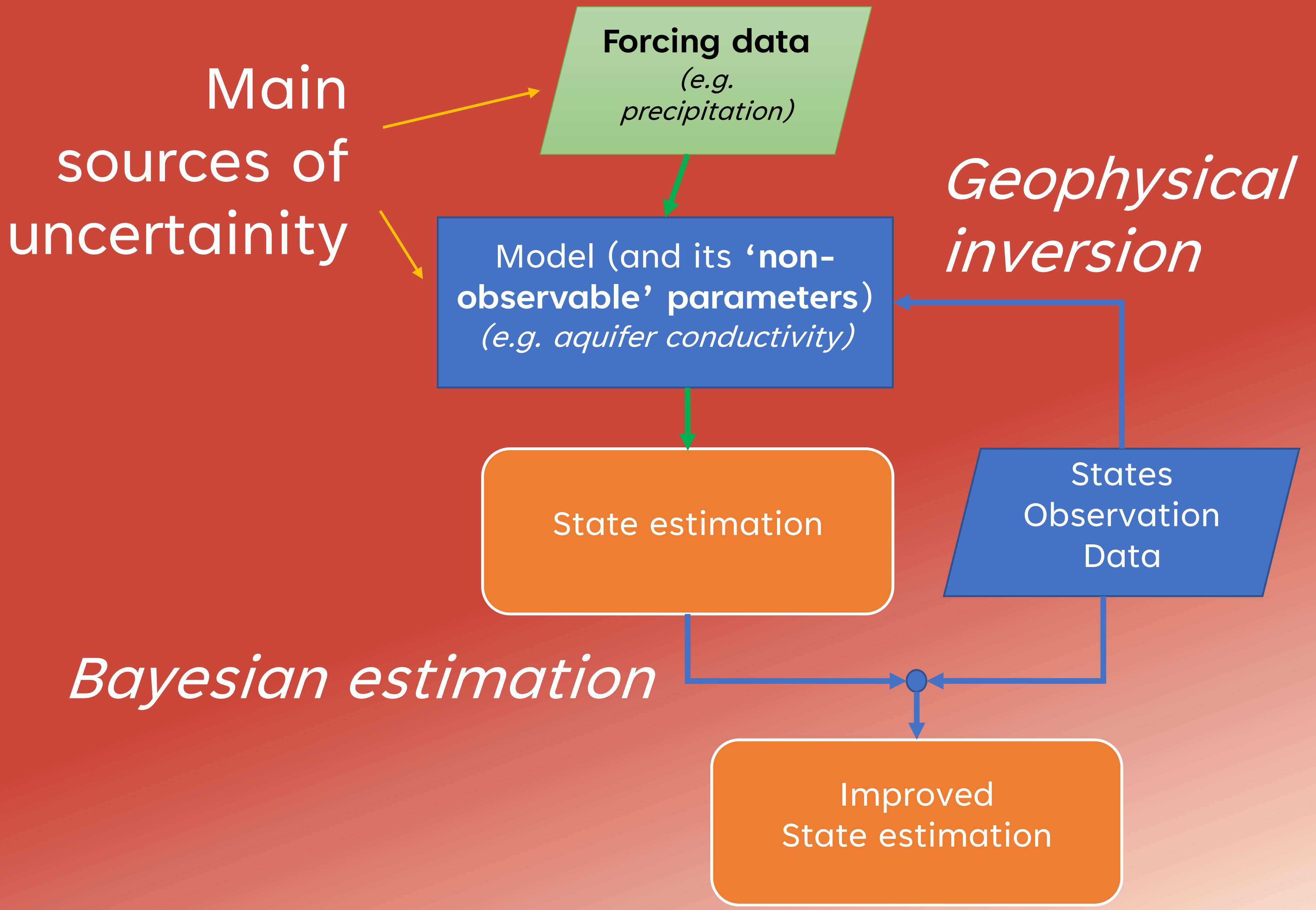


Courtesy of D. Entekhabi, MIT

Hydrology: what peculiarities for D.A.?

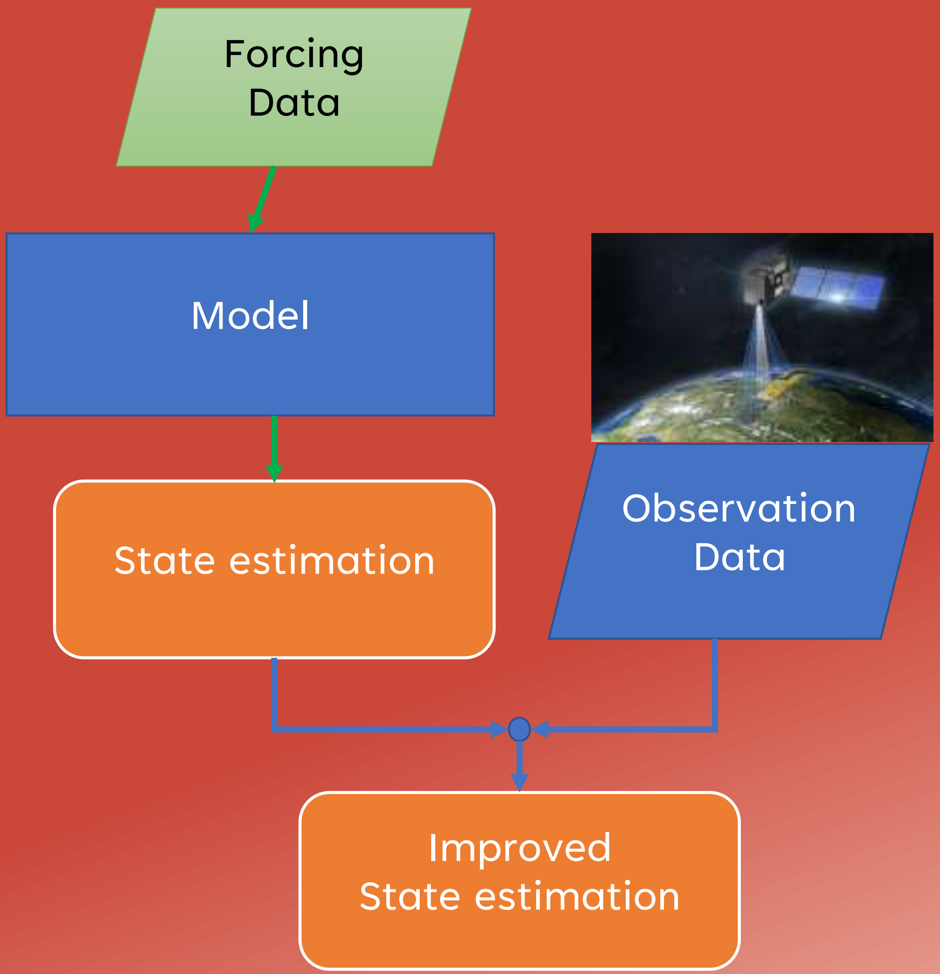
A pletora of problem/application specific models, with different level of conceptualizations and largely <u>unknown model biases</u>.





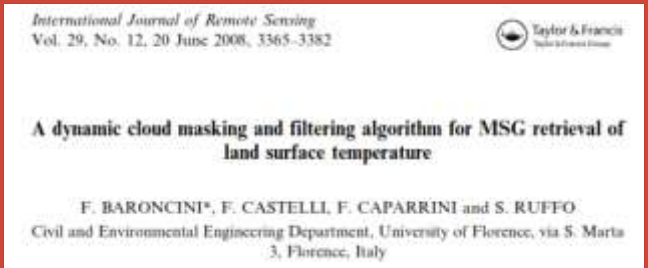
Main sources of uncertainity
Forcing data
(e.g. precipitation)
Model (and its 'non-observable' parameters)
(e.g. aquifer conductivity)
Geophysical inversion
State estimation
States Observation Data
Bayesian estimation
Improved State estimation

1. Improving remote sensing of land surface: dynamic (Kalman-based) filtering



- A linear example with a minimalistic model: *Filtering clouds from Meteosat-SEVIRI Land Surface Temperature*
- A non linear example with a global, distributed soil hydrology model: *Enhanced (Level-4) Soil Moisture products from the SMAP mission*

A pure filtering example:
Filtering cloud-contaminated LST observations from MSG-SEVIRI

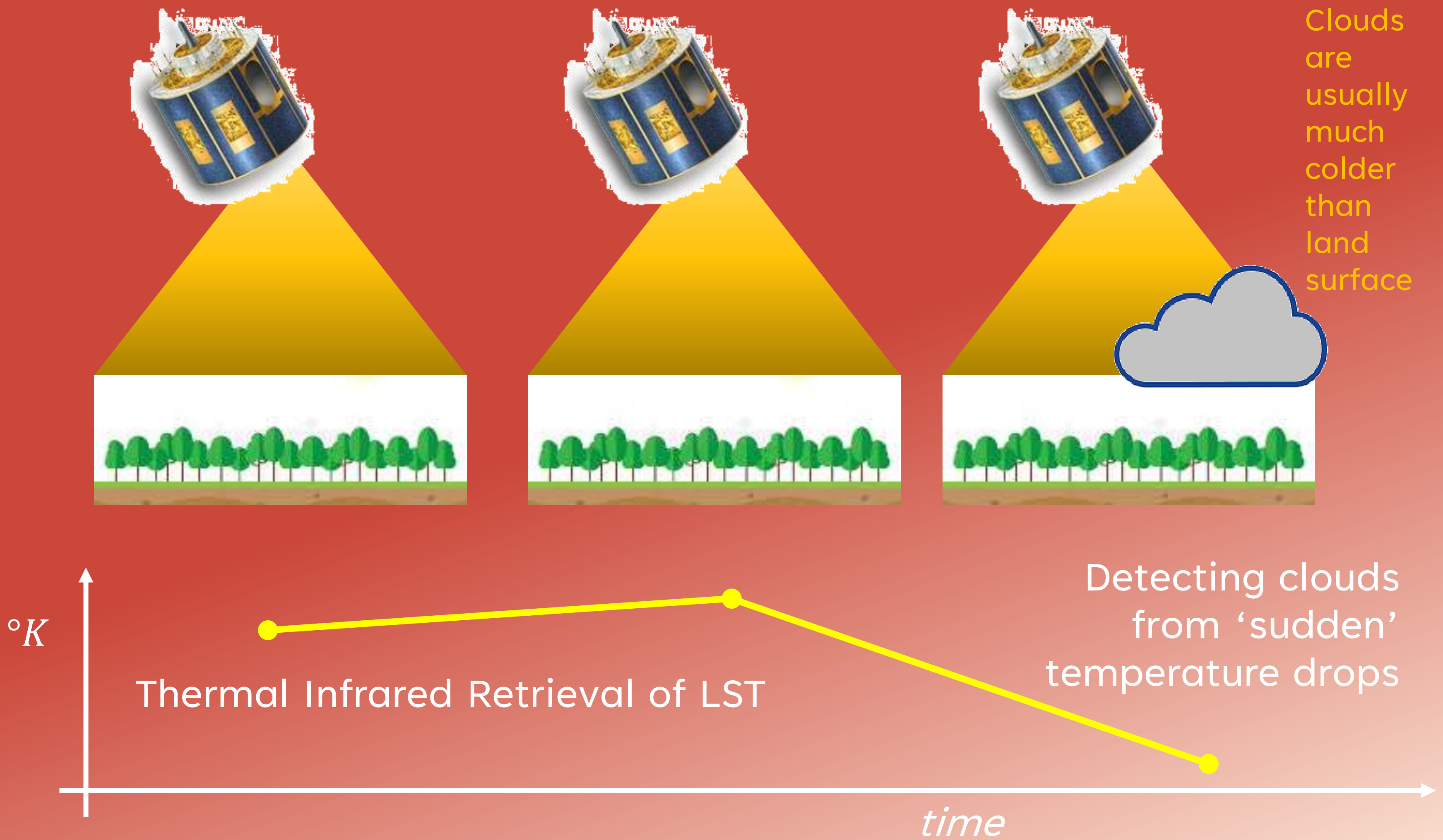


International Journal of Remote Sensing
Vol. 29, No. 12, 20 June 2008, 3365–3382

Taylor & Francis

A dynamic cloud masking and filtering algorithm for MSG retrieval of land surface temperature

F. BARONCINI*, F. CASTELLI, F. CAPARRINI and S. RUFFO
Civil and Environmental Engineering Department, University of Florence, via S. Marta 3, Florence, Italy



Incoming short-wave radiation

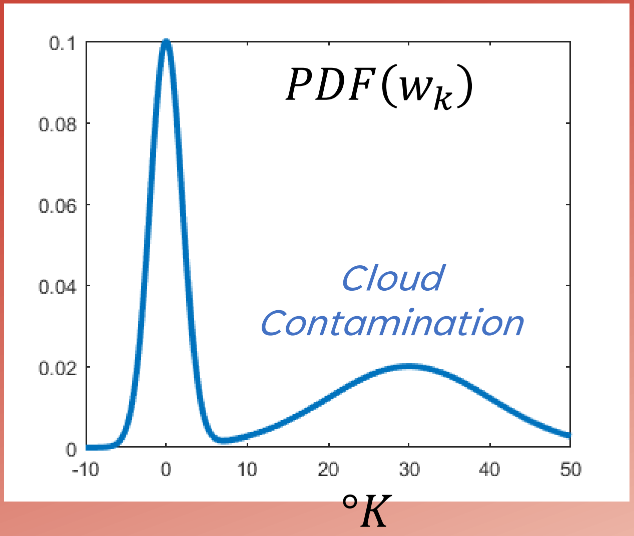
$$T_k = \alpha_k T_{k-1} + \beta S_k + v_k$$
$$T_k = LST_k + w_k$$

SEVIRI Land Surfce Temperature retrievals, 3km res., 30' revisit

Partial relaxation of Gaussian measurement error hypothesis:

- w_k comes from the mixture of a 'standard' $N(0, R^*)$ Gaussian noise ad a 'much larger' (non-0 mean!) cloud contamination error.
- Which error component is active at time k can be 'detected' with the innovation:

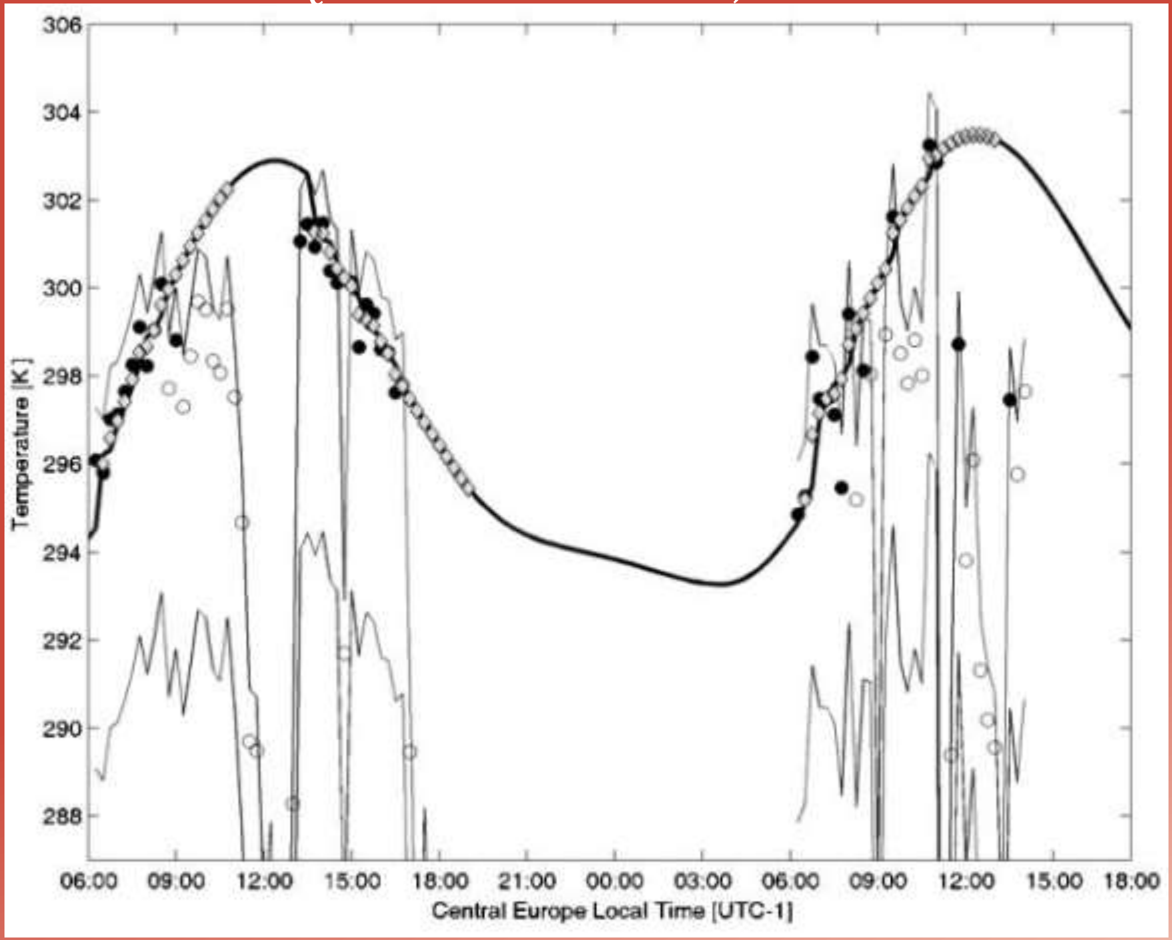
$$R_k = \begin{cases} R^* & if \quad T_k^- - LST_k \leq \delta(R^* + Q) \\ \infty & otherwise \end{cases}$$



$$R_k = \begin{cases} R^* & if \quad T_k^- - LST_k \leq \delta(R^* + Q)^{1/2} \\ \infty & otherwise \end{cases}$$

Kalman Gain is zero in this case, and the observation is discarded!

$$Q^{1/2} = R^{*\,1/2} = 1°K \quad ; \quad \delta = 2.5$$



Model prediction

Valid LST

Cloud contaminated (not used)

Reliable Analysis $(P_k^+ \leq R^* + Q)$

Table 1. Amount of validated LST estimates with different cloud-masking algorithms, as a percentage of the 533 724 total land pixels at SEVIRI resolution of the 28 ground-truth MODIS-based maps used for validation, and corresponding error statistics based on validation pixels with less than 50% MODIS cloud cover.

	Over all validation pixels	Over all validation pixels with less than 50% MODIS cloud cover	RMSE (K)	R^2
MODIS+Static	100% (533 724)	67.1% (358 390)	0	1.0
SEVIRI+Static	49.8% (265 911)	44.9% (239 492)	3.25	0.52
SEVIRI+CMKF (retained raw GSW estimate)	54.8% (292 569)	48.3% (257 651)	3.19	0.59
SEVIRI+CMKF (valid posterior estimate)	78.3% (417 846)	60.7% (324 171)	3.33	0.54

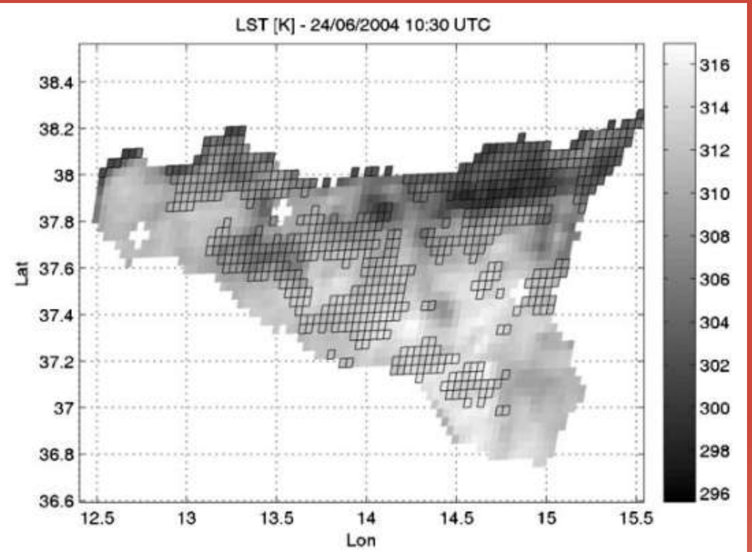
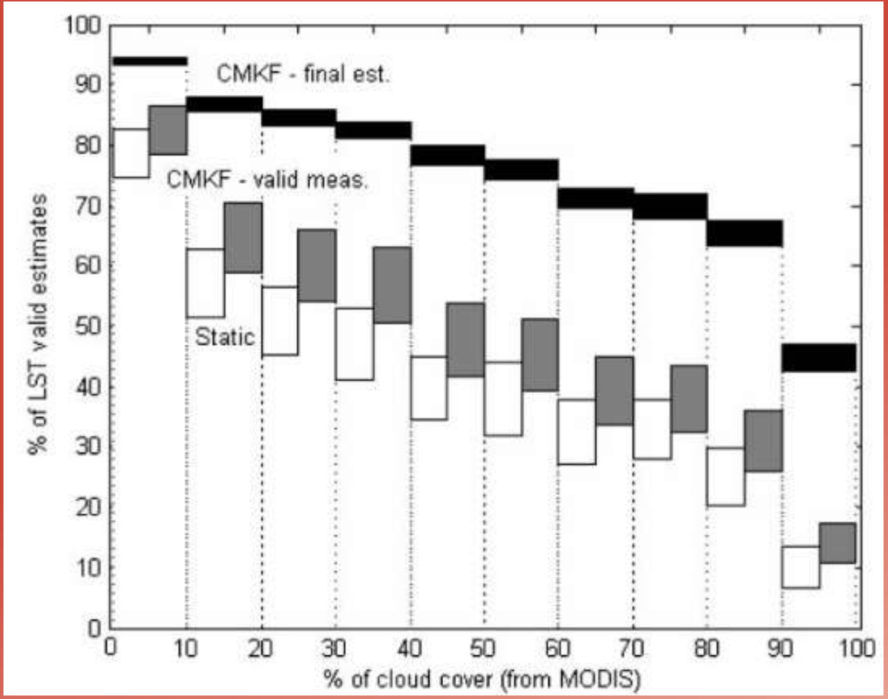
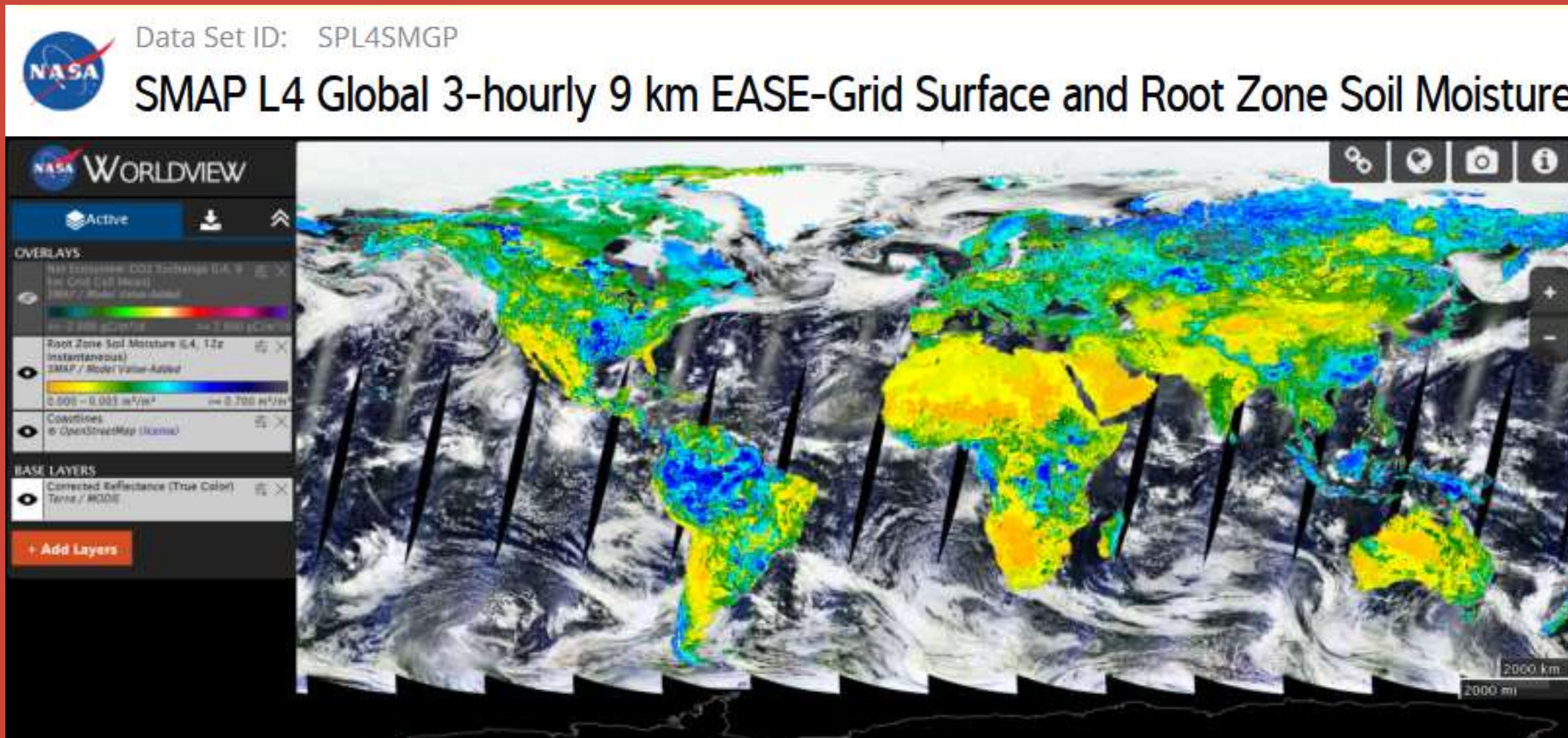


Figure 5. Enlargement of the map shown in figure 4 over Sicily. Boxed pixels identify regions where the LST retrieval was based on model prediction only, having recognized the GSW estimates as being too cloud contaminated.



Validation with MODIS

Sate estimation with sequential non-lienar filtering:
Enhanced Soil Moisture products from the SMAP mission



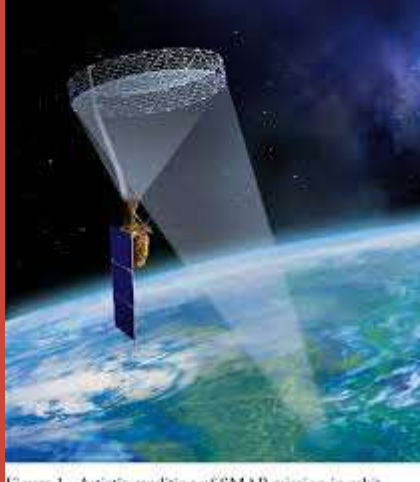
JAMES | Journal of Advances in Modeling Earth Systems

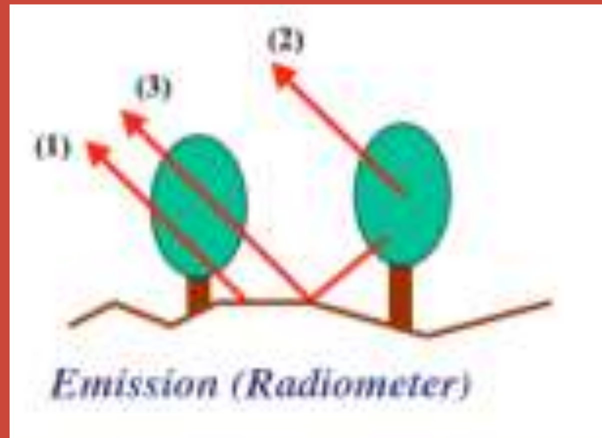
Research Article | Open Access

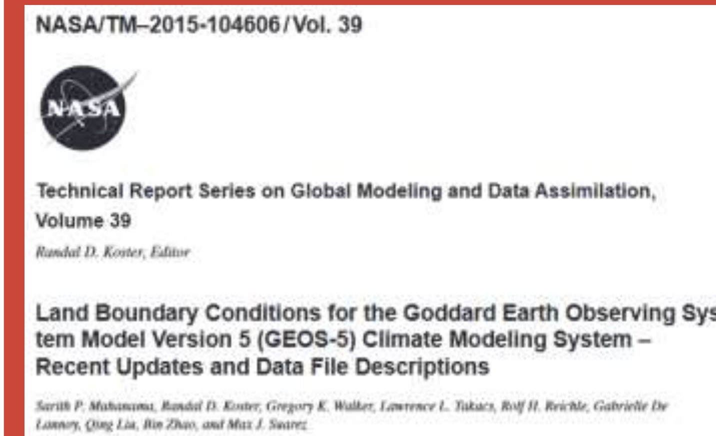
Version 4 of the SMAP Level-4 Soil Moisture Algorithm and Data Product

Rolf H. Reichle, Qing Liu, Randal D. Koster, Wade T. Crow, Gabrielle J. M. De Lannoy, John S. Kimball, Joseph V. Ardizzone, David Bosch, Andreas Colliander, Michael Cosh, Jana Kolassa, ... See all authors

First published: 02 September 2019 | https://doi.org/10.1029/2019MS001729 | Citations: 35






NASA/TM–2015-104606/Vol. 39

Technical Report Series on Global Modeling and Data Assimilation, Volume 39

Randal D. Koster, Editor

Land Boundary Conditions for the Goddard Earth Observing System Model Version 5 (GEOS-5) Climate Modeling System – Recent Updates and Data File Descriptions

Sarith P. Mahanama, Randal D. Koster, Gregory K. Walker, Lawrence L. Takacs, Rolf H. Reichle, Gabrielle De Lannoy, Qing Liu, Bin Zhao, and Max J. Suarez

- **SMAP L1C_TB observations**
- **Distributed ensemble Kalman filter (EnKF)**
- **NASA GEOS-5 Catchment land surface model**
- Surface meteorological forcing from NASA GEOS-5 Forward Processing system
- Precipitation corrections with NOAA Climate Prediction Center "Unified" global, 0.5 degree gauge-based product

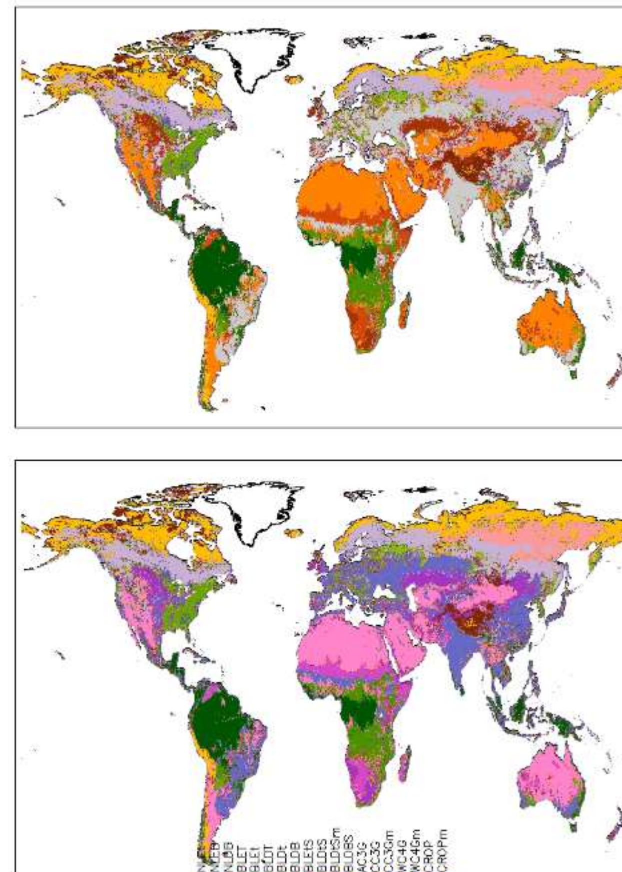
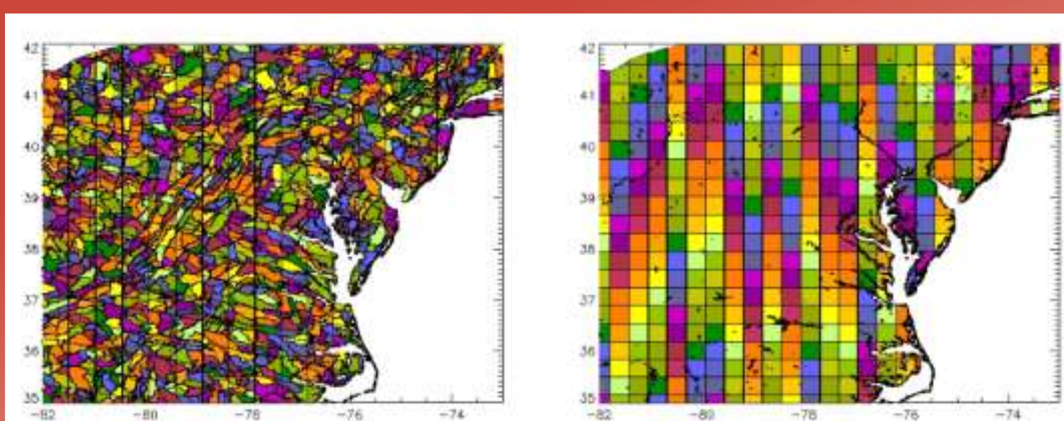
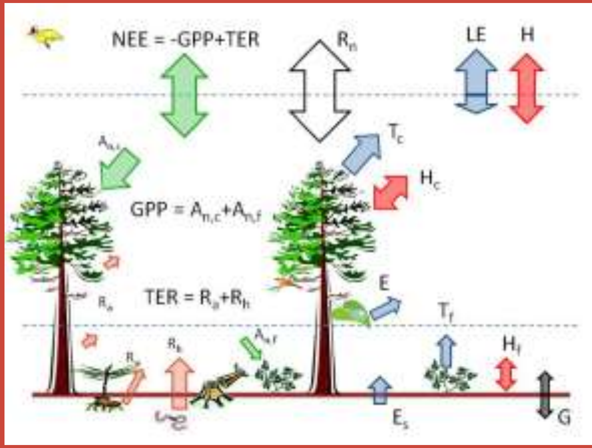


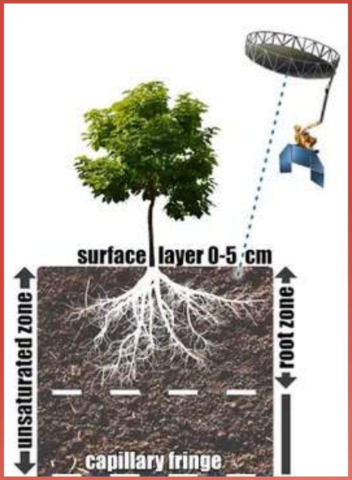
Figure 3: Catchment-CN (top) primary vegetation types, and (bottom) secondary vegetation types.



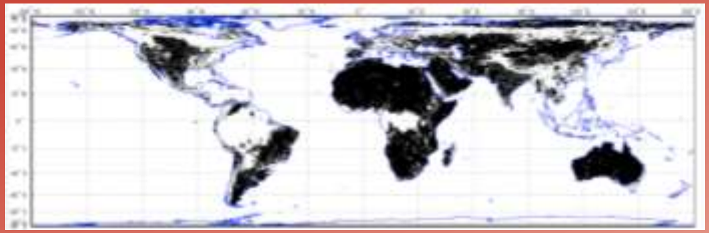
Assimilating Microwave TB in a hydrologic model resolving water, carbon and energy balance of the soil-vegetation layers improve the Level 2 original SMAP-Soil Moisture product by:

- Downscale the from the nominal 36km resolution of the TB data to the 9km* resolution of the distributed DA system
- Extend the estimate from the top 5 cm (influencing the mw emission) to various depths down to 100 cm
- Extend the estimate to regions where the TB data decorrelate with soil moisture (high vegetation, terrain slope variability)

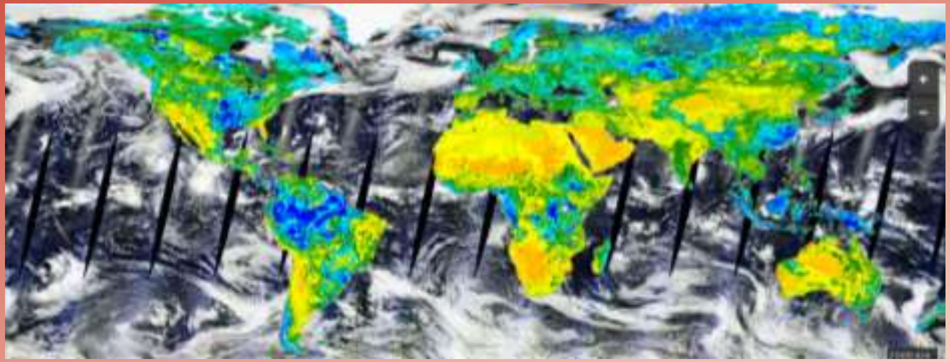




* Intended resolution of the merged active-passive product before radar failure

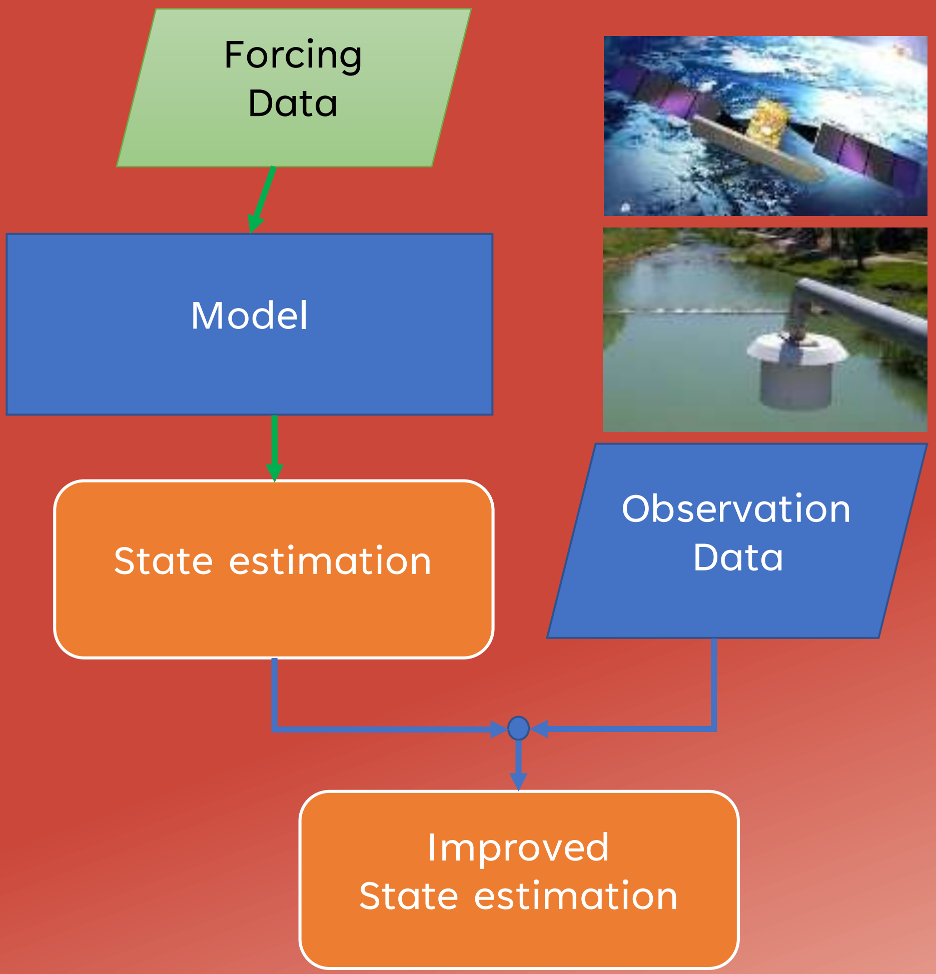


Level-2 Quality Mask



Sample of Level-4 Root-zone S.M.

2. State estimation with dynamic (Kalman-based) filtering and sparse observations



- *Near-realtime flood mapping with assimilation of river stage data (sparse in space) and high-resolution multispectral water extent (sparse in time) into a 2D hydrologic-hydraulic model*

https://doi.org/10.5194/hess-2021-125
Preprint. Discussion started: 8 A...

under final review

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

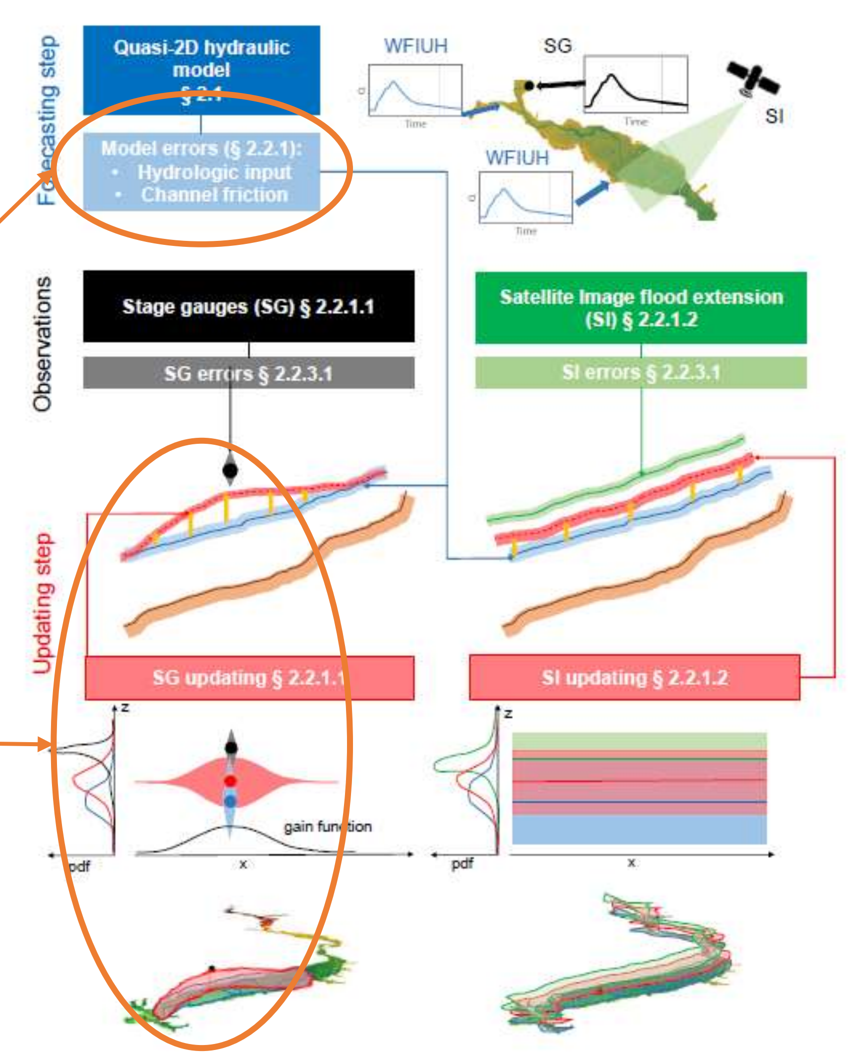
Simultaneous assimilation of water levels from river gauges and satellite flood maps for near-real time flood mapping

Antonio Annis[1,2], Fernando Nardi[1,3], and Fabio Castelli[2]

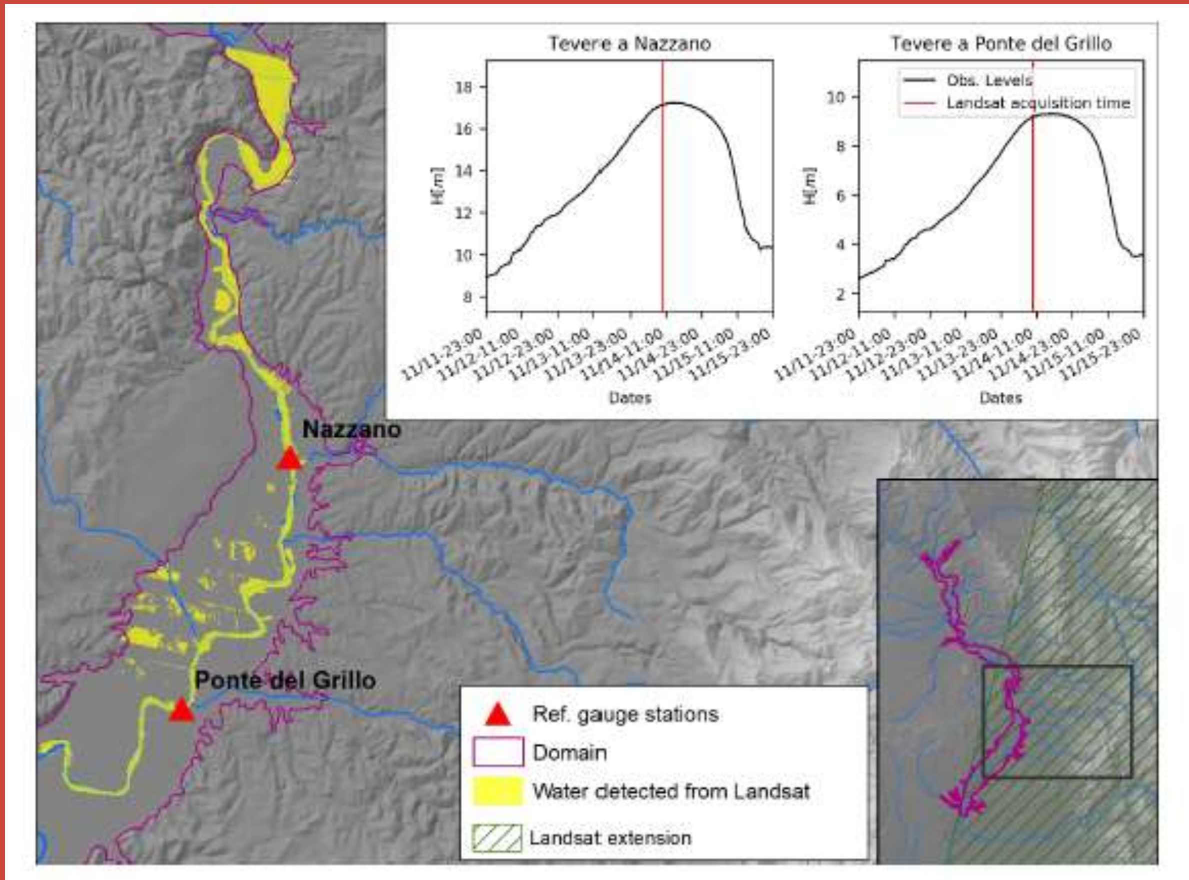
EnKF Scheme

Model error covariance estimated by perturbing hydrologic input (precipitation) and more uncertain model parameters (channel and floodplain roughness, terrain elevation)

River gauge observations are very sparse. Need of filter localization to avoid spurious error growth in limited-size ensemble sampling

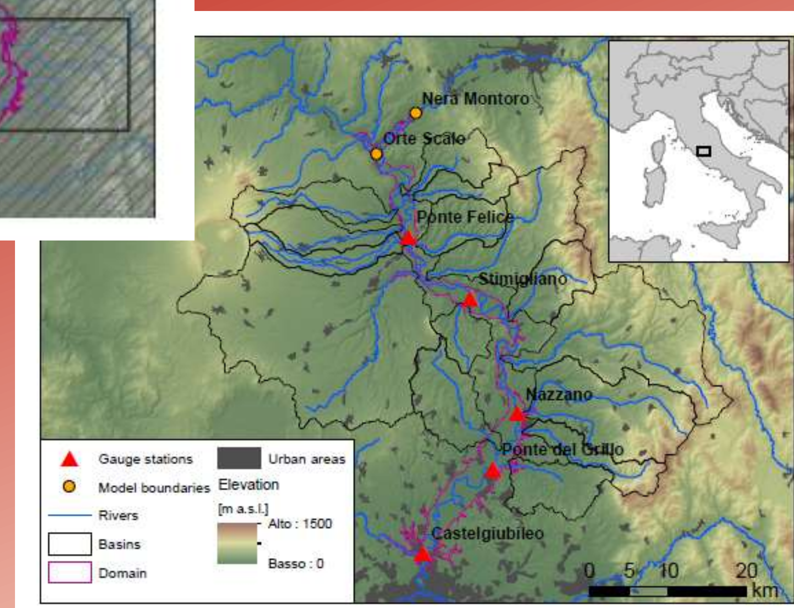


Case study: Nov. 2012 flooding of the Tiber valley (upstream of Rome)



Available data:

- 5 half-hourly water stage records from river gauges (4 inside the flooded area)
- 1 flooded area map from Landsat (partly cloudy) near time of flood peak

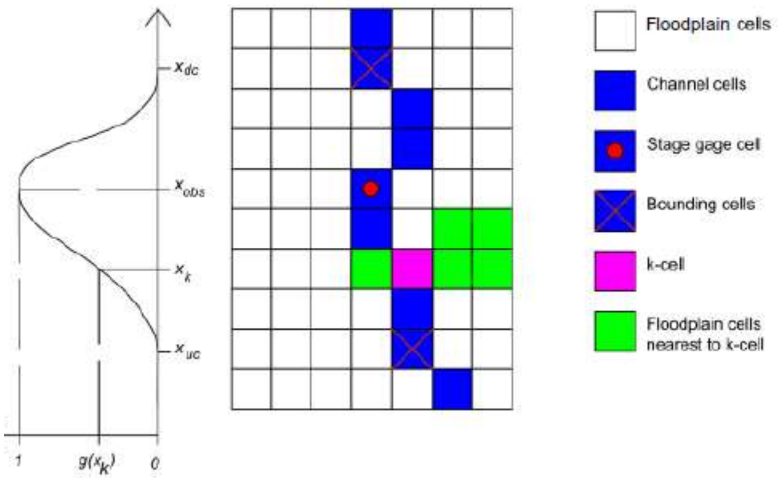


Localization applied in the EnKF update step (takes advantage of the **perfectly predictable flow direction** along the channel network)

Along-channel distance

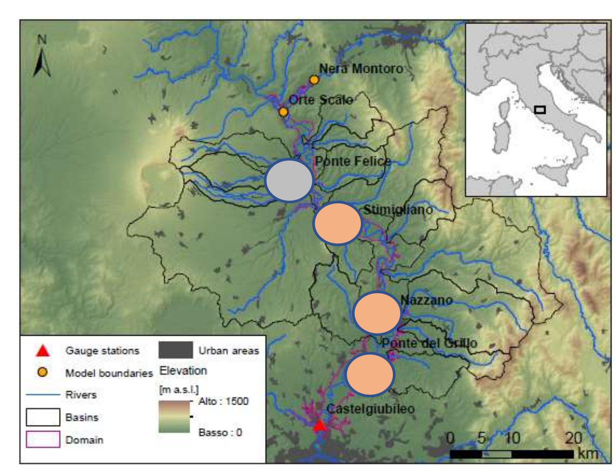
$$g(x_k) = A \cdot exp\left(-\frac{1}{2}\left(\frac{g'(x_k)}{1/3}\right)^2\right)$$

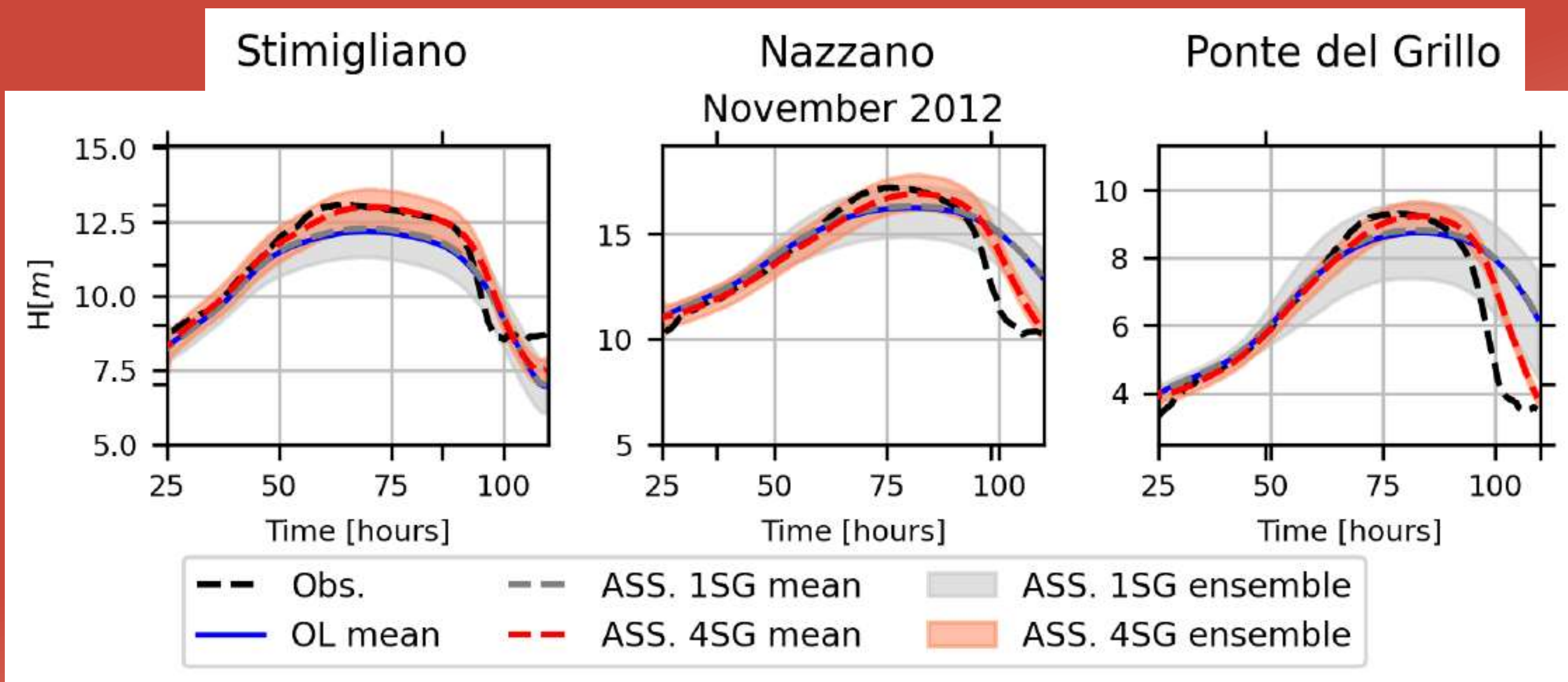
$$g'(x_k) = \begin{cases} \frac{x_{obs}-x_k}{x_{obs}-x_{uc}}, & x_{uc} \leq x_k \leq x_{obs} \\ \frac{x_k-x_{obs}}{x_{dc}-x_{obs}}, & x_{obs} \leq x_k \leq x_{dc} \end{cases}$$

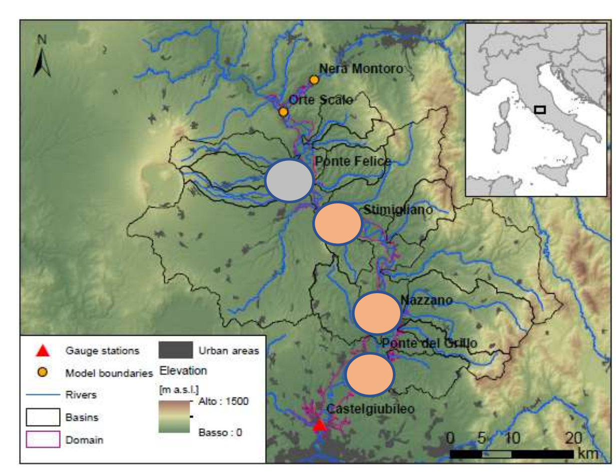


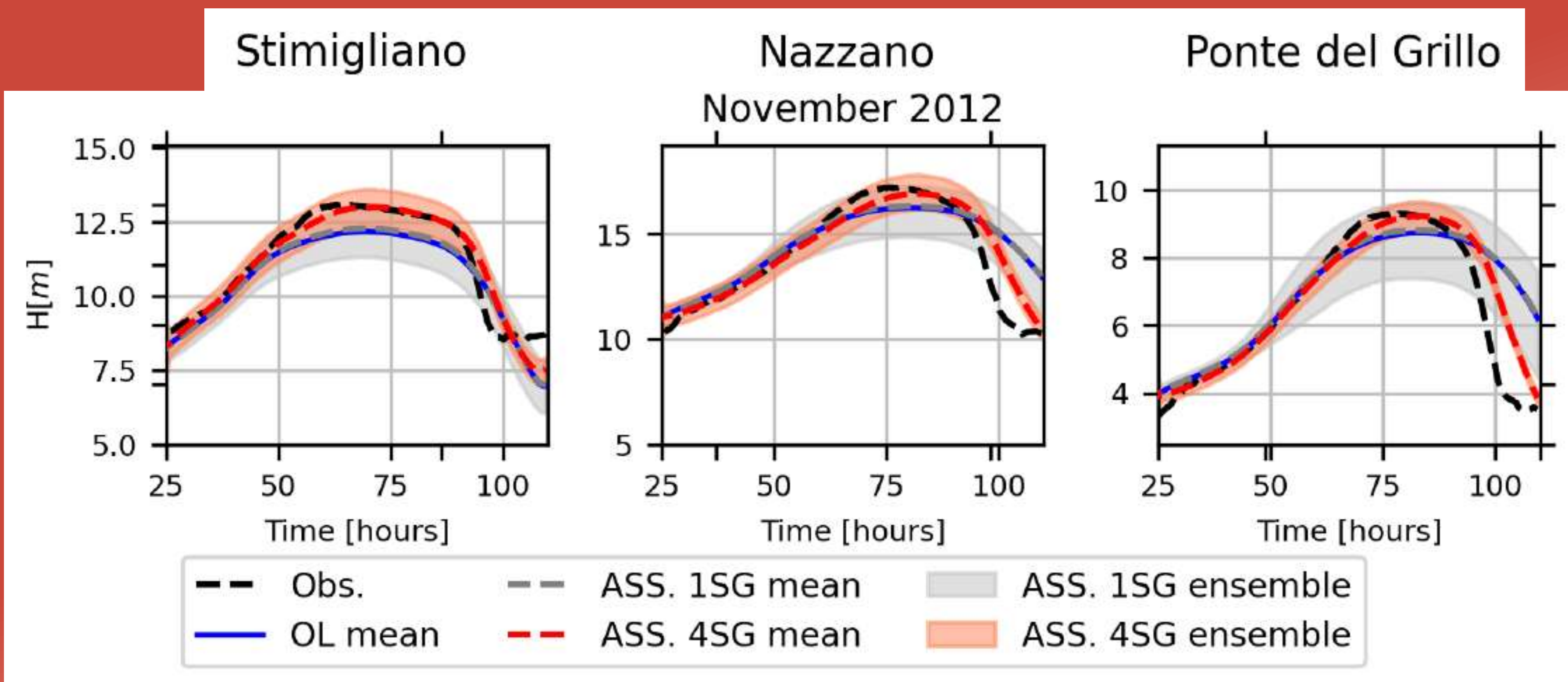
The ‘standard’ EnKF update is applied to the river gauges cells first, water level in neighboring cells is then updated as:

$$\Delta h(x_k) = \frac{\Delta h(x_{obs,u}) \cdot g(x_{k,u}) \cdot \frac{1}{x_k - x_{obs,u}} + \Delta h(x_{obs,d}) \cdot g(x_{k,d}) \cdot \frac{1}{x_{obs,d} - x_k}}{\frac{1}{x_{obs,d} - x_{obs,u}}}$$

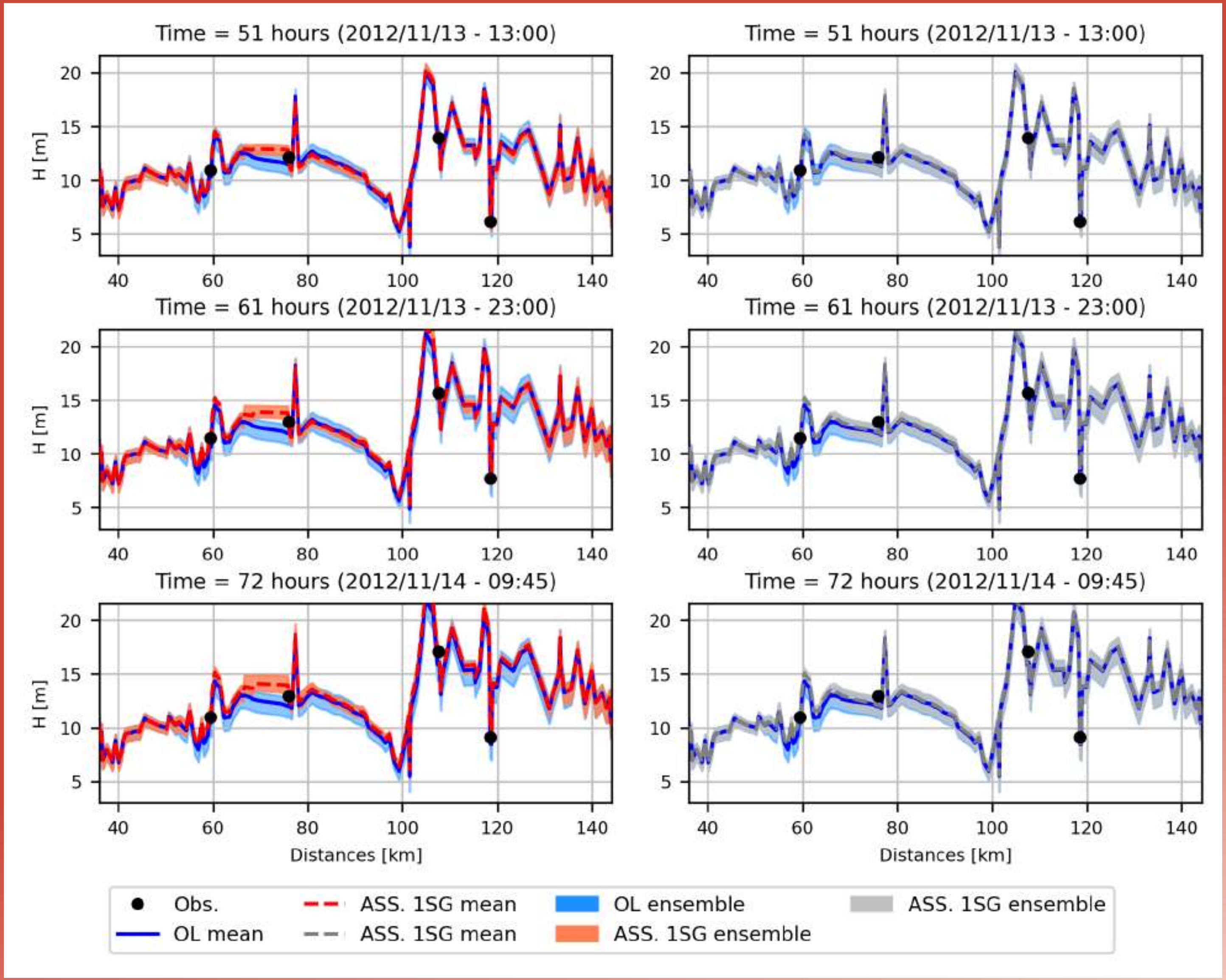

N
Nera Montoro
Orte Scalo
Ponte Felice
Stimigliano
Nazzano
Ponte del Grillo
Castelgiubileo
Gauge stations
Model boundaries
Rivers
Basins
Domain
Urban areas
Elevation
[m a.s.l.]
Alto : 1500
Basso : 0
0 5 10 20
km


Stimigliano
Nazzano
Ponte del Grillo
November 2012
H[m]
Time [hours]
Obs.
OL mean
ASS. 1SG mean
ASS. 4SG mean
ASS. 1SG ensemble
ASS. 4SG ensemble


N
Nera Montoro
Orte Scalo
Ponte Felice
Stimigliano
Nazzano
Ponte del Grillo
Castelgiubileo
Gauge stations
Model boundaries
Rivers
Basins
Domain
Urban areas
Elevation
[m a.s.l.]
Alto : 1500
Basso : 0
0 5 10 20
km

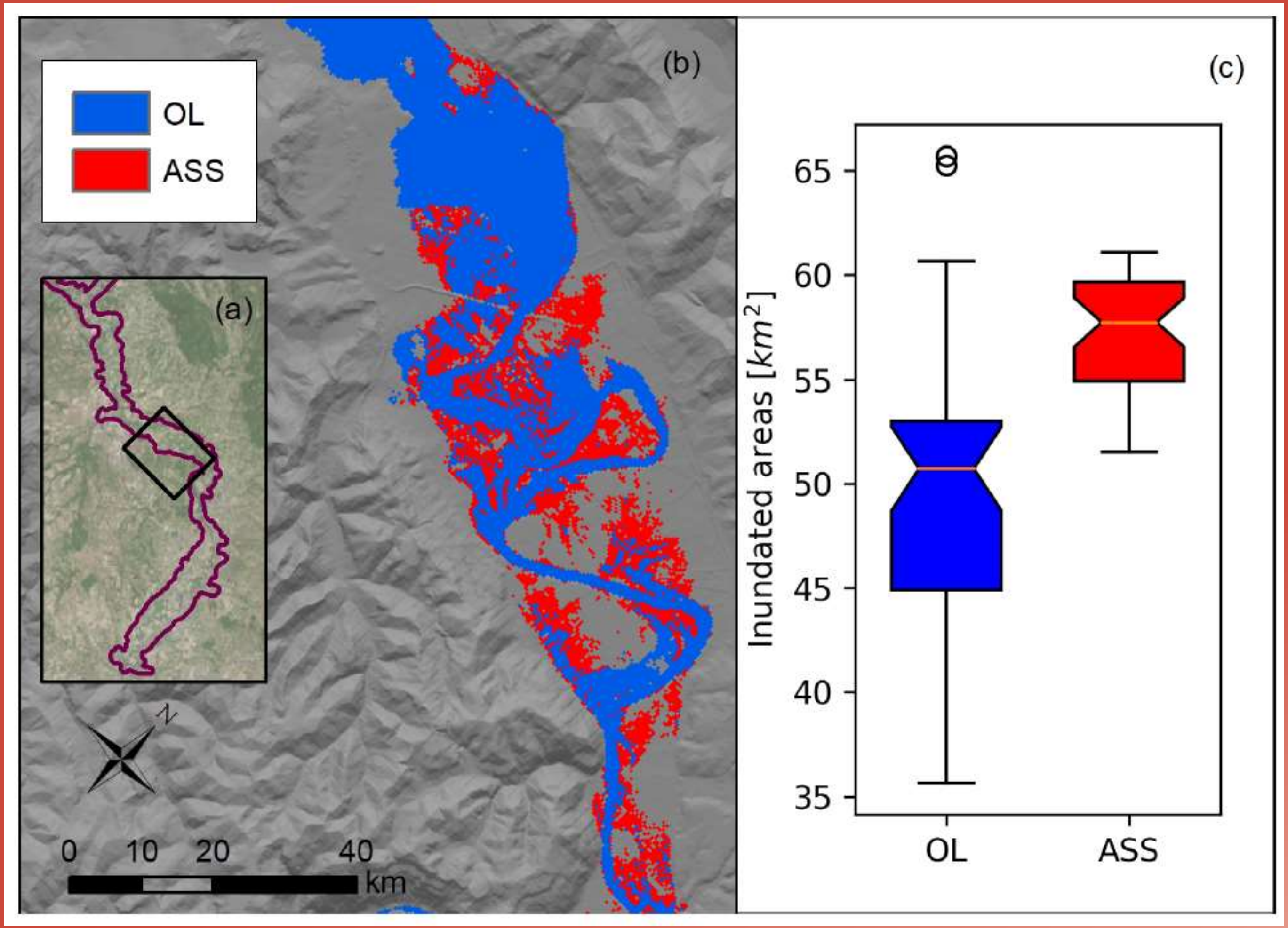

Stimigliano
Nazzano
Ponte del Grillo
November 2012
H[m]
Time [hours]
Obs.
OL mean
ASS. 1SG mean
ASS. 4SG mean
ASS. 1SG ensemble
ASS. 4SG ensemble

Water depth profiles along the main channel



Flow direction

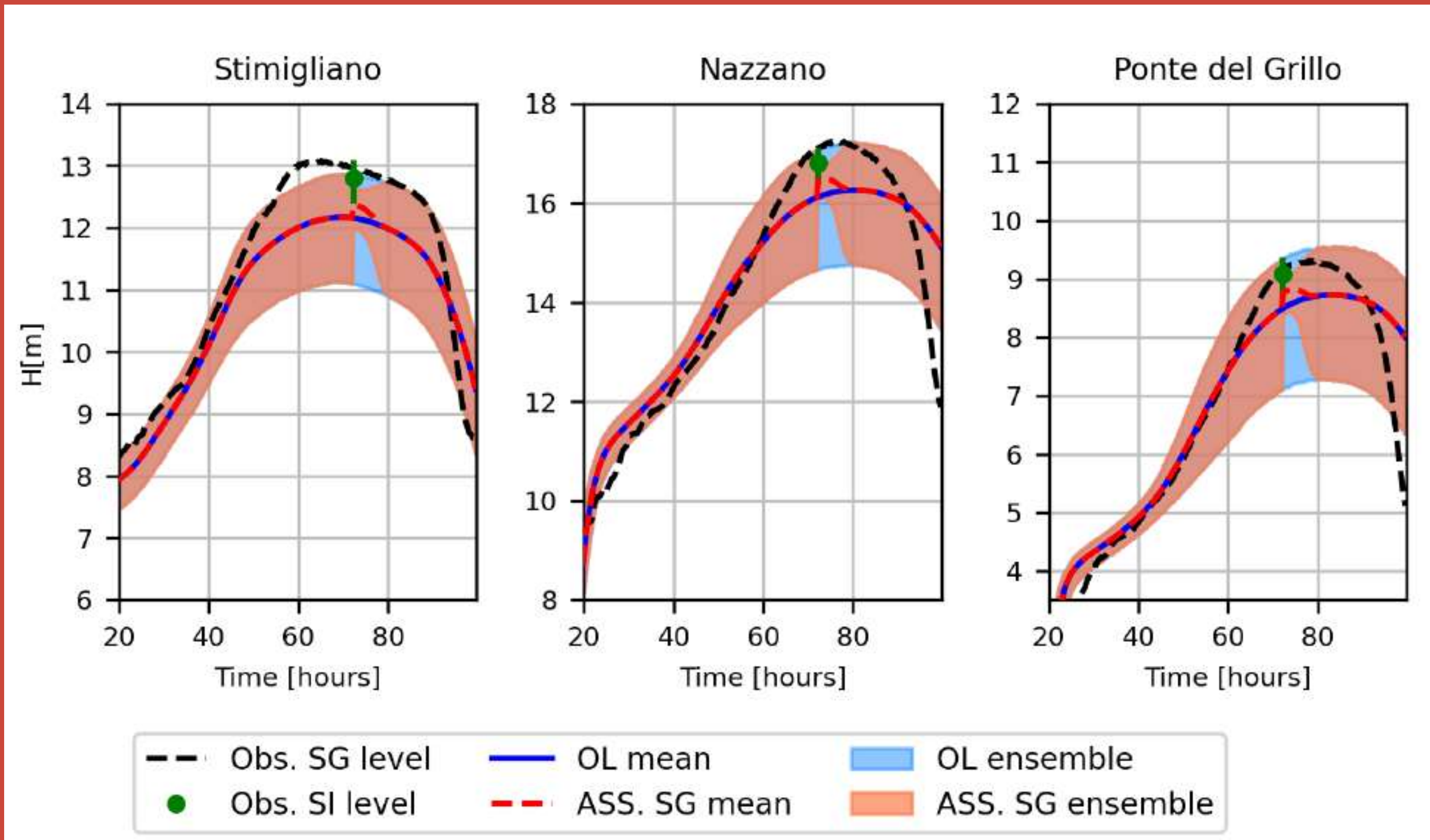
Assimilation of Landsat flood extent alone



Clear reduction of flood extent uncertainty.
Accuracy?

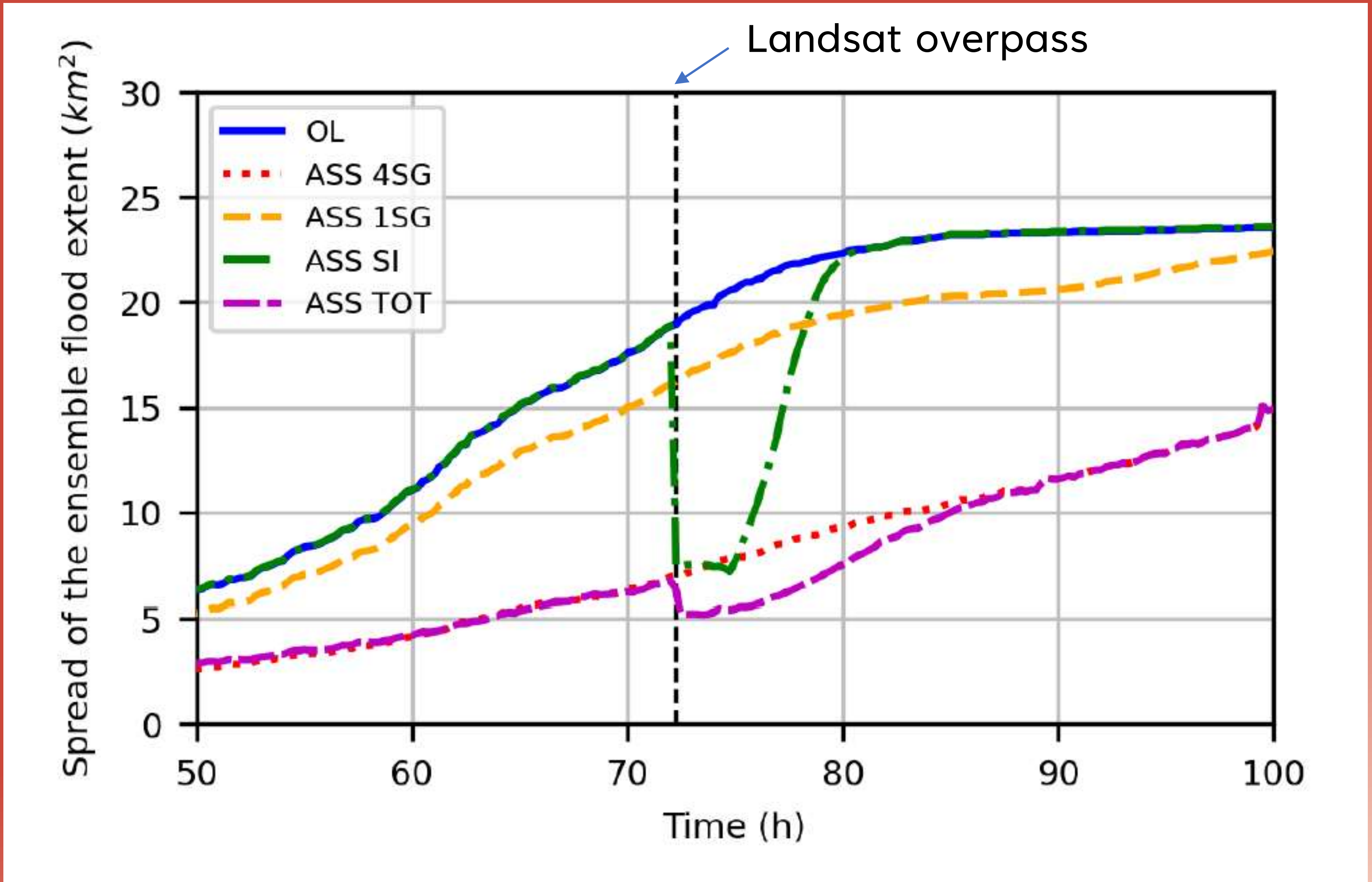
Assimilation of Landsat flood extent alone

Effects of flood extent assimilation on water stage at river gauges

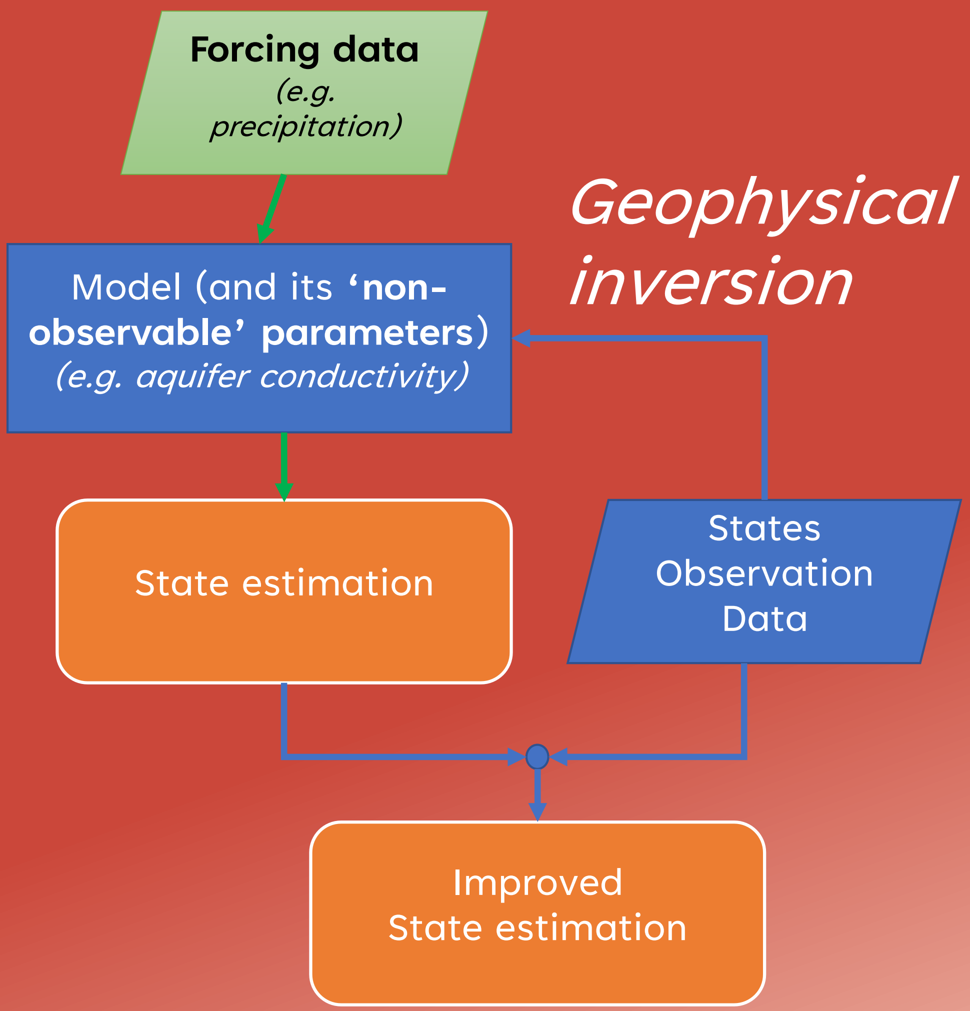


Quite short persistence of D.A. benefit
(... convergent system!)

Benefits of joint assimilation of ground water stages and satellite flood extent



3. Geophysical inversion as a causal identification and quantification tool (*true backward reasoning*)



- *Land subsidence induced by excessive groundwater abstraction*

Inverse problems: Reasoning backwards

Most people, if you describe a train of events to them will tell you what the result will be. There are few people, however that if you told them a result, would be able to evolve from their own inner consciousness what the steps were that led to that result. This power is what I mean when I talk of reasoning backward.

Sherlock Holmes,
A Study in Scarlet,
Sir Arthur Conan Doyle (1887)

Study area

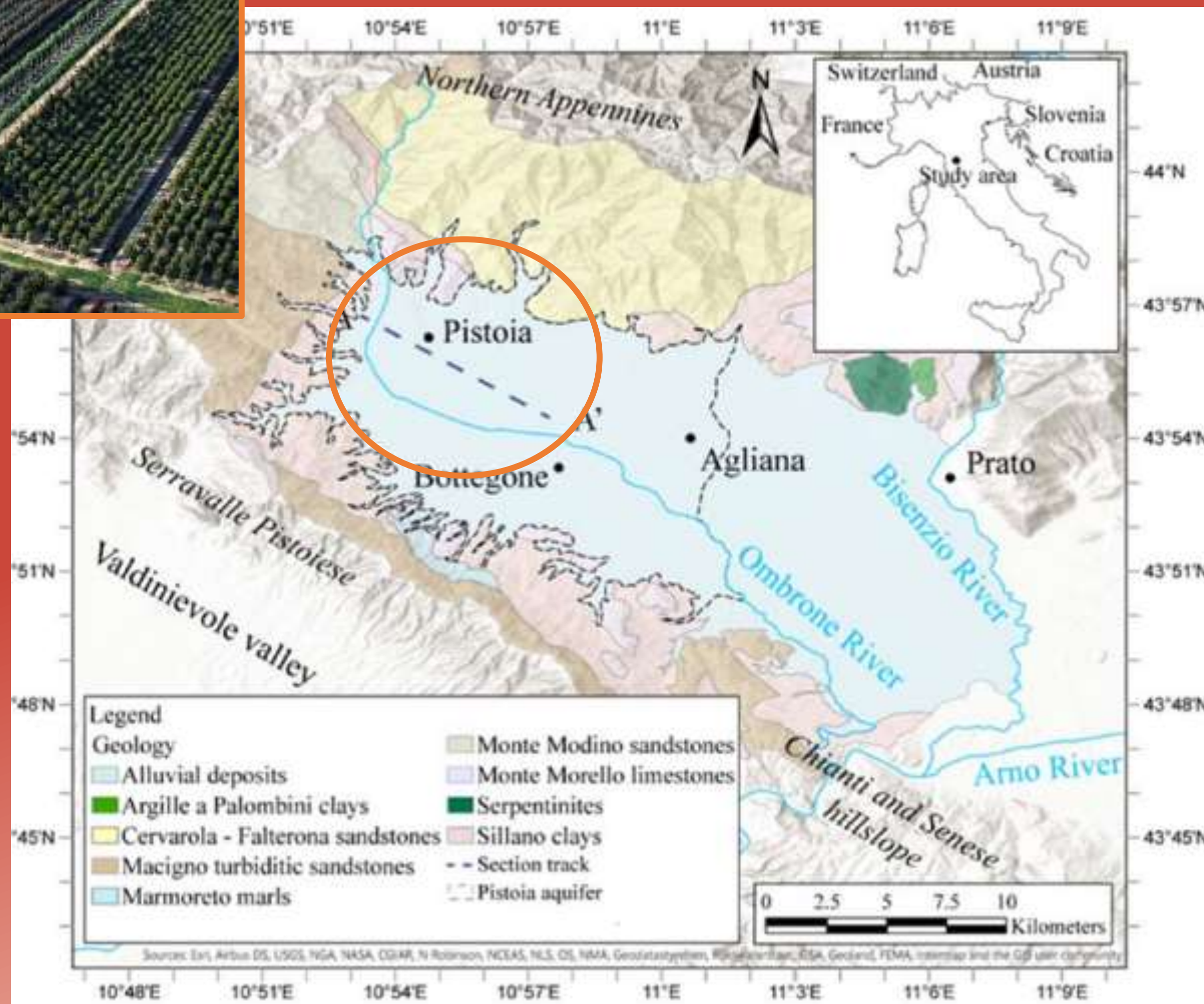
Hydrogeology Journal (2021) 29:629–649
https://doi.org/10.1007/s10040-020-02255-2

PAPER

Numerical modelling of land subsidence related to groundwater withdrawal in the Firenze-Prato-Pistoia basin (central Italy)

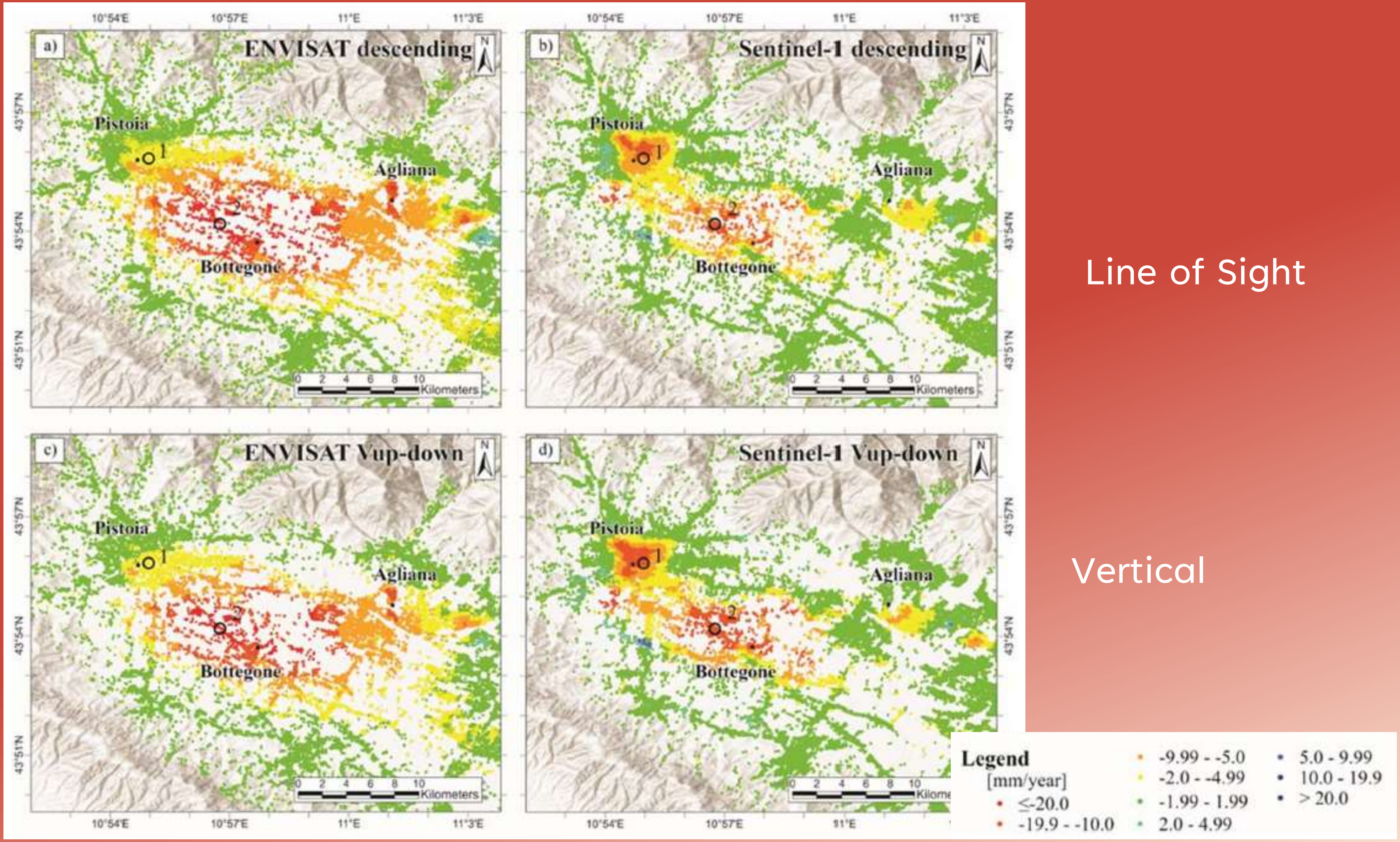
Mattia Ceccatelli[1] · Matteo Del Soldato[1] · Lorenzo Solari[2] · Riccardo Fanti[1] · Gaddo Mannori[3] · Fabio Castelli[4]

Strong water abstraction from an alluvial aquifer to support plant nursery irrigation in the Pistoia area



Satellite SAR interferometric estimates of surface deformation velocity

ENVISAT 2003-2010 Sentinel-1 2015-2017



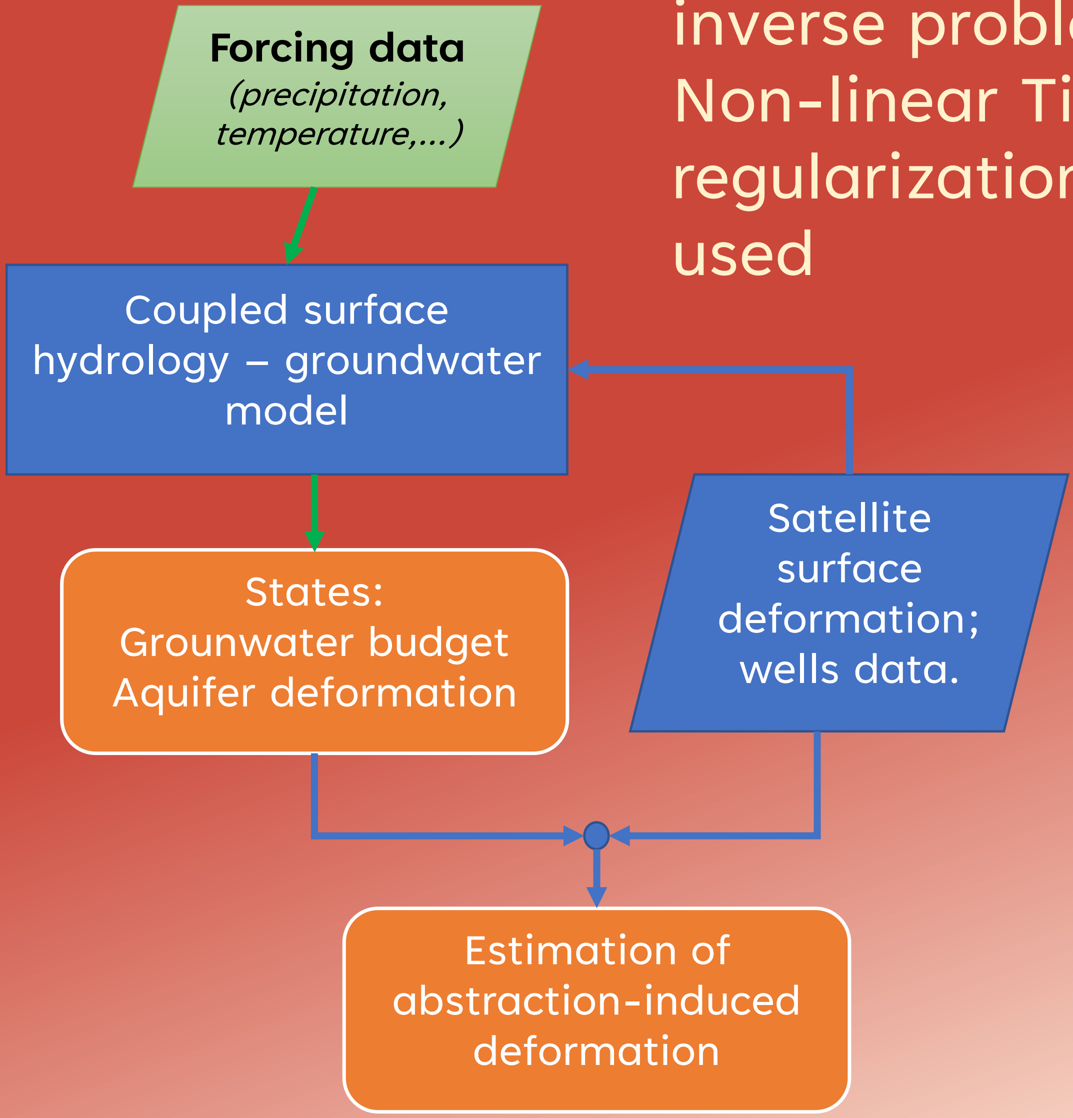
Line of Sight

Vertical

Variational Data Assimilation Framework

Aquifer parameter estimated with D.A. (spatially variable):

- Hydraulic conductivity
- Elastic and inelastic skeletal storage coefficients



Underdetermined inverse problem: Non-linear Tikhonov regularization is used

Tikhonov regularization (*general definition*)

In the simplest linear inverse problem: find $\boldsymbol{\theta}$ such that $\boldsymbol{G}(\boldsymbol{x})\boldsymbol{\theta} = \boldsymbol{z}$ the problem may be **ill-posed** (e.g. underdetermined system of equations due to a larger number of model parameters than observations, or collinear observations).

$\Rightarrow$ the inverse mapping operates as a high-pass filter that has the undesirable tendency of amplifying noise

This tendency can be alleviated by adding **a regularization term** to the square residuals to be minimized

$$\|\boldsymbol{G\theta} - \boldsymbol{z}\|^2 + \|\underset{\uparrow}{\boldsymbol{\Gamma}}\boldsymbol{\theta}\|^2 \Rightarrow \widehat{\boldsymbol{\theta}} = \underbrace{(\boldsymbol{G}^T\boldsymbol{G} + \boldsymbol{\Gamma}^T\boldsymbol{\Gamma})^{-1}}\boldsymbol{G}^T\boldsymbol{z}$$

Tikhonov matrix

A properly chosen Tikhonov term guarantees the existence and indetifiability of this inverse

Tikhonov regularization in a D.A. (Bayesian) framework

Solve the ill-posed inversion of $\boldsymbol{G}(\boldsymbol{x})\boldsymbol{\theta} = \boldsymbol{z}$, given:

- $\boldsymbol{\theta}_0$ = a prior estimate (expected value) of the unknown parameters $\boldsymbol{\theta}$
- $\boldsymbol{Q}$ = the covariance of the unknown $\boldsymbol{\theta}$
- $\boldsymbol{P}$ = the covariance of the data $\boldsymbol{z}$

Need to be prescribed!

Find the minimum of:

$$\|(\boldsymbol{G\theta} - \boldsymbol{z})^T \boldsymbol{P}(\boldsymbol{G\theta} - \boldsymbol{z})\| + \|(\boldsymbol{\theta} - \boldsymbol{\theta}_0)^T \boldsymbol{Q}(\boldsymbol{\theta} - \boldsymbol{\theta}_0)\|$$

$$\widehat{\boldsymbol{\theta}} = (\boldsymbol{G}^T\boldsymbol{PG} + \boldsymbol{Q})^{-1}(\boldsymbol{G}^T\boldsymbol{Pz} + \boldsymbol{Q\theta}_0) = \boldsymbol{\theta}_0 + (\boldsymbol{G}^T\boldsymbol{PG} + \boldsymbol{Q})^{-1}\left(\boldsymbol{G}^T\boldsymbol{P}(\boldsymbol{z} - \boldsymbol{G\theta}_0)\right)$$

In this case, the Tikhonov matrix would be a factorization of the covariance of the unknowns $\boldsymbol{Q} = \boldsymbol{\Gamma}^T\boldsymbol{\Gamma}$

Compare Tikhonov Linear Inverse with Kalman Filter Update

$$\widehat{\boldsymbol{\theta}} = \boldsymbol{\theta}_0 + (\boldsymbol{G}^T\boldsymbol{P}\boldsymbol{G} + \boldsymbol{Q})^{-1}\left(\boldsymbol{G}^T\boldsymbol{P}(\boldsymbol{z} - \boldsymbol{G}\boldsymbol{\theta}_0)\right)$$

Hint:
$\boldsymbol{G} = \boldsymbol{H}$

$$\boldsymbol{K}_k = \boldsymbol{P}_k^-\boldsymbol{H}^T(\boldsymbol{H}\boldsymbol{P}_k^-\boldsymbol{H}^T + \boldsymbol{Q})^{-1}$$

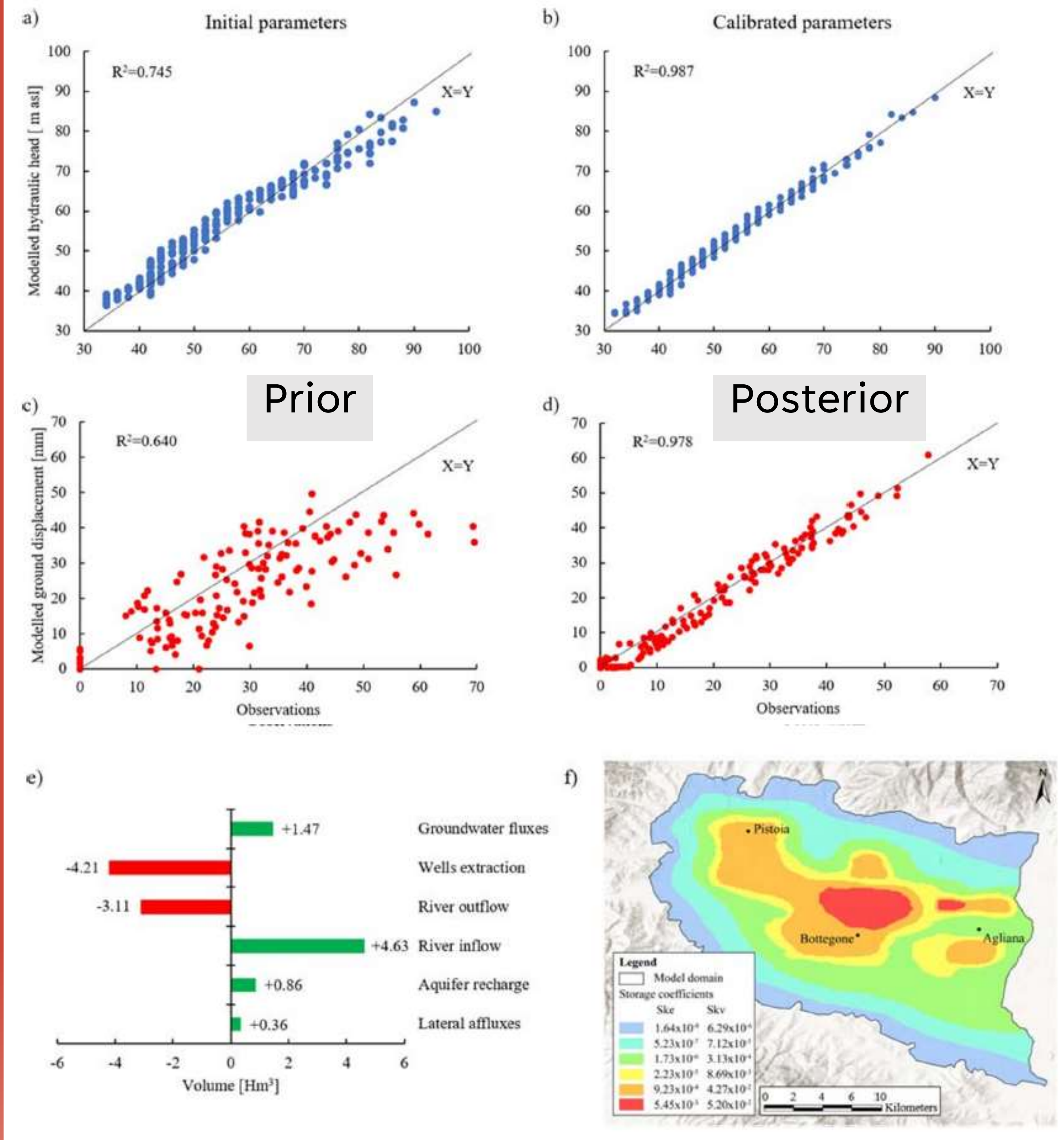
$$\widehat{\boldsymbol{x}}_{\boldsymbol{k}}^+ = \widehat{\boldsymbol{x}}_{\boldsymbol{k}}^- + \boldsymbol{K}_k(\boldsymbol{z}_k - \boldsymbol{H}\widehat{\boldsymbol{x}}_{\boldsymbol{k}}^-)$$

... or the dumped, minimum length, weighted Backus-Gilbert Generalized Inverse

$$\widehat{\boldsymbol{\theta}} = \boldsymbol{\theta}_0 + (\boldsymbol{G}^T\boldsymbol{W}_d\boldsymbol{G} + \varepsilon^2\boldsymbol{W}_x)^{-1}\left(\boldsymbol{G}^T\boldsymbol{W}_d(\boldsymbol{z} - \boldsymbol{G}\boldsymbol{\theta}_0)\right)$$

In **non-linear ill-posed problems**, $\boldsymbol{F}(\boldsymbol{x},\boldsymbol{\theta}) = \boldsymbol{z}$ a Variational (iterative) approach is used to find the minimum of:

$$\|(\boldsymbol{F}(\boldsymbol{x},\boldsymbol{\theta}) - \boldsymbol{z})^T\boldsymbol{P}(\boldsymbol{F}(\boldsymbol{x},\boldsymbol{\theta}) - \boldsymbol{z})\| + \|(\boldsymbol{\theta} - \boldsymbol{\theta}_0)^T\boldsymbol{Q}(\boldsymbol{\theta} - \boldsymbol{\theta}_0)\|$$

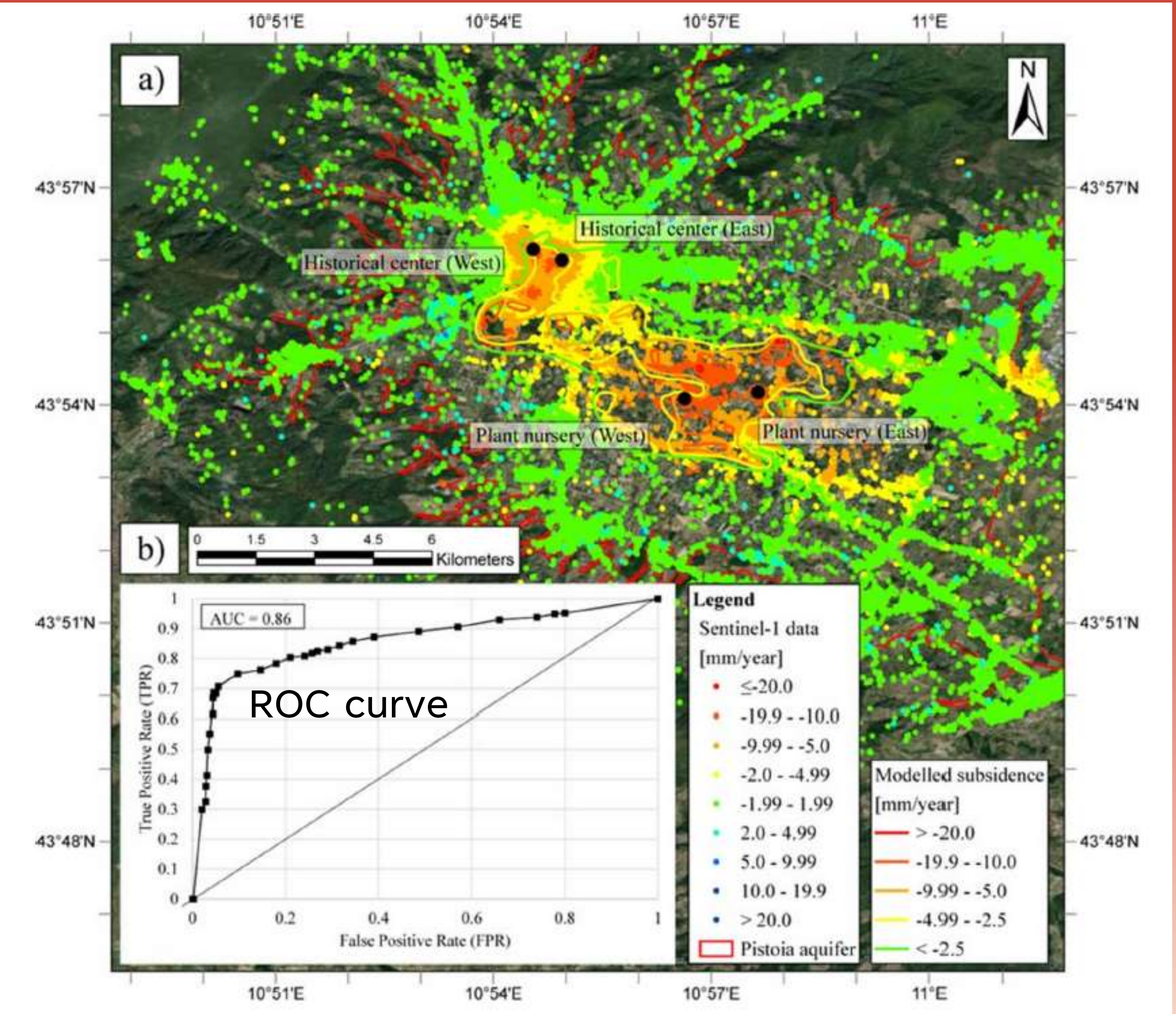


Water table elevation

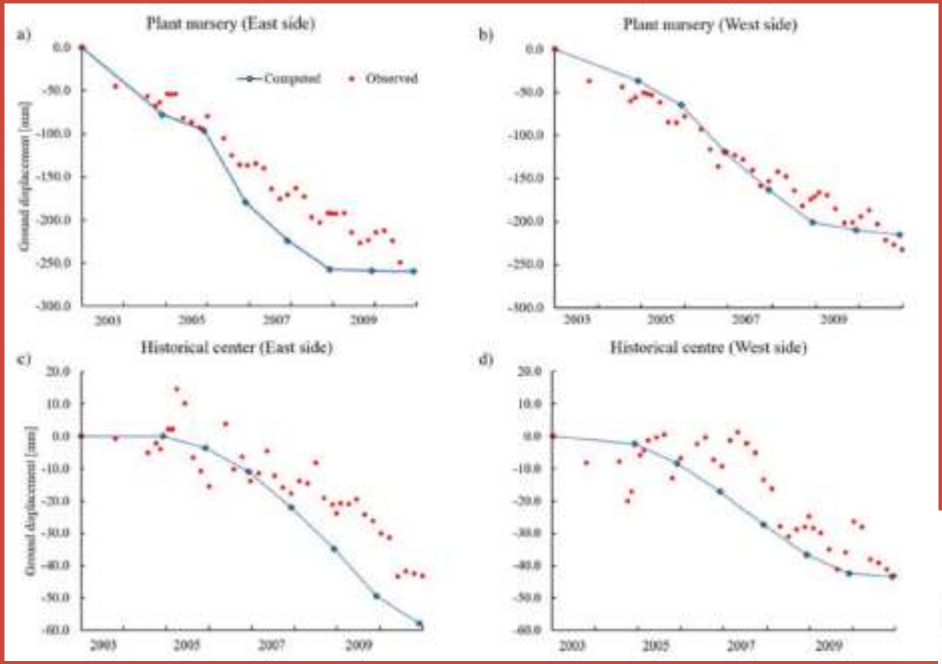
Surface displacement

Posterior storage coefficient field

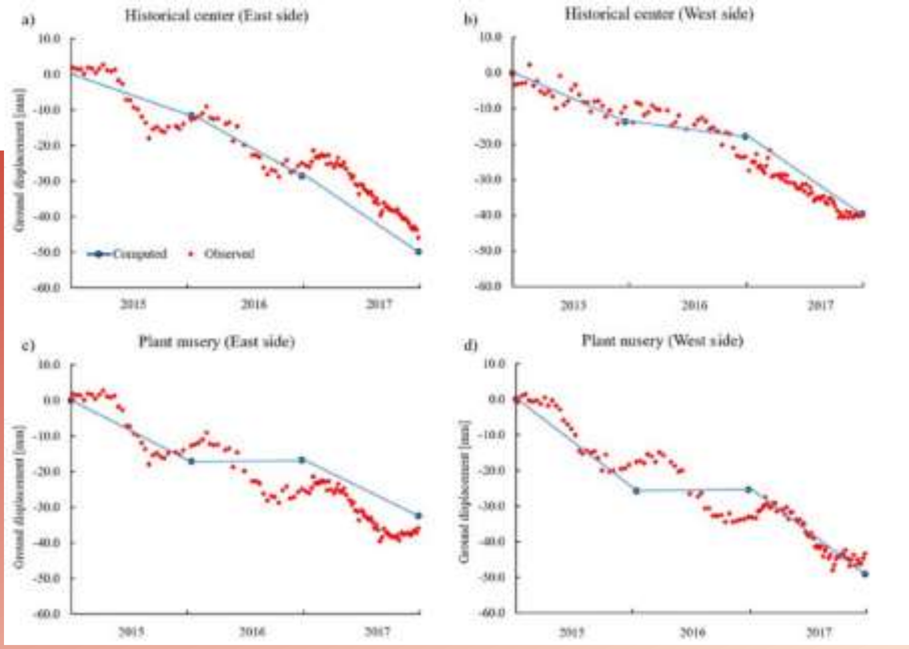
How much the improved state estimation 'explains' the measured ground displacement spatial pattern?



How much the improved state estimation 'explains' the measured ground displacement dynamics?



ENVISAT 2003-2010



Sentinel-1 2015-2017

Means, motive, and opportunity

Means, motive, and opportunity is a popular cultural summation of the three aspects of a crime needed to convince a jury of guilt in a criminal proceeding.

Motive: Aquifer deformation physics.

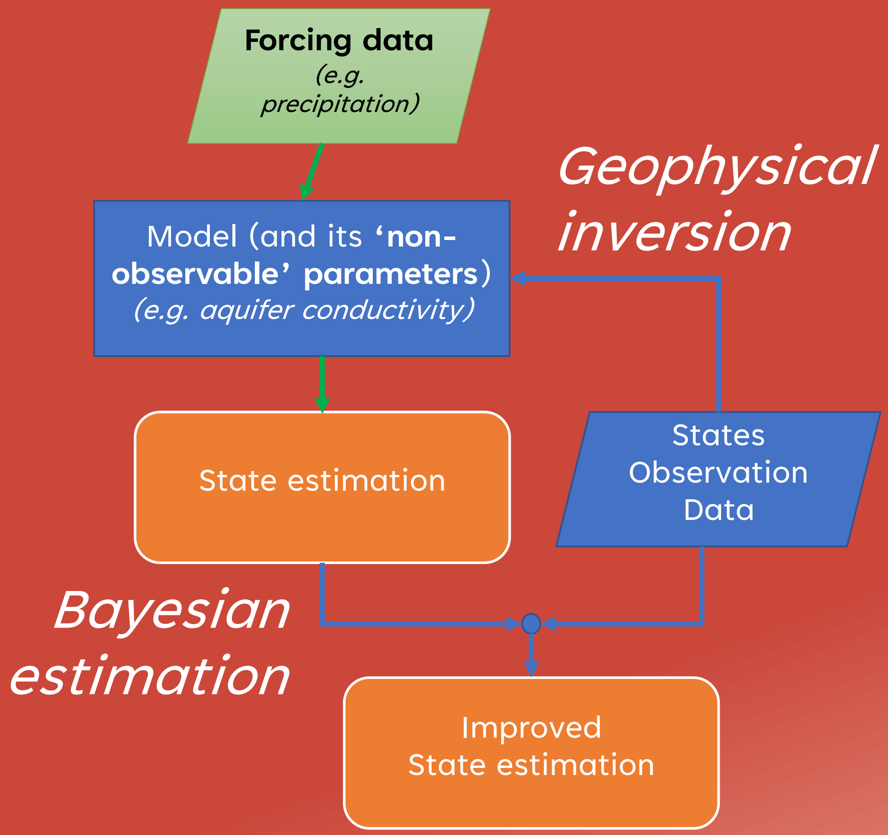
Opportunity: Spatial and temporal coherence of subsidence and groundwater dynamics.

Means: The geophysical inversion demonstrated that the groundwater abstraction had the means to induce the measured amounts of surface deformation.

Alibi ?

4. Geophysical inversion and state estimation together

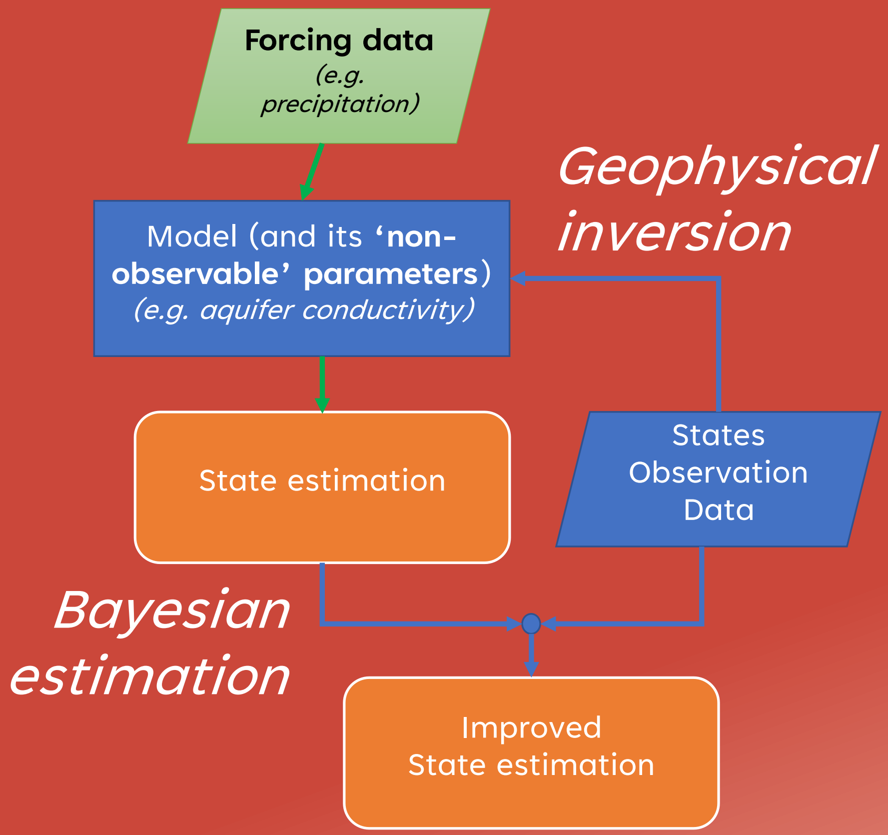
(for the reasons recalled in the introduction, many D.A. solutions in hydrology need to address both).



Common approaches:

- *Filter augmentation*
- ***Variational assimilation with an adjoint***

4. Geophysical inversion and state estimation together



- An example of satellite Var.D.A. for mapping surface, soil moisture-controlled, energy balance and soil/atmosphere interaction.
- An example of river flow Var.D.A. for flood prediction.


AGU PUBLICATIONS

Water Resources Research

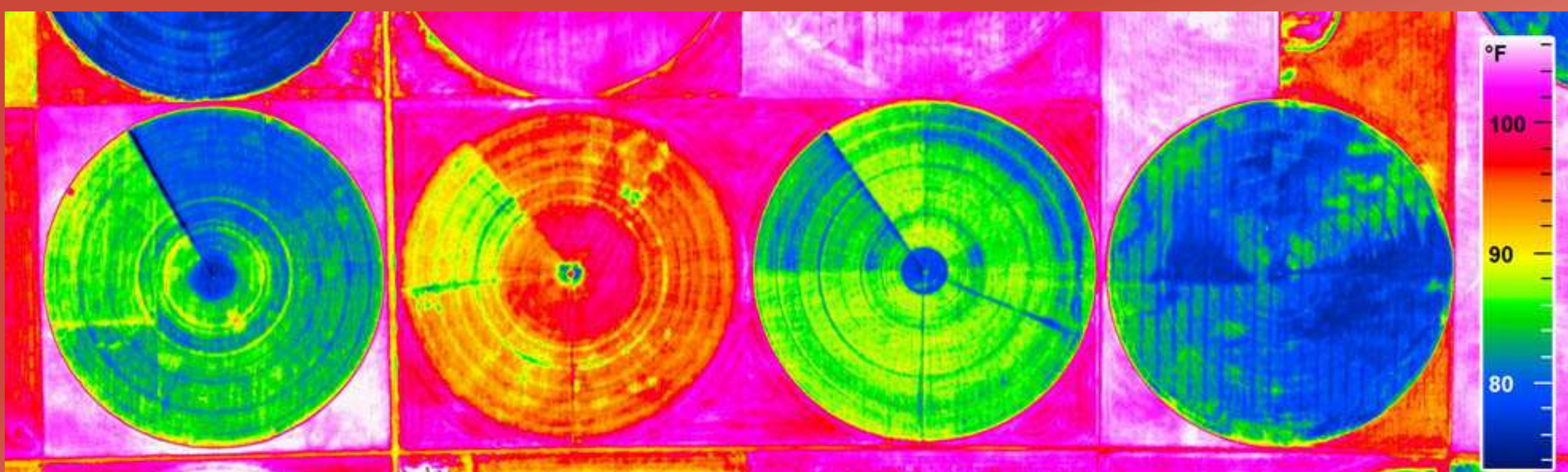
RESEARCH ARTICLE
10.1002/2013WR014573

Special Section:
Patterns in Soil-Vegetation-Atmosphere Systems:

Coupled estimation of surface heat fluxes and vegetation dynamics from remotely sensed land surface temperature and fraction of photosynthetically active radiation

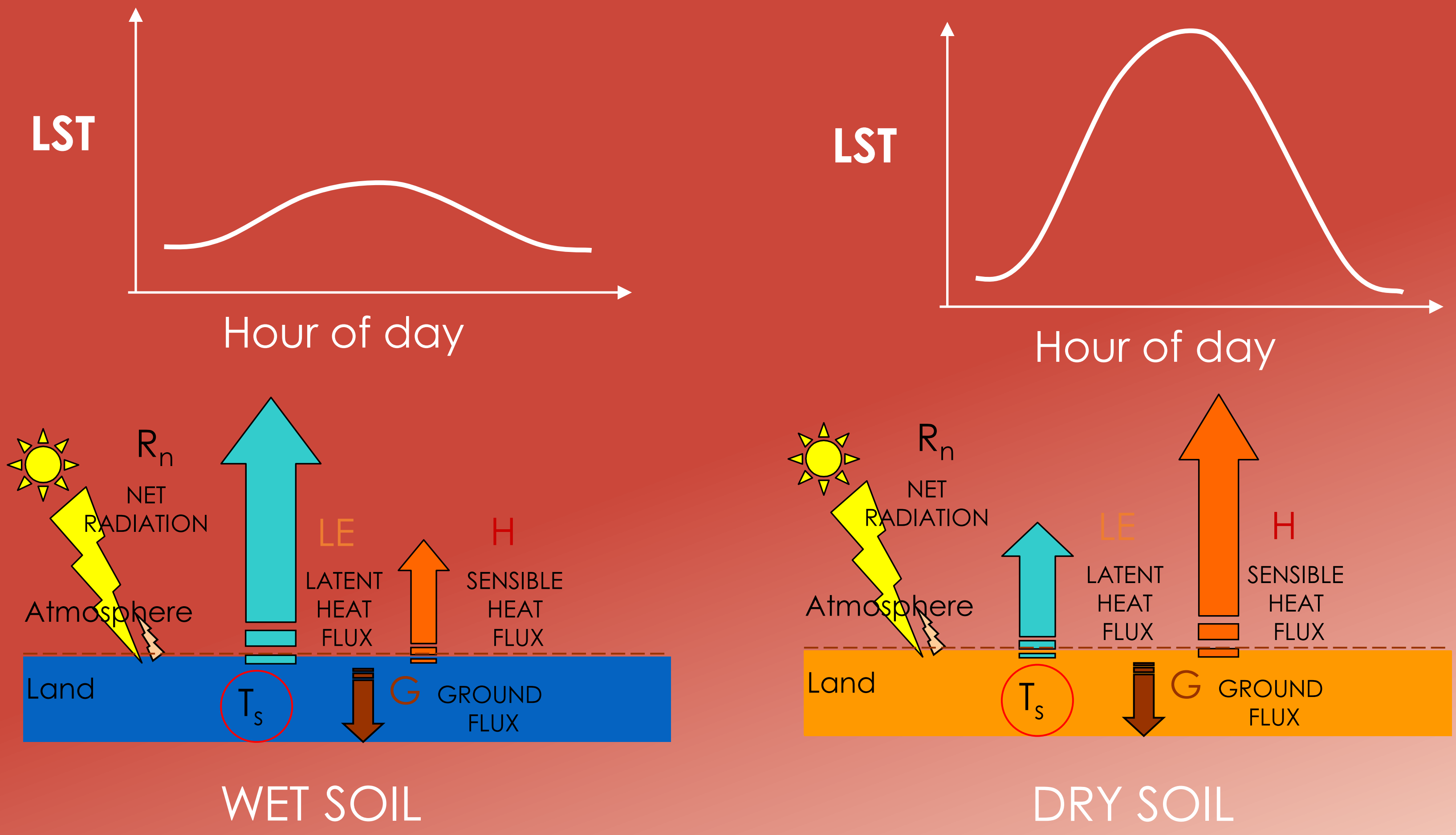
S. M. Bateni[1], D. Entekhabi[2], S. Margulis[3], F. Castelli[4], and L. Kergoat[5]

Guiding principle:
Land Surface Temperature (LST) diurnal dynamics bears information on soil moisture and its control on surface turbulent fluxes



Thermal imaging of irrigated crops

Soil moisture controls the partinioning of available surface energy (Net Radiation minus Ground Heat Flux) among Turbulent Latent (evapotranspiration) and Sensible Heat Flux



Multivariate 1D-Var

State estimation as a time-continuous initial value problem

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{F}(\boldsymbol{x}, \boldsymbol{\theta}, \boldsymbol{u}) + \boldsymbol{w} \quad t \in (t_0, t_1) \quad \boldsymbol{x}(t_0) = \boldsymbol{x}_0$$
$$\boldsymbol{z} = \boldsymbol{G}(\boldsymbol{x}) + \boldsymbol{v}$$

$\boldsymbol{\theta}$ is the 'non-observable' parameters set

Global penalty function with adjoined model constraint through Lagrange multipliers

Assimilate a number of observations $\boldsymbol{z}_k = \boldsymbol{z}(t_k)$, $t_k \in (t_0, t_1)$, $k = 1, \ldots, N$ through the minimization of:

$$J(\boldsymbol{x}, \boldsymbol{\theta}, \boldsymbol{\lambda} | \boldsymbol{z}_k) = (\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}})\boldsymbol{\Gamma}_\theta(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}})^T + \sum_k (\boldsymbol{G}(\boldsymbol{x}) - \boldsymbol{z}_k)\boldsymbol{\Gamma}_z(\boldsymbol{G}(\boldsymbol{x}) - \boldsymbol{z}_k)^T +$$
$$+ \int_{t_0}^{t_1} \boldsymbol{\lambda} \left[\frac{d\boldsymbol{x}}{dt} - \boldsymbol{F}(\boldsymbol{x}, \boldsymbol{\theta}, \boldsymbol{u})\right] dt + i.c.$$

Global minimization by setting independent variates to zero

$$\boldsymbol{\delta J}(\boldsymbol{x}, \boldsymbol{\theta}, \boldsymbol{\lambda} | \boldsymbol{z}_k) = 0 \Rightarrow \quad {}^{\partial \boldsymbol{J}}\!/_{\partial \boldsymbol{x}} = 0 \quad {}^{\partial \boldsymbol{J}}\!/_{\partial \boldsymbol{\theta}} = 0 \quad {}^{\partial \boldsymbol{J}}\!/_{\partial \boldsymbol{\lambda}} = 0$$

Euler-Lagrange Equations

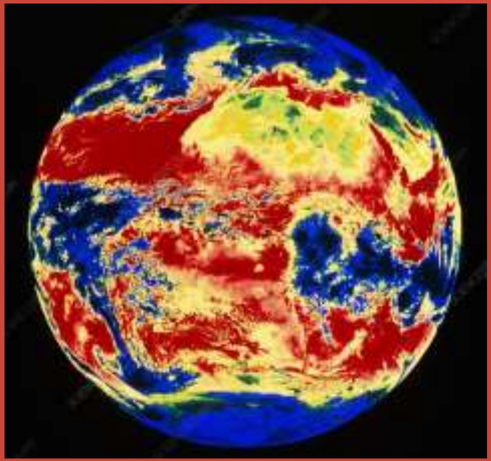
Forward Model $\frac{d\boldsymbol{x}}{dt} = \boldsymbol{F}(\boldsymbol{x}, \boldsymbol{\theta}, \boldsymbol{u})$

Backward Adjoint Model $\frac{d\boldsymbol{\lambda}}{dt} = -\boldsymbol{\lambda}\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{x}} - \boldsymbol{\Gamma}_z(\boldsymbol{G}(\boldsymbol{x}) - \boldsymbol{z}_k)\delta(t - t_k)$

Parameters Update $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}} + \boldsymbol{\Gamma}_\theta^{-1}\int_{t_0}^{t_1} \boldsymbol{\lambda}\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{\theta}}dt$

Iterate until $\boldsymbol{\lambda} \rightarrow 0$

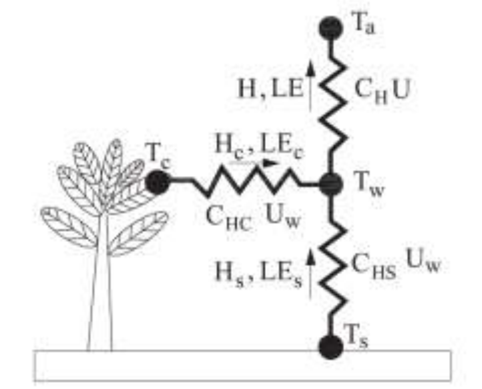
Assimilation of remotely sensed Land Surface Temperature (**LST**) and Fraction of Photosynthetically Active Radiation absorbed by vegetation (**FPAR**) from MeteosatSG-SEVIRI (geostationary, 15’ revisit, 3km resolution)



Forward model

$$c\frac{\partial T_s(z,t)}{\partial t}=\frac{\partial}{\partial z}\left(p\frac{\partial T_s(z,t)}{\partial z}\right)$$

Heat diffusion into the soil with b.c. provided by surface energy balance



$$\frac{dLAI}{dt}=1.52a_a c_g(1-g_a)x\frac{m'_a}{m_i-m_a}E_c-(m_s+d_T)LAI$$

Vegetation dynamics (Leaf Area Index)

Retrieved daily states and parameters

FPAR $\mathbf{FPAR}=1-\exp\left(-k_e\mathbf{LAI}\right)$

Evaporative fractions

$$EF_c=\frac{LE_c}{LE_c+H_c}$$

$$EF_s=\frac{LE_s}{LE_s+H_s}$$

Multivariate penalty function

$$J(\mathbf{T}, \mathbf{LAI}, \mathbf{R}, \mathbf{EF}_S, \mathbf{EF}_C, \mathbf{c_g}, \boldsymbol{\Lambda 1}, \boldsymbol{\Lambda 2}) =$$

$$+\sum_{i=1}^{N}\int_{\tau_0}^{\tau_1}[\mathbf{T}_i(0,t)-\mathbf{T}_{\mathbf{obs,i}}(0,t)]^T\mathbf{K}_T^{-1}[\mathbf{T}_i(0,t)-\mathbf{T}_{\mathbf{obs,i}}(0,t)]dt$$

$$+\sum_{i=1}^{N}[\mathbf{FPAR}_i-\mathbf{FPAR}_{obs,i}]^T\mathbf{K}_{FPAR}^{-1}[\mathbf{FPAR}_i-\mathbf{FPAR}_{obs,i}]+(\mathbf{R}-\mathbf{R}')^T\mathbf{K}_R^{-1}(\mathbf{R}-\mathbf{R}')$$

$$+\sum_{i=1}^{N}(\mathbf{EF}_{S_i}-\mathbf{EF}'_{S_i})^T\mathbf{K}_{EFs}^{-1}(\mathbf{EF}_{S_i}-\mathbf{EF}'_{S_i})+\sum_{i=1}^{N}(\mathbf{EF}_{C_i}-\mathbf{EF}'_{C_i})^T\mathbf{K}_{EFc}^{-1}(\mathbf{EF}_{C_i}-\mathbf{EF}'_{C_i})$$

$$+(\mathbf{c}_g\text{-}\mathbf{c}'_g)^T K_{C_g}^{-1}(\mathbf{c}_g\text{-}\mathbf{c}'_g)+\sum_{i=1}^{N}\int_{\tau_0}^{\tau_1}\int_0^l \boldsymbol{\Lambda 1}_i\left[c\frac{\partial \mathbf{Ts_i}(z,t)}{\partial t}-\frac{\partial}{\partial z}\left(p\frac{\partial \mathbf{Ts_i}(z,t)}{\partial z}\right)\right]dzdt$$

$$+\int_{i=1}^{N}\boldsymbol{\Lambda 2}_i\left(\frac{d\mathbf{LAI}_i}{dt}-\frac{1.52a_a(1-g_a)xm'_aE_c}{m_i-m_a}c_g+(m_s+d_T)\mathbf{LAI}_i\right)dt$$

Figure 2. Graphical location of the Gourma mesoscale site (surrounded by red lines) in West Africa. Figure adapted from http://www.ecmwf.int/research/ESA_projects/SMOS/calval/smos-amma_index.html.

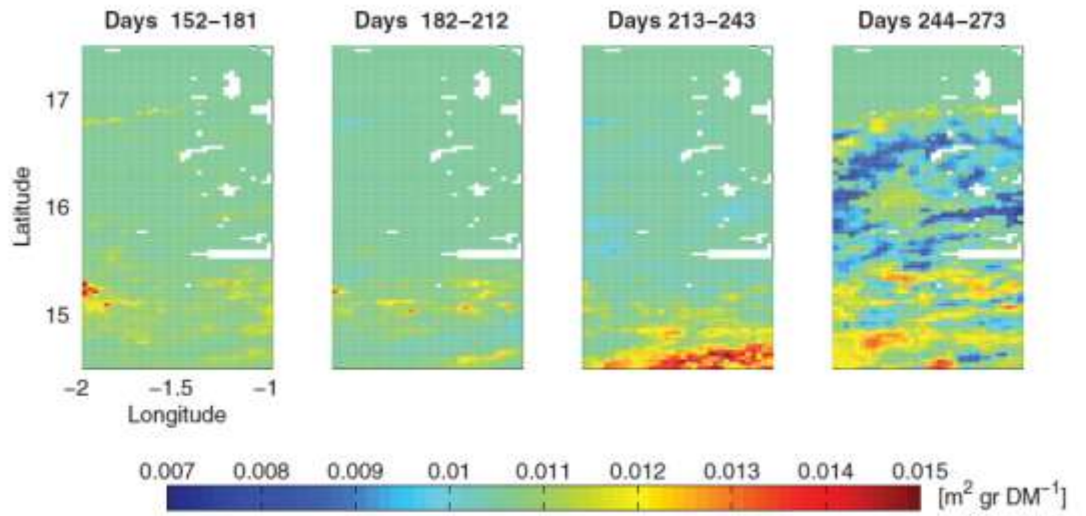


Figure 4. Retrieved values for the specific leaf area of the green biomass (c_g) over four periods.

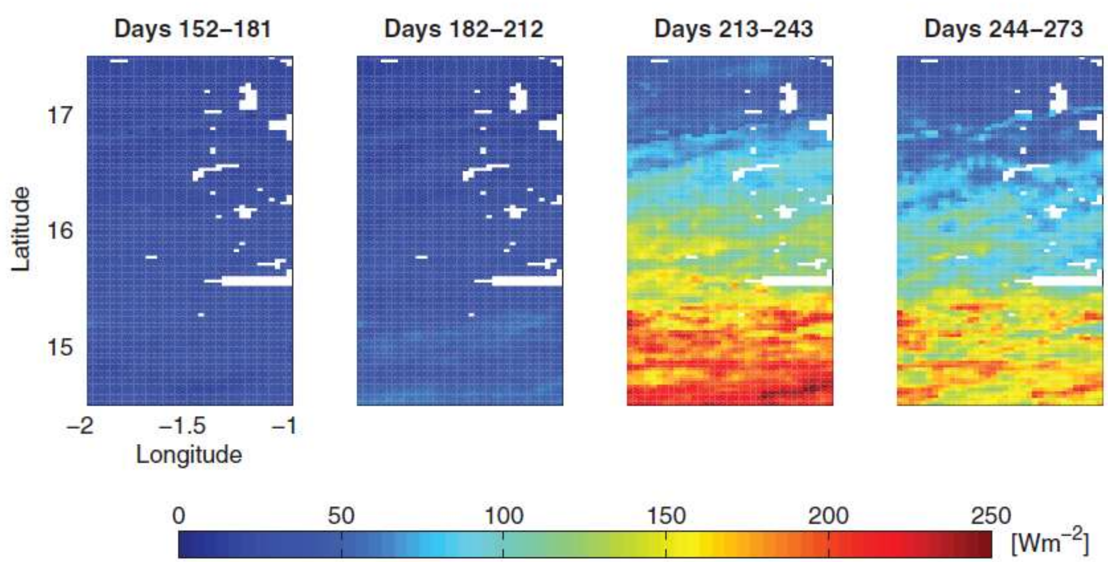
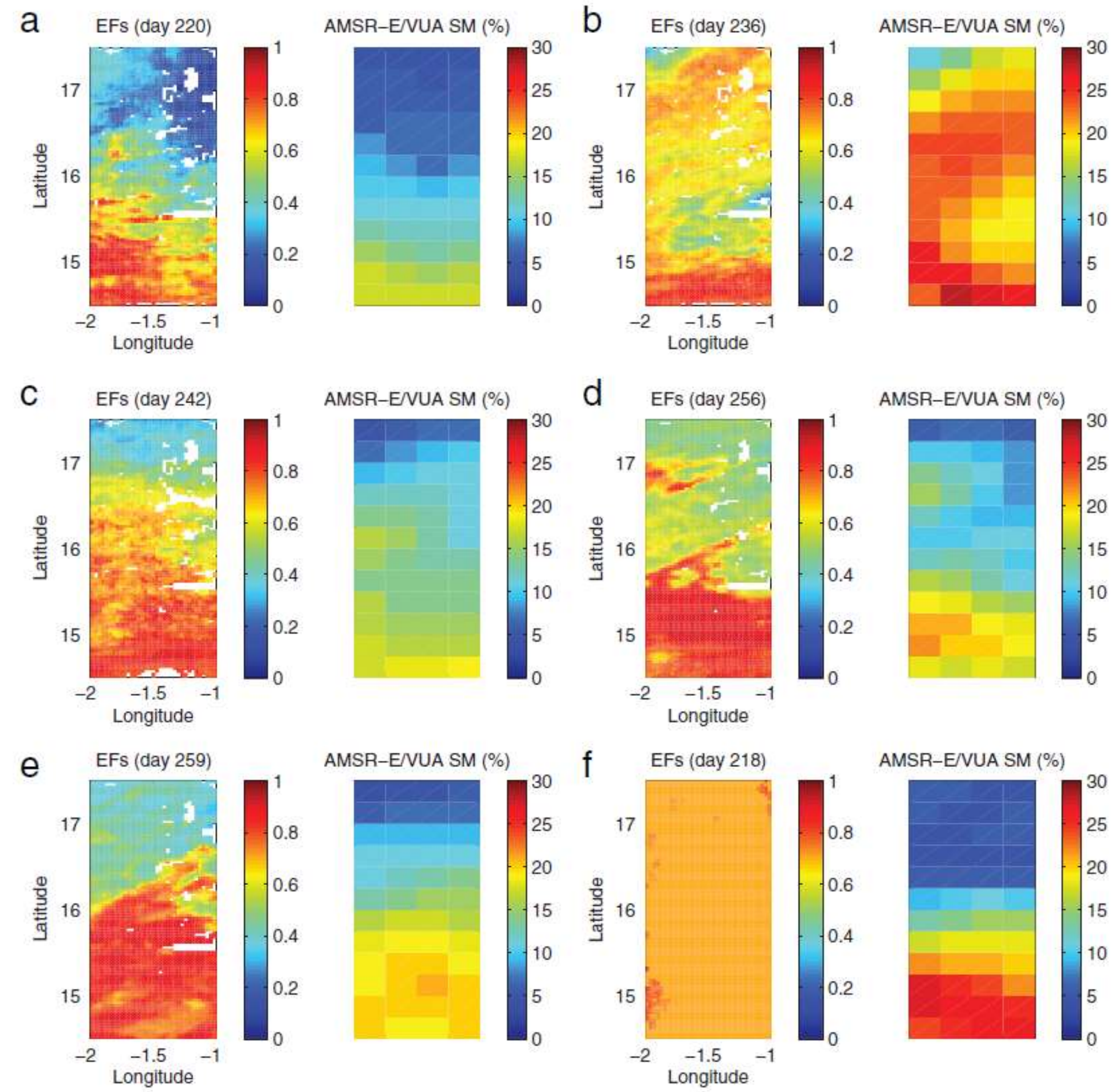


Figure 11. Maps of average daytime latent heat flux (Wm^{-2}).

Figure 2. Graphical location of the Gourma mesoscale site (surrounded by red lines) in West Africa. Figure adapted from http://www.ecmwf.int/research/ESA_projects/SMOS/calval/smos-amma_index.html.



Retrieved Evaporative Fractions vs. Satellite microwave (AMSR-E) soil moisture estimates

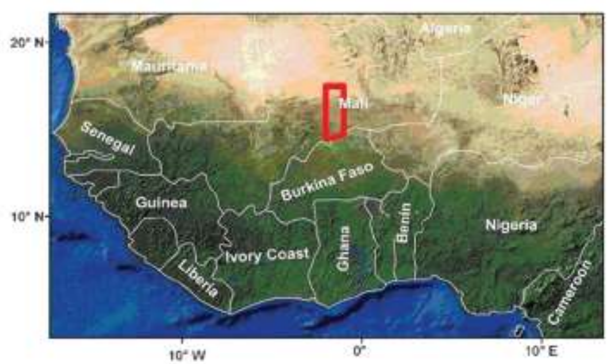
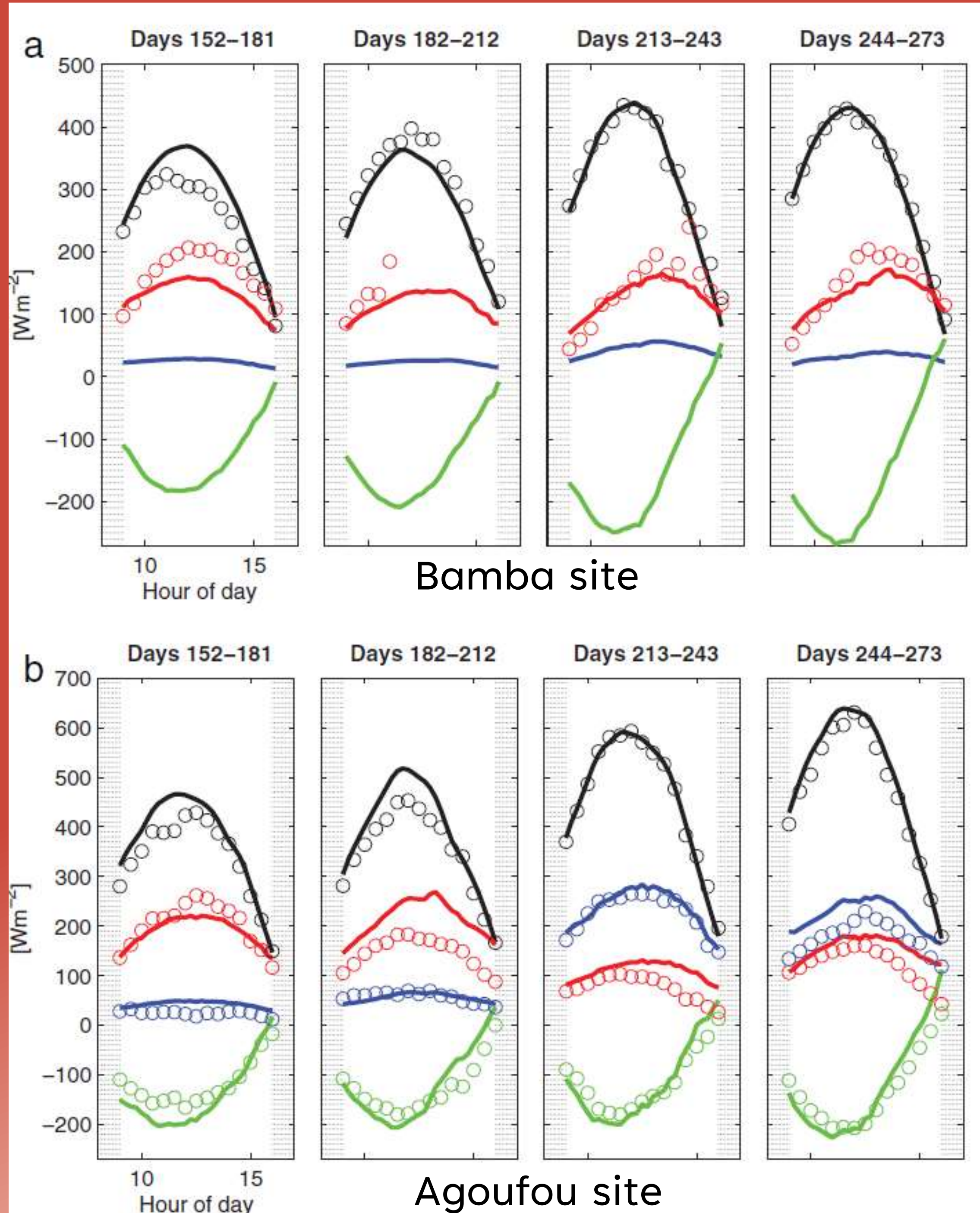


Figure 2. Graphical location of the Gourma mesoscale site (surrounded by red lines) in West Africa. Figure adapted from http://www.ecmwf.int/research/ESA_projects/SMOS/calval/smos-amma_index.html.



Retrieved diurnal flux cycles
vs.
ground flux tower data

Net Radiation
Sensible Heat
Latent Heat
Ground Heat

Water Resources Research

RESEARCH ARTICLE
10.1002/2016WR019208

Variational assimilation of streamflow data in distributed flood forecasting

Giulia Ercolani[1] and Fabio Castelli[1]

Key Points:
• Variational assimilation of streamflow

5 River flow gauges along the main channel

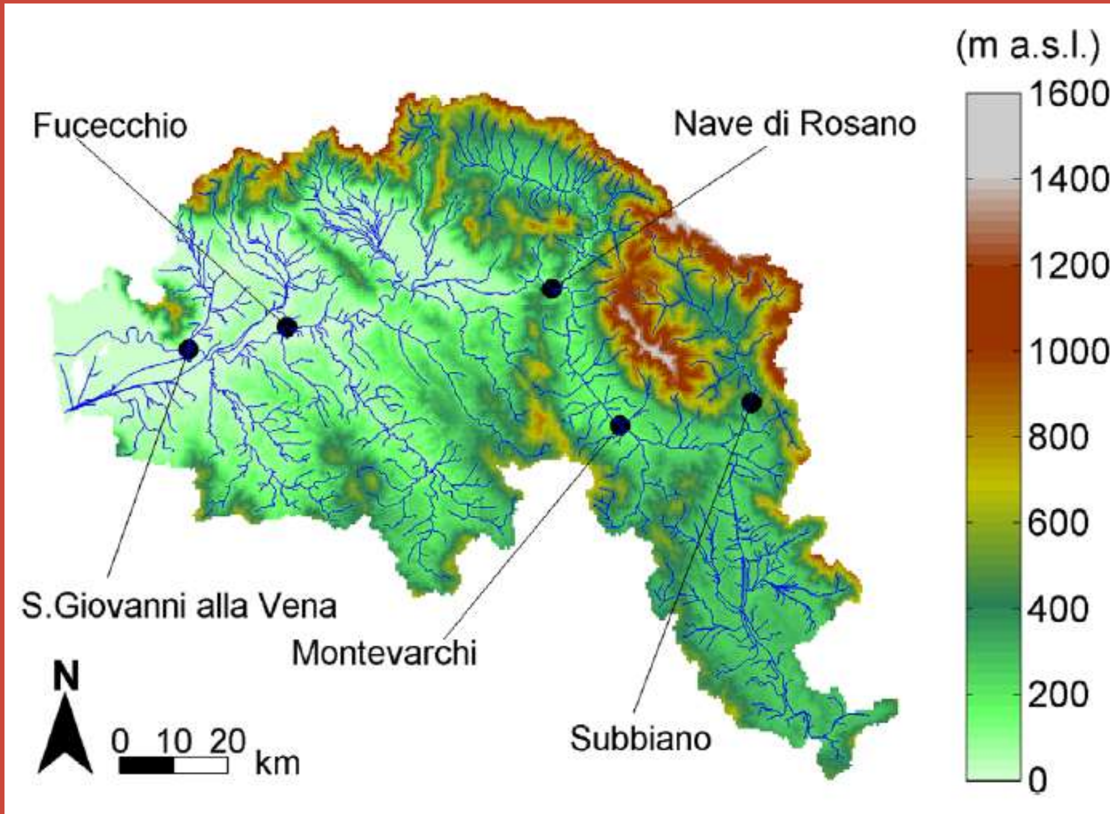


Figure 2. Digital Elevation Model (DEM) and river network of Arno basin used in the simulations. Black circles indicate the available flow measurement stations. DEM is plotted at the resolution employed in the simulations, i.e., 500 m.

Operational flood forecasting with a distributed, soil moisture accounting and river routing model

Total area: 9,374 km^2

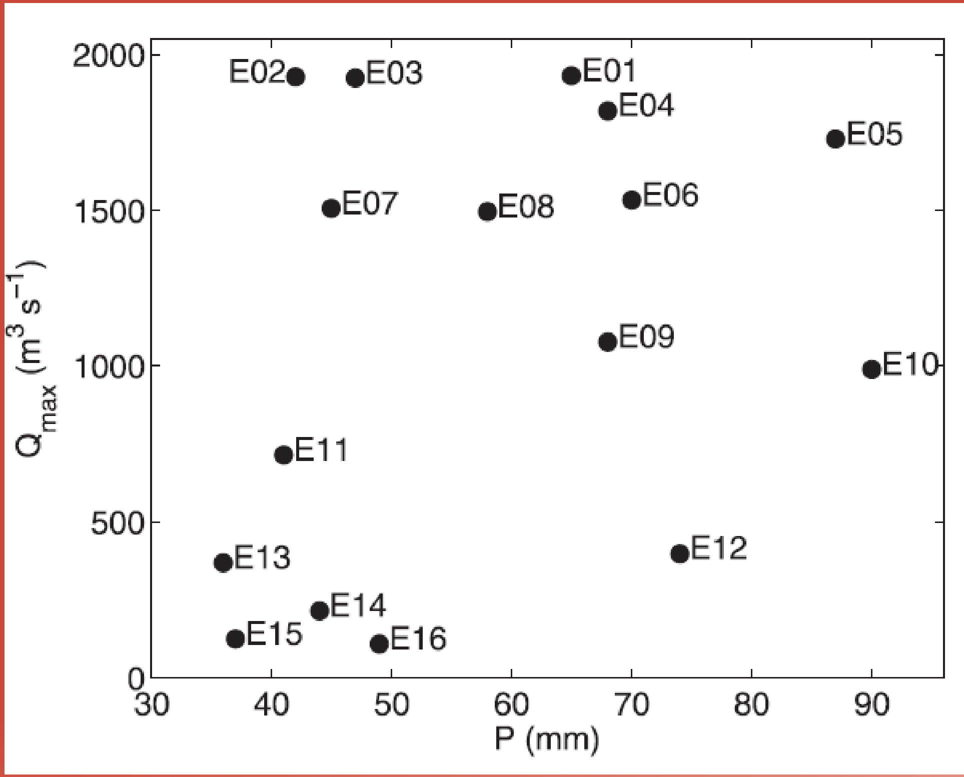
Cell size: 0.25 km^2

Total river network length: 3,692 km

Drainage density: 0.394 km^{-1}

Hindcast experiments with 16 recorded high-flow events, assuming rainfall can be perfectly predicted

Recorded peak flow at downstream gauge



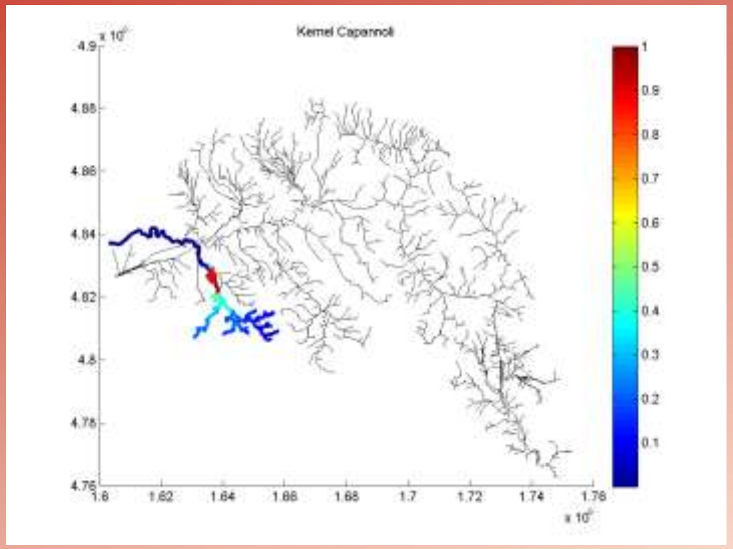
Basin-average cumulative precipitation

Predicting floods is not just predicting rainfall!

One key simplification with respect to other 'fluid flow' problems:

Knowing the drainage structure, the problem may be reduced to a system of coupled ODEs

Multivariate 1D-VAR instead of 3D or 4D-VAR



$$K(\hat{x}, \hat{x}_i) = exp\left(-\frac{\alpha_{up,down}}{\Delta t}\int_{\hat{x}_i}^{\hat{x}} \frac{ds}{C(s)}\right)$$

Dendritic Asymmetric Assimilation Kernel
(local updates need to efficiently propagate upstream)

Penalty function

$$J=\frac{1}{t_1-t_0}\int_{t_0}^{t_1}\left[\left(\mathbf{Q}(t)-\mathbf{Q}^{obs}(t)\right)^T\frac{\mathbf{K_Q}}{2}\left(\mathbf{Q}(t)-\mathbf{Q}^{obs}(t)\right)\right]dt$$
$$+\left[\left(\mathbf{Q}(t_1)-\mathbf{Q}^{obs}(t_1)\right)^T\frac{\mathbf{K_Q}}{2}\left(\mathbf{Q}(t_1)-\mathbf{Q}^{obs}(t_1)\right)\right]$$
$$+\left[\left(\mathbf{Q}(t_0)-\mathbf{Q}'(t_0)\right)^T\frac{\mathbf{K_{Q_0}}}{2}\left(\mathbf{Q}(t_0)-\mathbf{Q}'(t_0)\right)\right]$$
$$+\frac{1}{t_1-t_0}\int_{t_0}^{t_1}\left[\left(\mathbf{q_L}(t)-\mathbf{q}_L'(t)\right)^T\frac{\mathbf{K_{q_L}}}{2}\left(\mathbf{q_L}(t)-\mathbf{q}_L'(t)\right)\right]dt$$
$$+\int_{t_0}^{t_1}\left[\boldsymbol{\lambda}^T(t)\left(\frac{d\mathbf{Q}(t)}{dt}-\mathbf{F}(\mathbf{A},\mathbf{Q}(t),\mathbf{q_L}(t))\right)\right]dt$$

Euler-Lagrange equations (iterates for $\lambda \rightarrow 0$)

Forward model

$$\frac{d\mathbf{Q}(t)}{dt}=\mathbf{A}[\mathbf{q_L}(t)+\mathbf{UQ}(t)-\mathbf{Q}(t)]=\mathbf{F}(\mathbf{A},\mathbf{Q}(t),\mathbf{q_L}(t))$$

Adjoint backward model

$$\frac{d\boldsymbol{\lambda}(t)}{dt}=\frac{\mathbf{K_Q}^T\left(\mathbf{Q}(t)-\mathbf{Q}^{obs}(t)\right)}{t_1-t_0}-\frac{\partial\mathbf{F}^T(t)}{\partial\mathbf{Q}}\boldsymbol{\lambda}(t)$$

$$\boldsymbol{\lambda}(t_1)=-\mathbf{K_Q}^T\left(\mathbf{Q}(t_1)-\mathbf{Q}^{obs}(t_1)\right)$$

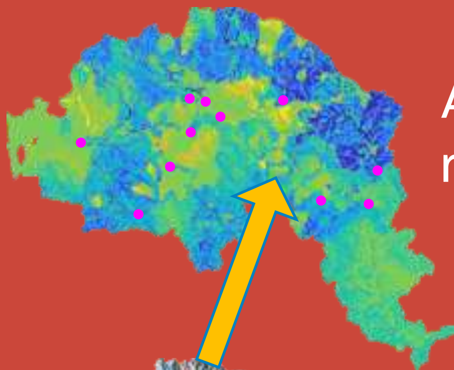
Updates

$$\mathbf{Q}(t_0)=\mathbf{Q}(t_0)'+\left(\mathbf{K_{Q_0}}^T\right)^{-1}\boldsymbol{\lambda}(t_0)$$

$$\mathbf{q_L}(t)=\mathbf{q_L}'(t)+(t_1-t_0)\left(\mathbf{K_{q_L}}^T\right)^{-1}\frac{\partial\mathbf{F}^T(t)}{\partial\mathbf{q_L}}\boldsymbol{\lambda}(t)$$

The sensitivities are coded explicitly, not estimated from discrete variations

D.A. of river flow can improve state estimation at hillslope level (hillslope runoff and initial soil moisture)?



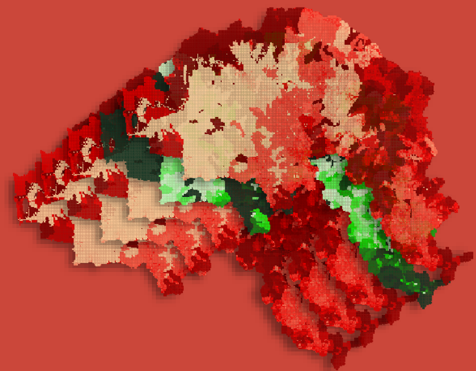
Analysis increment of hillslope runoff

Difficult to adjoin, but at least mass conservation and rainfall distribution need to be maintained

Analysis increment of river flows throught the network

Data of river flow at multiple locations

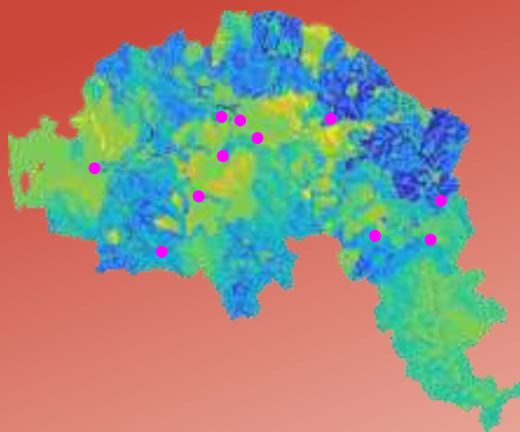
Mixed VAR – Particle Filter

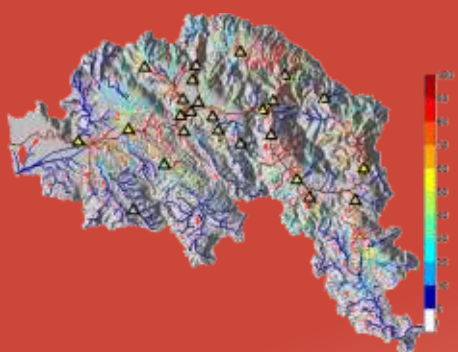


Montecarlo sampling of Antecedent Soil Moisture, rainfall interpolation

Likelihood

Analysis increment of hillslope runoff



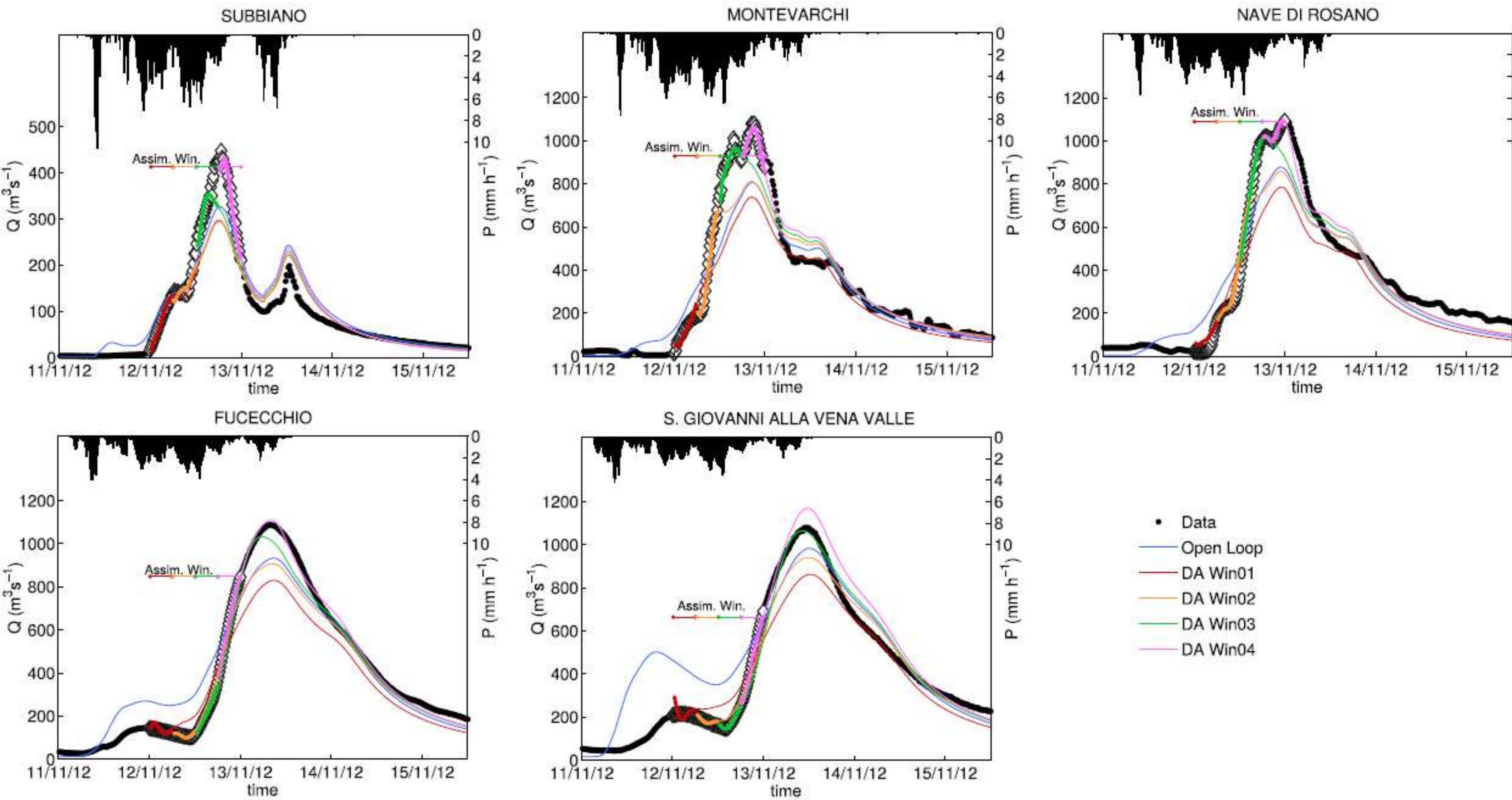


Analysis increment of river flows

Variational assimilation with an adjoint

Hydrometric data

Prediction updates with different D.A. windows



Prediction error reduction with D.A. as a function of prediction lead time

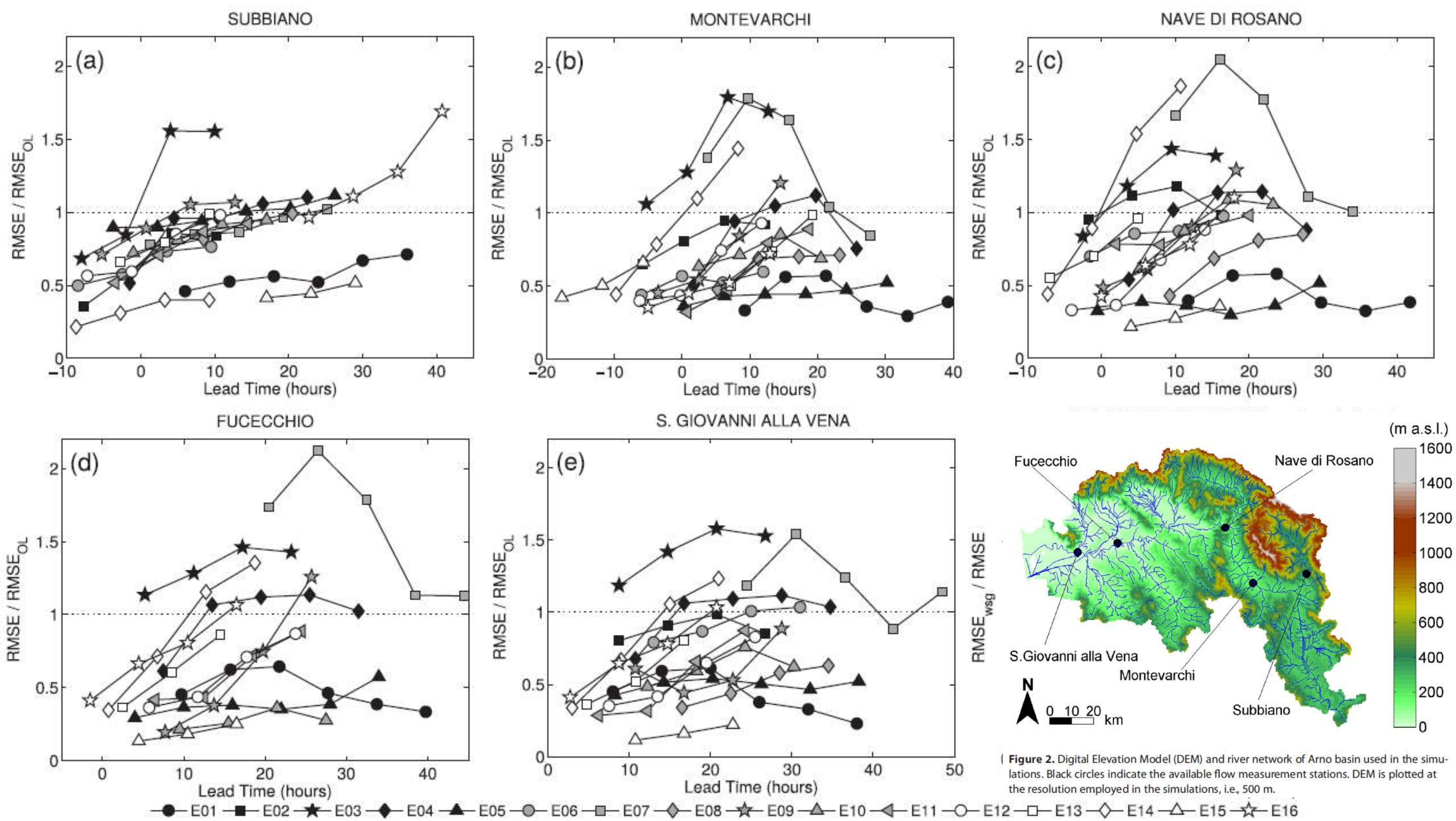


Figure 2. Digital Elevation Model (DEM) and river network of Arno basin used in the simulations. Black circles indicate the available flow measurement stations. DEM is plotted at the resolution employed in the simulations, i.e., 500 m.

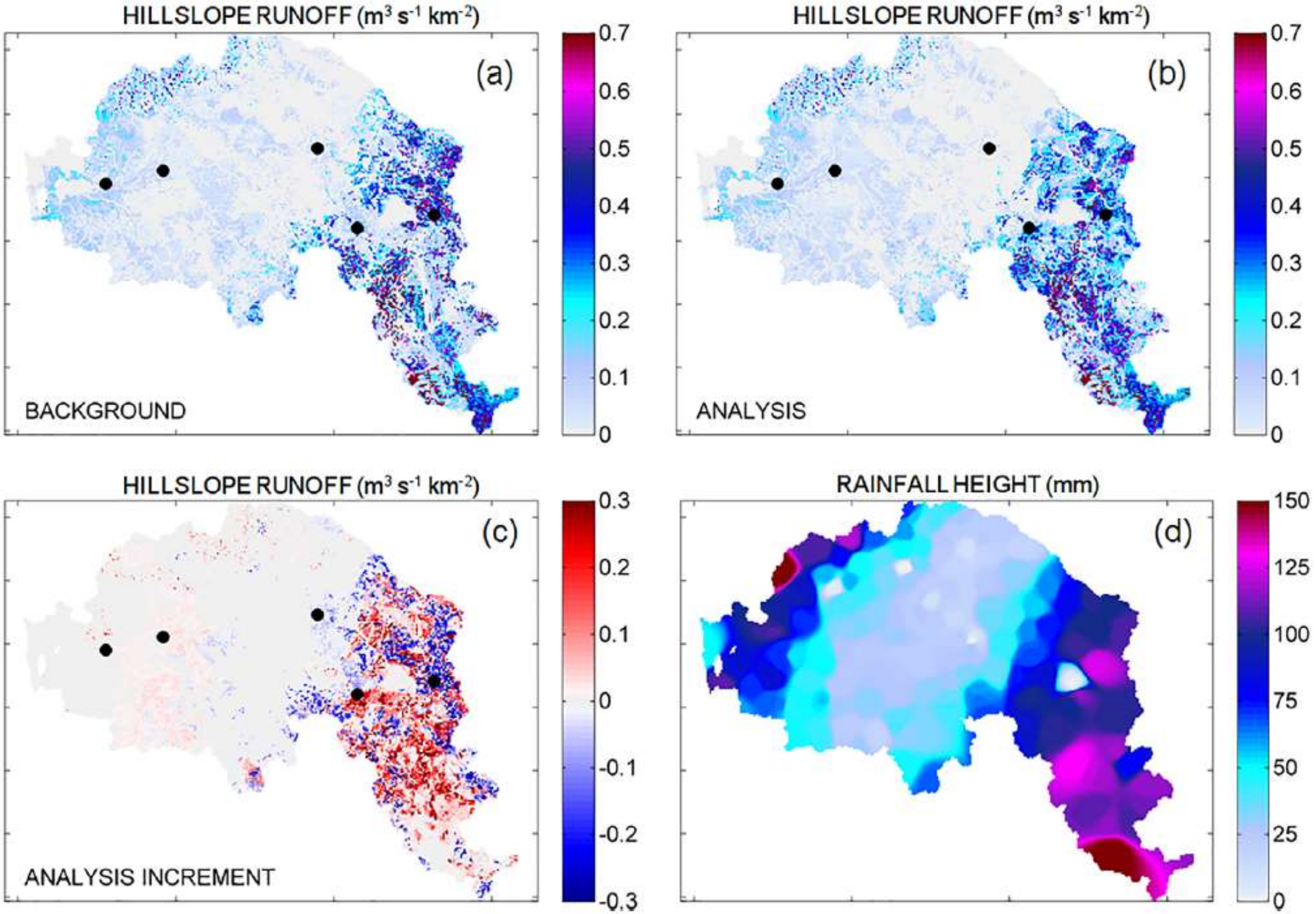


Figure 12. Hillslope runoff ($m^3s^{-1}km^{-2}$) of the high flow event E_{09} at the final instant of the fourth assimilation window. (a) Is the background map (open loop simulation), (b) is the analysis map (data assimilation simulation), and (c) is the analysis increment (analysis minus background). Black circles indicate the position of the gauge stations. (d) Is the cumulative rainfall height (mm) at the same instant.

Concluding remarks: (some) trending topics

Models

Hydrologic Digital Twins: Toward holistic, application independent, hydrological modeling

Data

New types of data: UGC (User Generated Contents), Opportunity, Citizen's science.

D.A. Techniques

Inclusions of AI-Learning algorithms to improve performances of D.A. in physically-based models (e.g. learning from repeated steps in sequential D.A. approaches).