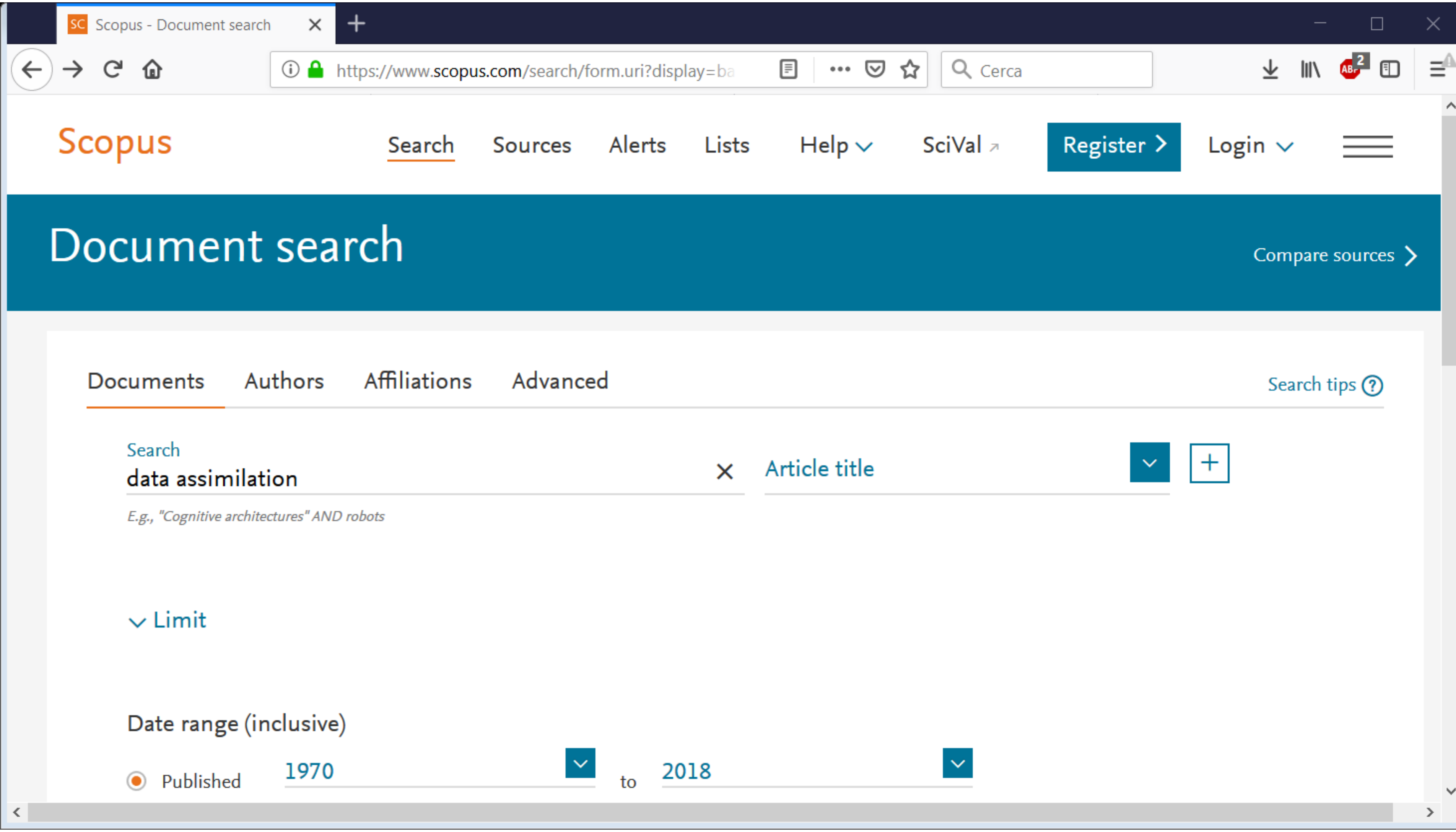
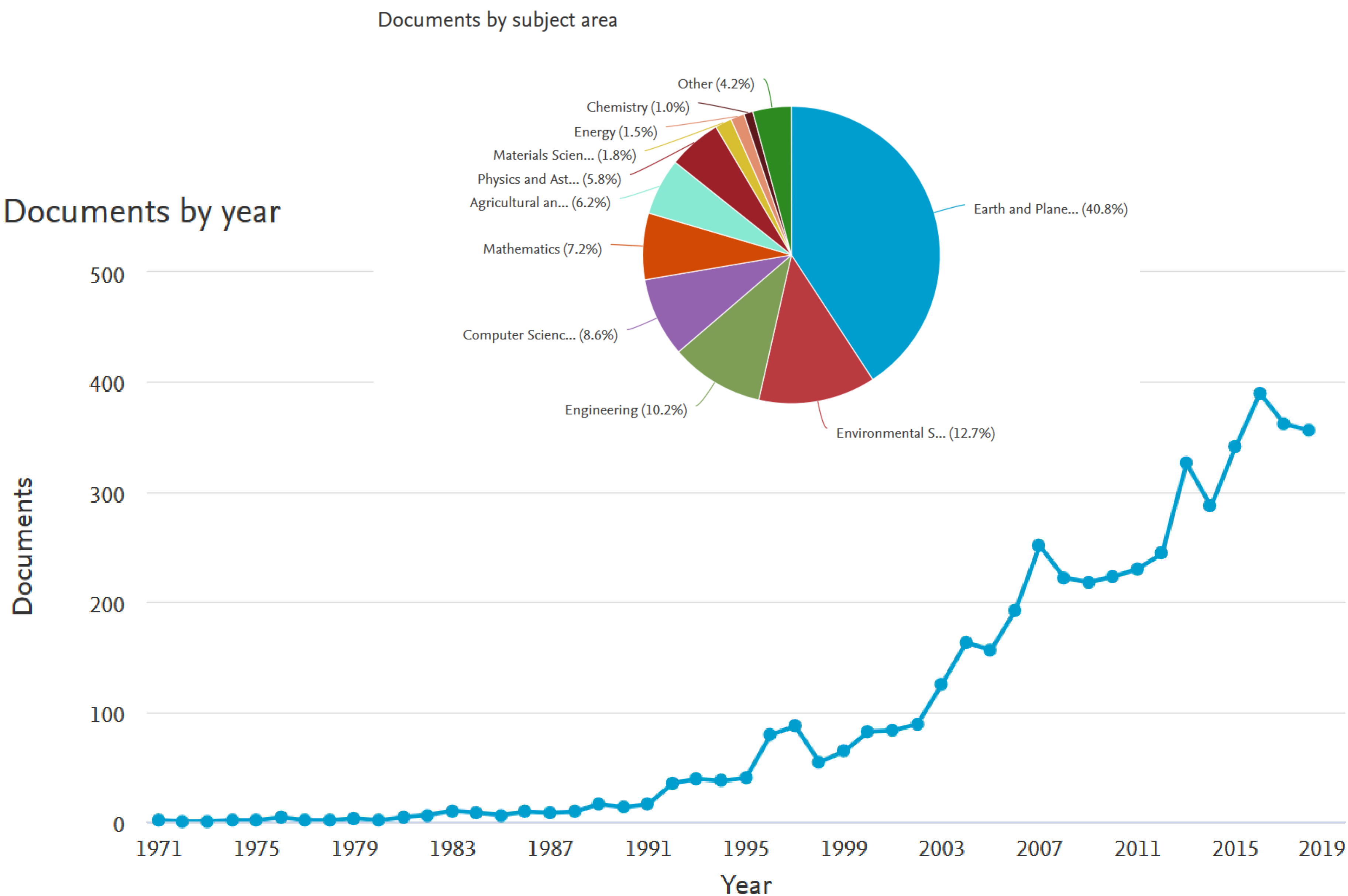
International Doctoral Winter School 2019 -Data Rich Hydrology

Villa Colombella, Perugia - 28.01-01.02, 2019

Remote sensing and data assimilation in hydrology

Fabio.Castelli@unifi.it

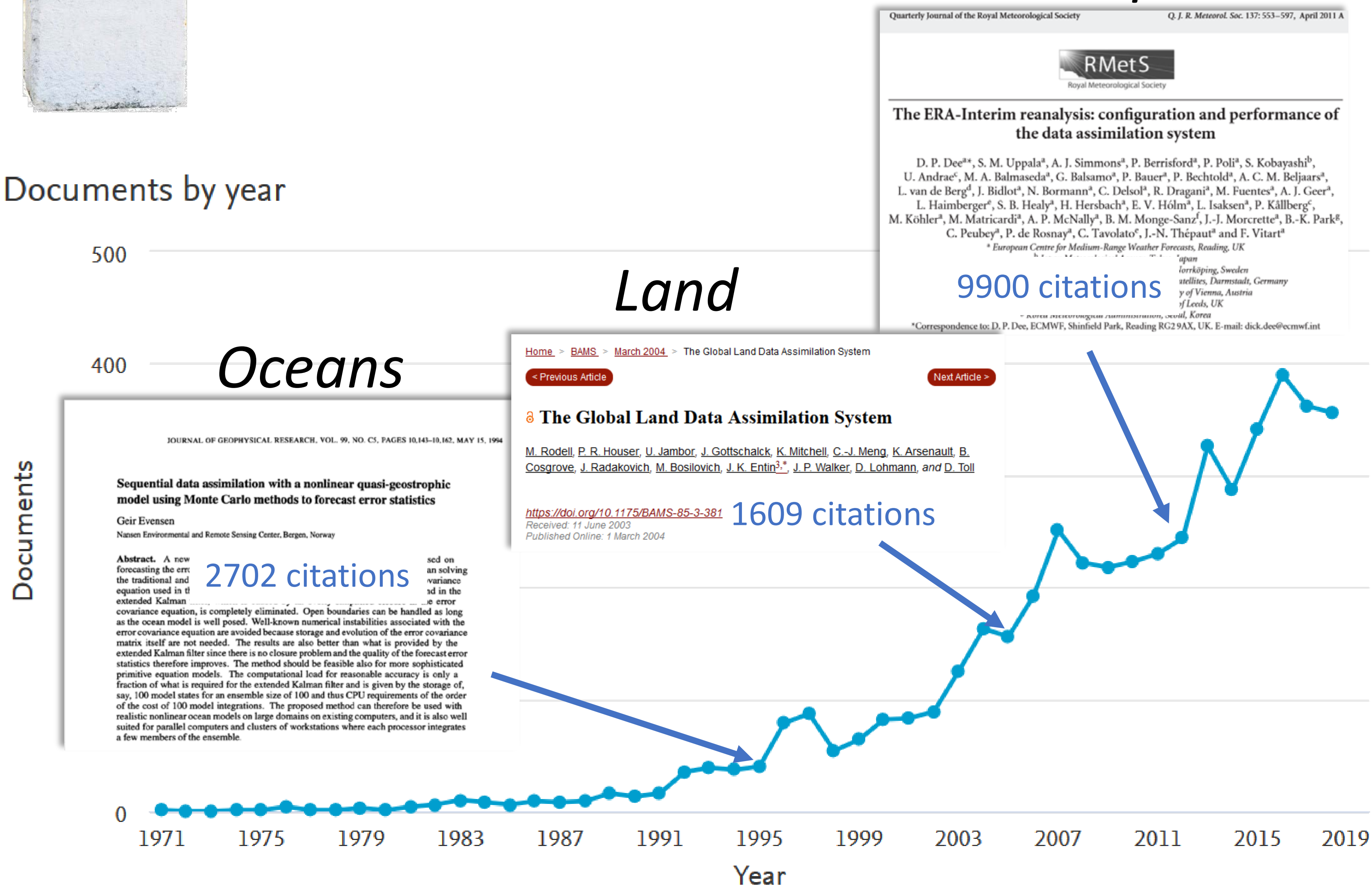

Scopus - Document search
https://www.scopus.com/search/form.uri?display=ba
Cerca
Scopus
Search
Sources
Alerts
Lists
Help
SciVal
Register
Login
Document search
Compare sources
Documents
Authors
Affiliations
Advanced
Search tips
Search
data assimilation
Article title
E.g., "Cognitive architectures" AND robots
Limit
Date range (inclusive)
Published
1970
to
2018



Documents by subject area
Other (4.2%)
Chemistry (1.0%)
Energy (1.5%)
Materials Scien... (1.8%)
Physics and Ast... (5.8%)
Agricultural an... (6.2%)
Mathematics (7.2%)
Computer Scienc... (8.6%)
Engineering (10.2%)
Environmental S... (12.7%)
Earth and Plane... (40.8%)
Documents by year
Documents
500
400
300
200
100
0
1971
1975
1979
1983
1987
1991
1995
1999
2003
2007
2011
2015
2019
Year

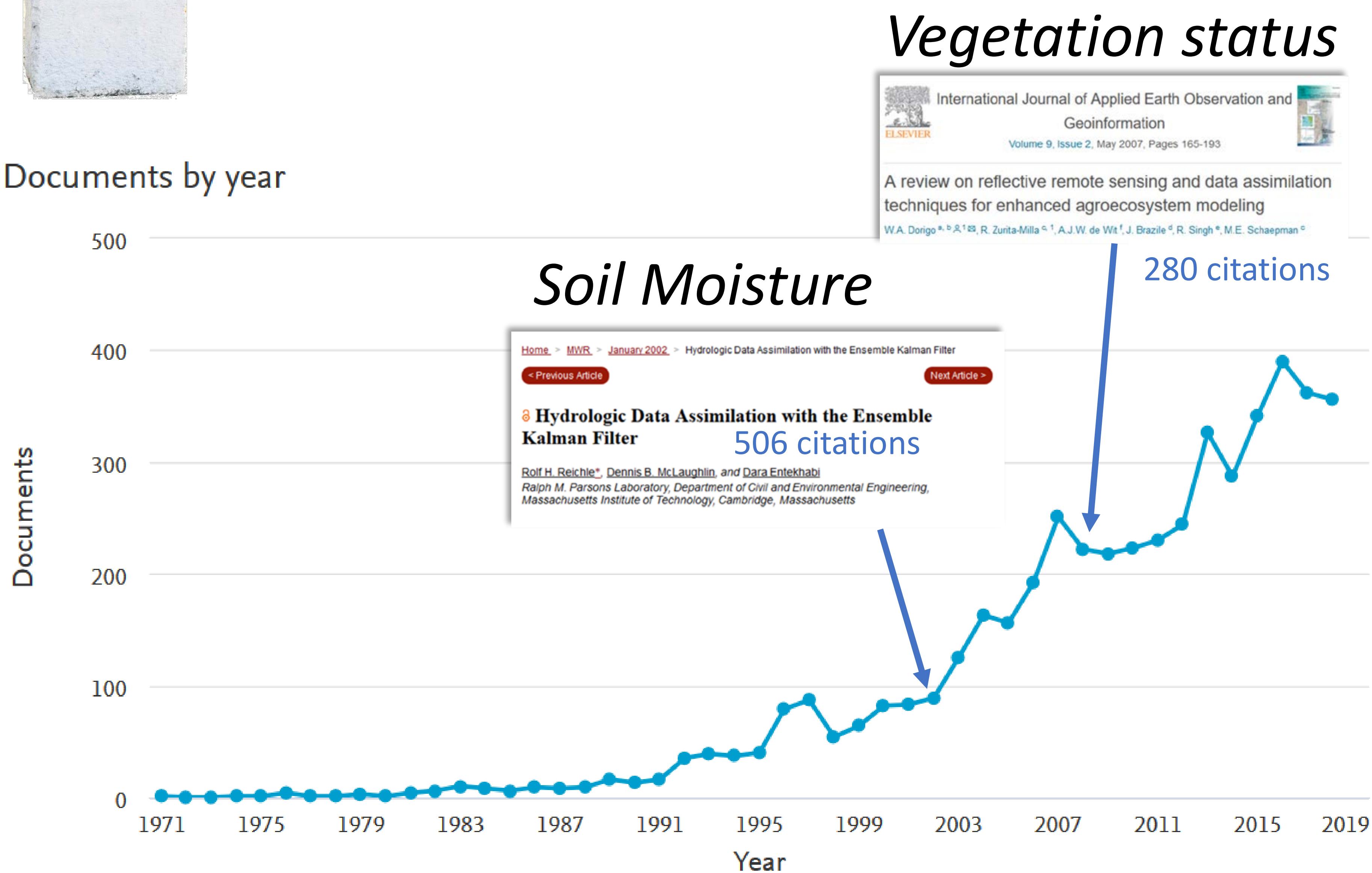
Milestones: the large assimilation systems

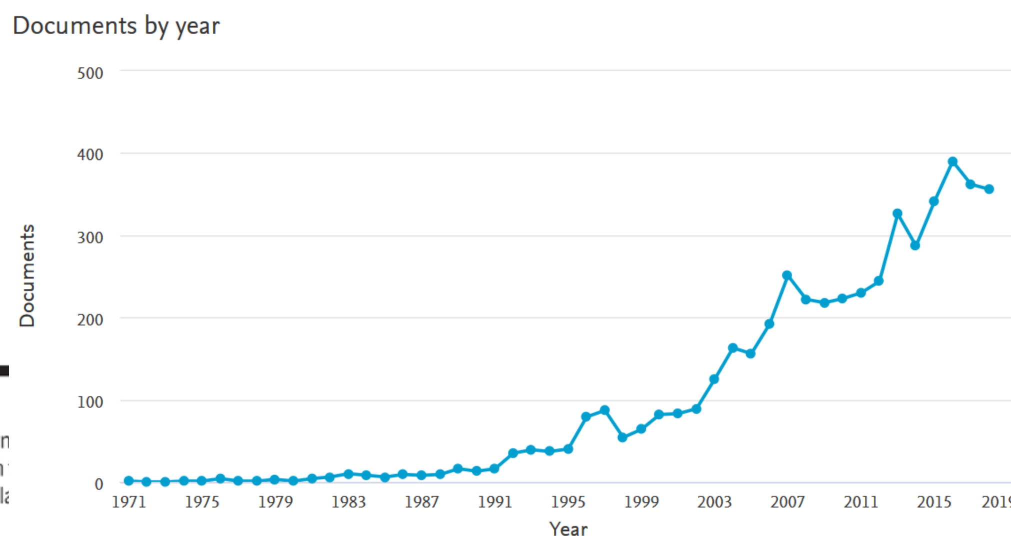
(the three most quoted papers)



Milestones: focus on hydrology

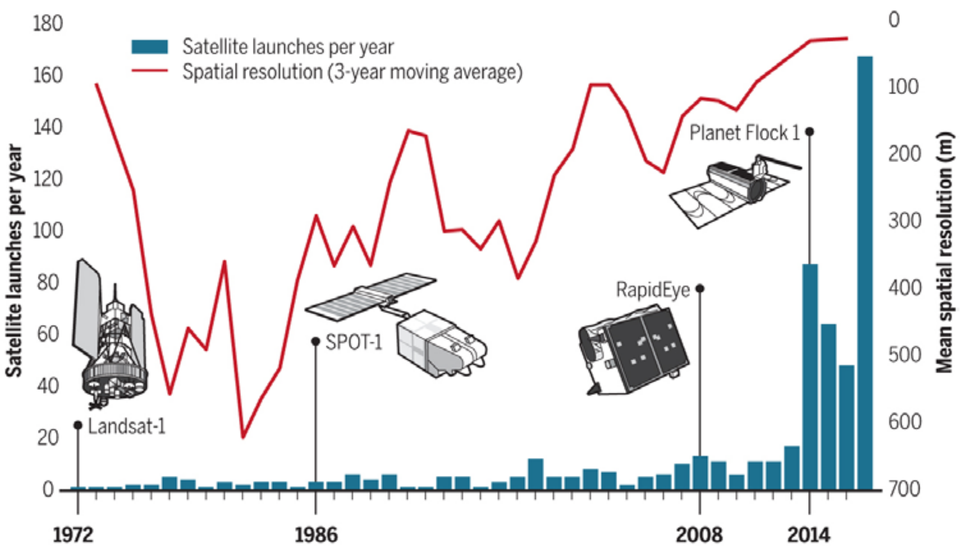
(among most quoted papers)

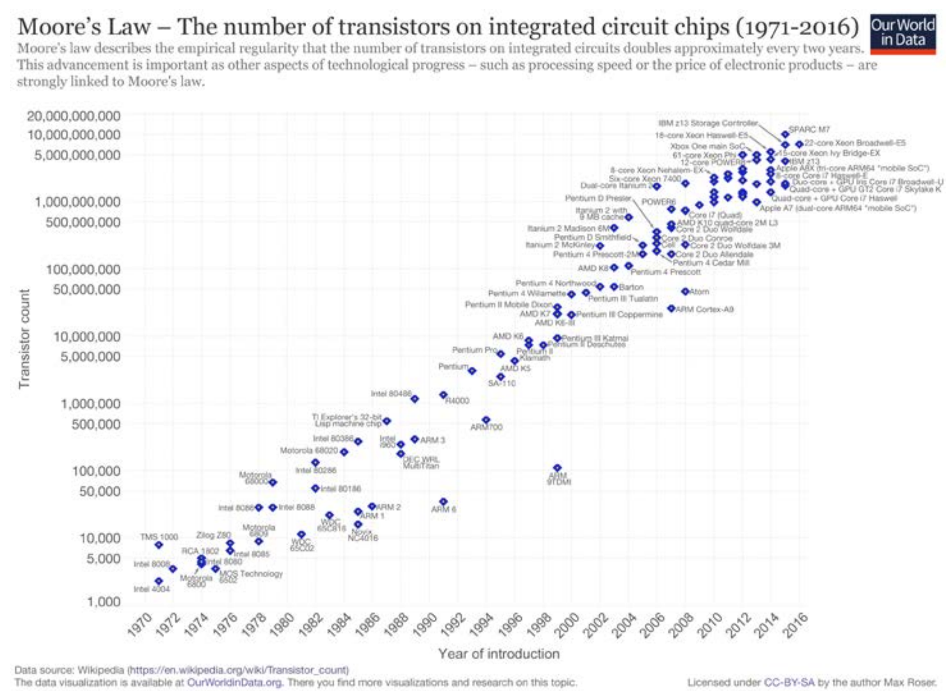




Trends in earth observation satellites

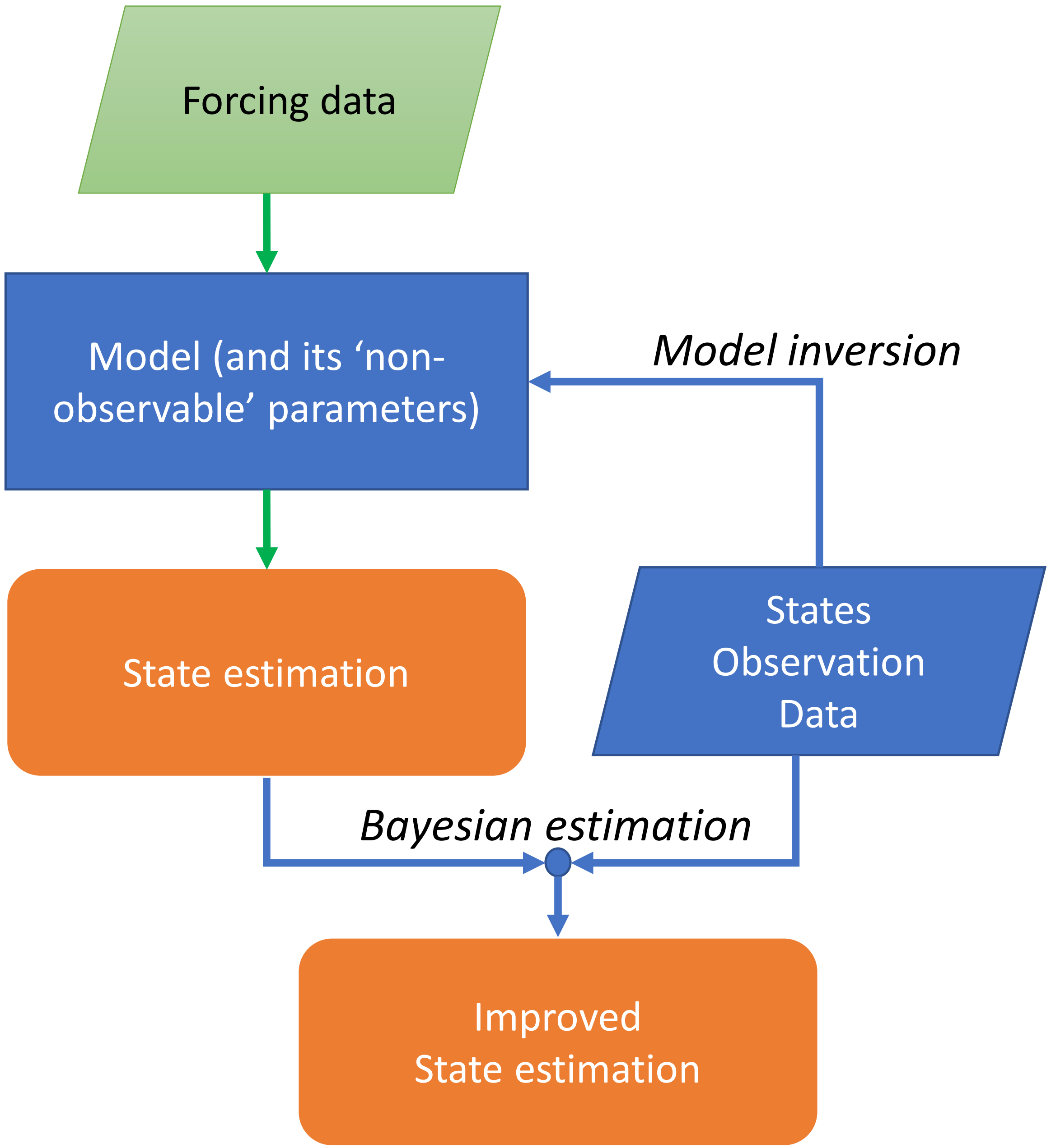
Data reflect 488 earth observation satellites launched since 1972 by commercial and govern
providers (excluding military). We followed methods established in (5) and added satellites from
of Concerned Scientists database and public launch information from SpaceFlightNow and Pla
See the supplementary materials for details.





One working definition of **data assimilation**

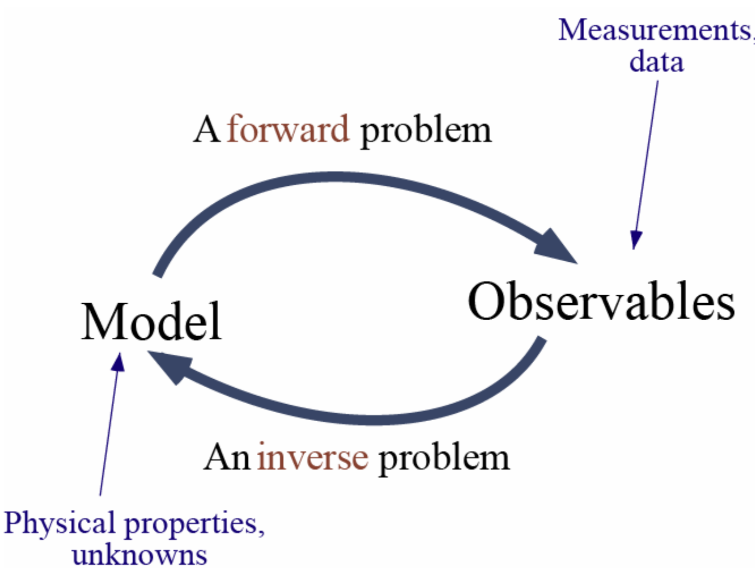
The set techniques the combine data with some underlying process model to provide optimal estimates of the true state and/or parameters of that model.



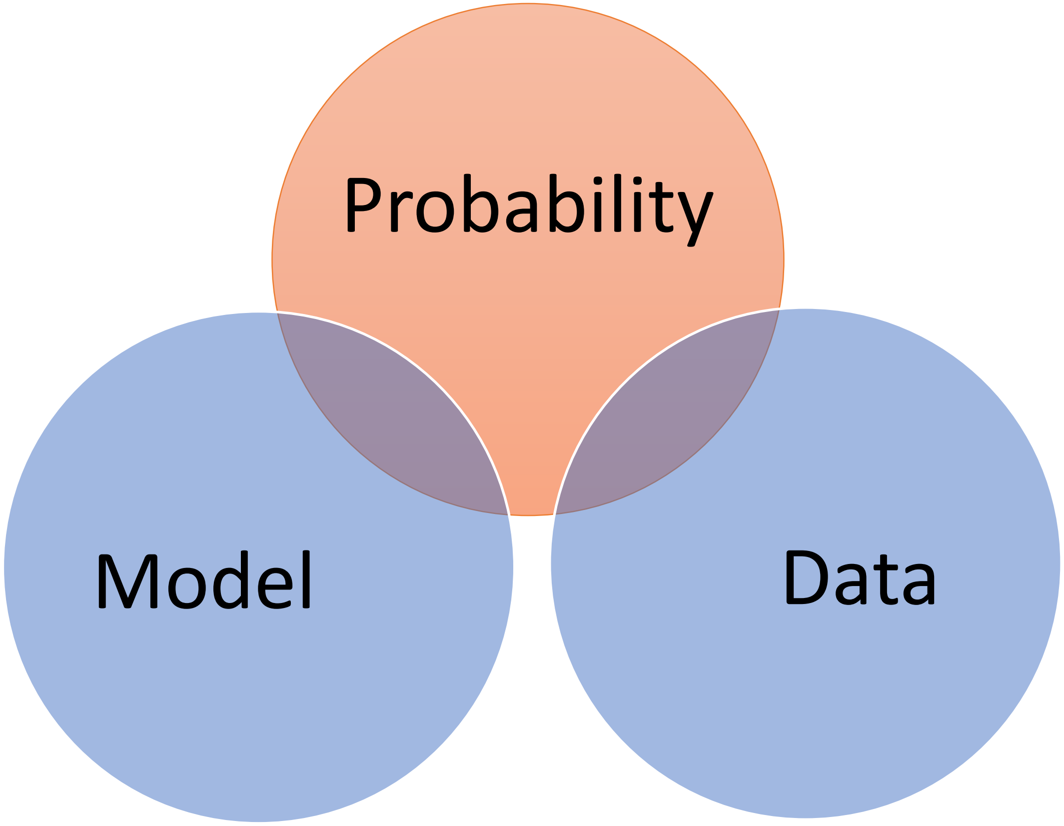
<u>Inverse problems: Reasoning backwards</u>

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)

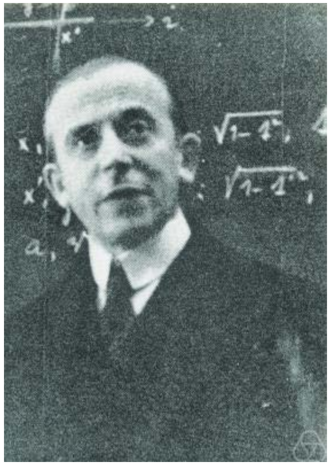


Key underlying concept: both data and model have **errors**, hopefully of different origin, so model and data can complement each other to reduce the overall **uncertainty**



Historic approaches to probability

Von Mises vs. Bayes


1883-1953


1702-1761

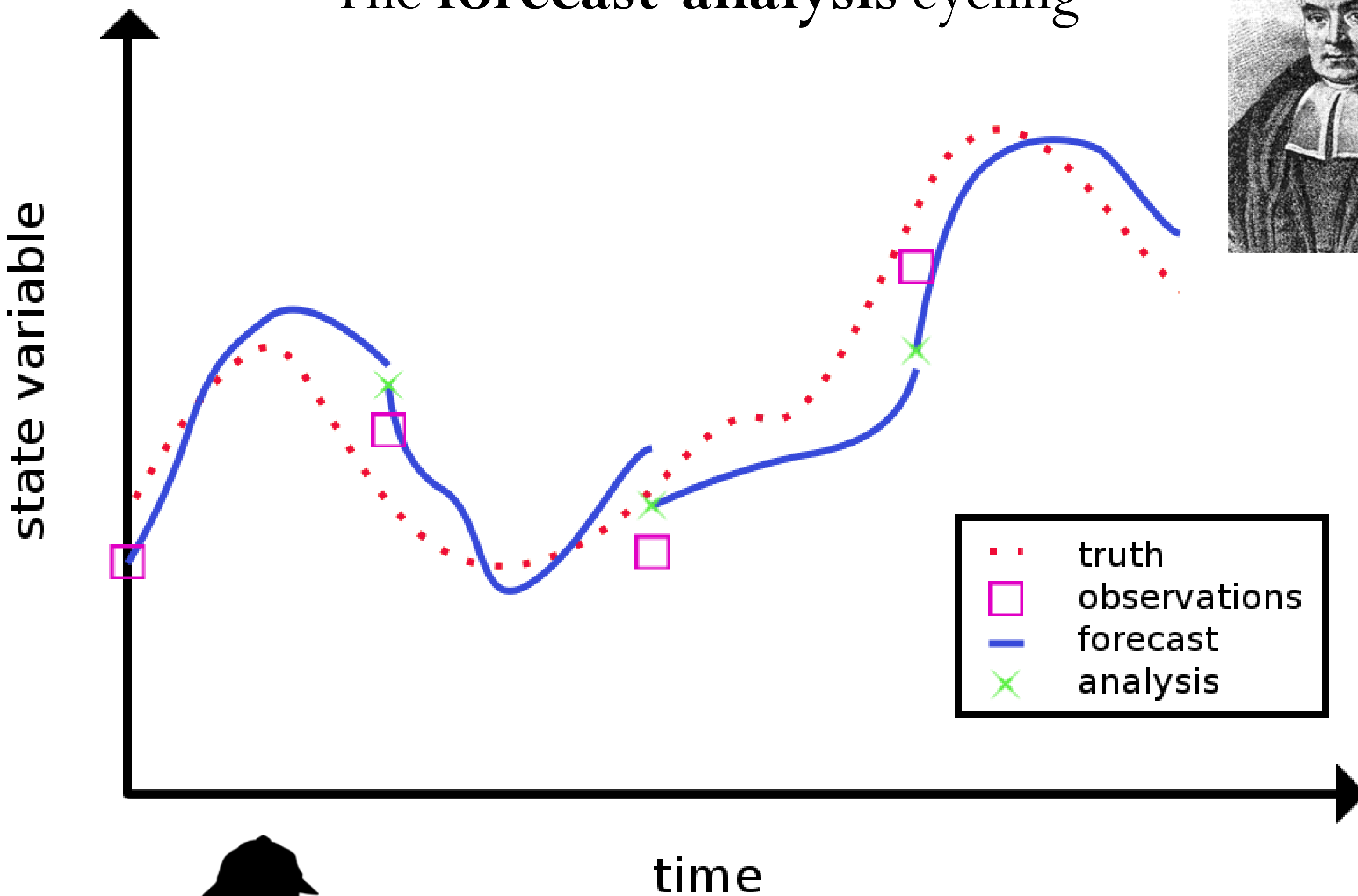
The frequentist approach:

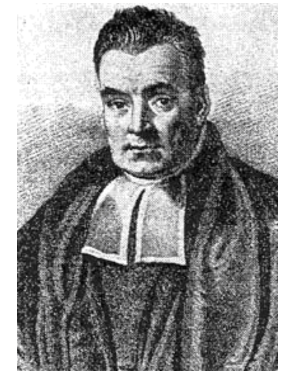
- Strictly objective probabilities
- Predictions are repeated in time, many times
- Prediction erros can be precisely described by statistical means

The Bayesian approach:

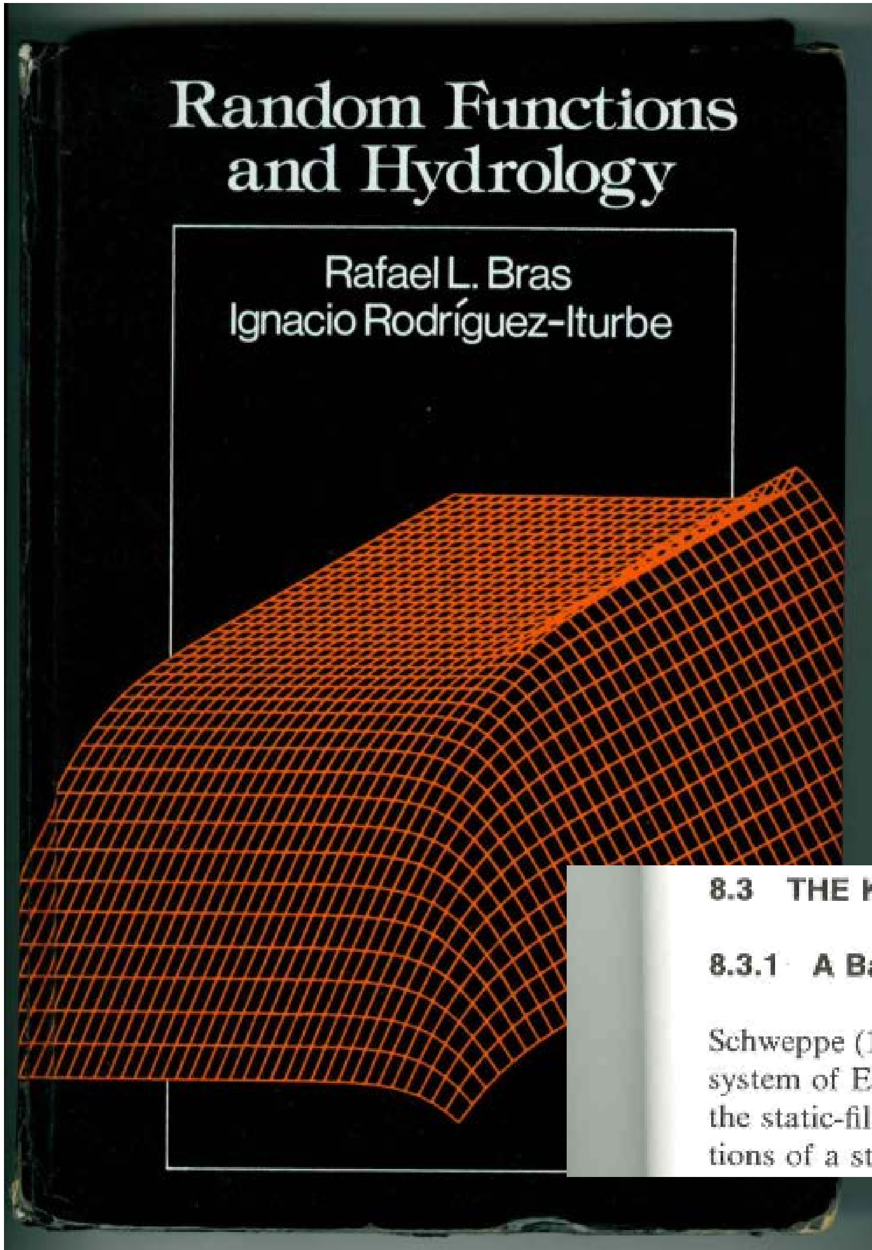
- Proabality as 'degree of belief' (largely subjective)
- Predictions are sporadic in time (even once in a lifetime)
- Utility comes on …

The **forecast-analysis** cycling





How to perform the analysis for a control state variable that is not observed?



CHAPTER 8

Estimation of Dynamic Hydrologic Systems

8.1 THE STATE–SPACE REPRESENTATION OF A STOCHASTIC LINEAR DYNAMIC SYSTEM

Many physical and geophysical systems can be represented by linear differential equations of the form

$$\frac{d^nX(t)}{dt^n} + a_{n-1}(t)\frac{d^{n-1}X(t)}{dt^{n-1}} + \cdots + a_1(t)\frac{dX(t)}{dt} + a_0(t)X(t) = L(t)U(t).$$

8.3 THE KALMAN FILTER

8.3.1 A Bayesian Approach for the Discrete Filter

Schweppe (1973) derives the filtering algorithm for the dynamic discrete linear system of Eq. (8.25) with discrete observations $\mathbf{Z}(k) = \mathbf{H}(k)\mathbf{X}(k) + \mathbf{V}(k)$ using the static-filter results of Chapter 7. The idea is to combine repetitive observations of a state vector $\mathbf{X}(k)$.

Recursive Bayes Estimation

Discrete Linear Gaussian

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.

Introduction

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.

In his pioneering work, Wiener [1][3] showed that problems (i) and (ii) lead to the so-called Wiener-Hopf integral equation; he also gave a method (spectral factorization) for the solution of this integral equation in the practically important special case of stationary statistics and rational spectra.

Many extensions and generalizations followed Wiener's basic work. Zadeh and Ragazzini solved the finite-memory case [2]. Concurrently and independently of Bode and Shannon [3], they also gave a simplified method [2] of solution. Booton discussed the nonstationary Wiener-Hopf equation [4]. These results are now in standard texts [5-6]. A somewhat different approach along these main lines has been given recently by Darlington [7]. For extensions to sampled signals, see, e.g., Franklin [8], Lees [9]. Another approach based on the eigenfunctions of the Wiener-Hopf equation (which applies also to nonstationary problems whereas the preceding methods in general don't), has been pioneered by Davis [10] and applied by many others, e.g., Shinbrot [11], Blum [12], Pugachev [13], Solodovnikov [14].

In all these works, the objective is to obtain the specification of a linear dynamic system (Wiener filter) which accomplishes the prediction, separation, or detection of a random signal.[4]

Present methods for solving the Wiener problem are subject to a number of limitations which seriously curtail their practical usefulness:

(1) The optimal filter is specified by its impulse response. It is not a simple task to synthesize the filter from such data.

(2) Numerical determination of the optimal impulse response is often quite involved and poorly suited to machine computation. The situation gets rapidly worse with increasing complexity of the problem.

(3) Important generalizations (e.g., growing-memory filters, nonstationary prediction) require new derivations, frequently of considerable difficulty to the nonspecialist.

(4) The mathematics of the derivations are not transparent. Fundamental assumptions and their consequences tend to be obscured.

This paper introduces a new look at this whole assemblage of problems, sidestepping the difficulties just mentioned. The following are the highlights of the paper:

(5) *Optimal Estimates and Orthogonal Projections.* The Wiener problem is approached from the point of view of conditional distributions and expectations. In this way, basic facts of the Wiener theory are quickly obtained; the scope of the results and the fundamental assumptions appear clearly. It is seen that all statistical calculations and results are based on first and second order averages; no other statistical data are needed. Thus difficulty (4) is eliminated. This method is well known in probability theory (see pp. 75–78 and 148–155 of Doob [15] and pp. 455–464 of Loève [16]) but has not yet been used extensively in engineering.

(6) *Models for Random Processes.* Following, in particular, Bode and Shannon [3], arbitrary random signals are represented (up to second order average statistical properties) as the output of a linear dynamic system excited by independent or uncorrelated random signals ("white noise"). This is a standard trick in the engineering applications of the Wiener theory [2–7]. The approach taken here differs from the conventional one only in the way in which linear dynamic systems are described. We shall emphasize the concepts of *state* and *state transition;* in other words, linear systems will be specified by systems of first-order difference (or differential) equations. This point of view is

[1] This research was supported in part by the U. S. Air Force Office of Scientific Research under Contract AF 49 (638)-382.

[2] 7212 Bellona Ave.

[3] Numbers in brackets designate References at end of paper.

[4] Of course, in general these tasks may be done better by nonlinear filters. At present, however, little or nothing is known about how to obtain (both theoretically and practically) these nonlinear filters.

Contributed by the Instruments and Regulators Division and presented at the Instruments and Regulators Conference, March 29– April 2, 1959, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, February 24, 1959. Paper No. 59—IRD-11.

Transactions of the ASME–Journal of Basic Engineering, 82 (Series D): 35-45. Copyright © 1960 by ASME

Least-squares estimation: from Gauss to Kalman

The Gaussian concept of estimation by least squares, originally stimulated by astronomical studies, has provided the basis for a number of estimation theories and techniques during the ensuing 170 years—probably none as useful in terms of today's requirements as the Kalman filter

H. W. Sorenson University of California, San Diego

This discussion is directed to least-squares estimation theory, from its inception by Gauss[1] to its modern form, as developed by Kalman.[2] To aid in furnishing the desired perspective, the contributions and insights provided by Gauss are described and related to developments that have appeared more recently (that is, in the 20th century). In the author's opinion, it is enlightening to consider just how far (or how little) we have advanced since the initial developments and to recognize the truth in the saying that we "stand on the shoulders of giants."

have made use of since the year 1795, has lately been published by Legendre in the work *Nouvelles méthodes pour la determination des orbites des cometes, Paris, 1806,* where several other properties of this principle have been explained which, for the sake of brevity, we here omit." This reference angered Legendre who, with great indignation, wrote to Gauss and complained[3] that "Gauss, who was already so rich in discoveries, might have had the decency not to appropriate the method of least-squares." It is interesting to note that Gauss, who is now regarded as one of the "giants" of mathematics, felt that he had

Discrete Kalman Filter

Estimate the state $\boldsymbol{x} \in \Re^n$ of a
linear stochastic difference equation (model)

$$\boldsymbol{x}_k = \boldsymbol{A}\boldsymbol{x}_{k-1} + \boldsymbol{B}\boldsymbol{u}_k + \boldsymbol{w}_{k-1}$$

where $\boldsymbol{w}$ is a $N(0, \boldsymbol{Q})$ process noise (model error) with zero mean and covariance matrix $\boldsymbol{Q} \in \Re^n \times \Re^n$,

given measurements (data) $\boldsymbol{z} \in \Re^m$ related to states through the
linear observation equation

$$\boldsymbol{z}_k = \boldsymbol{H}\boldsymbol{x}_k + \boldsymbol{v}_k$$

where $\boldsymbol{v}$ is a $N(0, \boldsymbol{R})$ measurement noise (data error) with zero mean and covariance matrix $\boldsymbol{R} \in \Re^m \times \Re^m$

Estimates

$\hat{\boldsymbol{x}}_{\boldsymbol{k}} \in \Re^n$ is the estimated state at time step k

$\hat{\boldsymbol{x}}_{\boldsymbol{k}}^{-} \in \Re^n$ .... after prediction, before observation (prior, background)

$\hat{\boldsymbol{x}}_{\boldsymbol{k}}^{+} \in \Re^n$ .... after prediction and observation (posterior, analysis)

Errors

$\boldsymbol{e}_{\boldsymbol{k}}^{-} = \boldsymbol{x}_k - \hat{\boldsymbol{x}}_{\boldsymbol{k}}^{-}$ with covariance matrix $\boldsymbol{P}_k^{-} = \mathbb{E}\left[\boldsymbol{e}_{\boldsymbol{k}}^{-} \boldsymbol{e}_{\boldsymbol{k}}^{-T}\right]$

$\boldsymbol{e}_{\boldsymbol{k}}^{+} = \boldsymbol{x}_k - \hat{\boldsymbol{x}}_{\boldsymbol{k}}^{+}$ with covariance matrix $\boldsymbol{P}_k^{+} = \mathbb{E}\left[\boldsymbol{e}_{\boldsymbol{k}}^{+} \boldsymbol{e}_{\boldsymbol{k}}^{+T}\right]$

Goal of Kalman Filter

$$\begin{bmatrix} \hat{\boldsymbol{x}}_{\boldsymbol{k}}^{-} \\ \boldsymbol{P}_k^{-} \end{bmatrix} \Rightarrow \begin{bmatrix} \hat{\boldsymbol{x}}_{\boldsymbol{k}}^{+} \\ \boldsymbol{P}_k^{+} \end{bmatrix}; min\|\boldsymbol{P}_k^{+}\|$$

Time Update (Prediction)

$\boldsymbol{x}_k = \boldsymbol{A}\boldsymbol{x}_{k-1} + \boldsymbol{B}\boldsymbol{u}_k + \boldsymbol{w}_{k-1}$
$\boldsymbol{Q} = \mathbb{E}[\boldsymbol{w}\,\boldsymbol{w}^T]$

$$\widehat{\boldsymbol{x}}_k^- = \boldsymbol{A}\widehat{\boldsymbol{x}}_{k-1}^+ + \boldsymbol{B}\boldsymbol{u}_k$$

$$\boldsymbol{P}_k^- = \boldsymbol{A}\boldsymbol{P}_{k-1}^+\boldsymbol{A}^T + \boldsymbol{Q}$$

← Can you derive this, at least in the scalar case?

Kalman Gain

$\boldsymbol{z}_k = \boldsymbol{H}\boldsymbol{x}_k + \boldsymbol{v}_k$
$\boldsymbol{R} = \mathbb{E}[\boldsymbol{v}\,\boldsymbol{v}^T]$

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

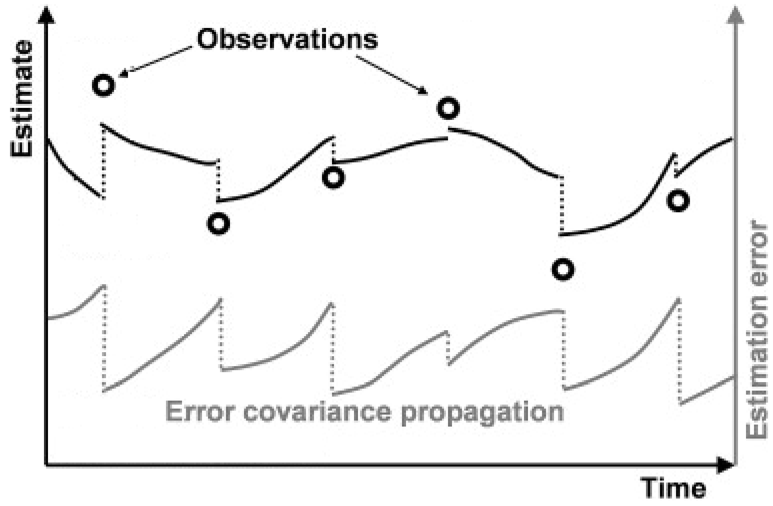
← Proportional to .....?

Measurement Update (Analysis)

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

$$\boldsymbol{P}_k^+ = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H})\boldsymbol{P}_k^-$$

← How this motivates the term 'Kalman Gain'?



Adapted from *Reichle, Adv. Water res., 2009.*

Kalman Filter main hypotheses:

- Linear model & observation equations
- Gaussian error distributions (needed to ensure $min\|\boldsymbol{P}_k^+\|$)

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

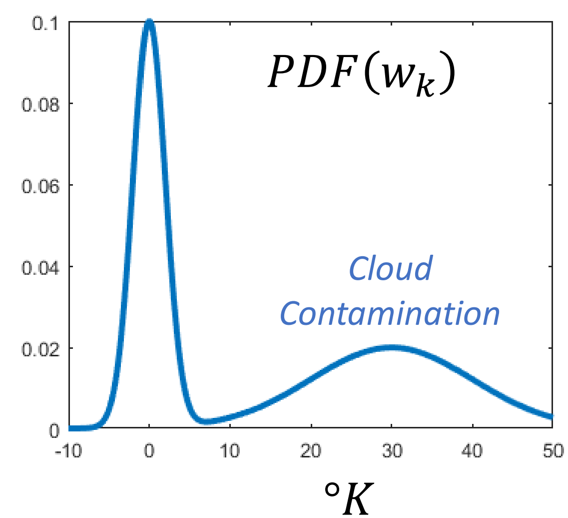
$$T_k = \alpha_k T_{k-1} + \beta S_k + v_k$$
$$T_k = LST_k + w_k$$

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:

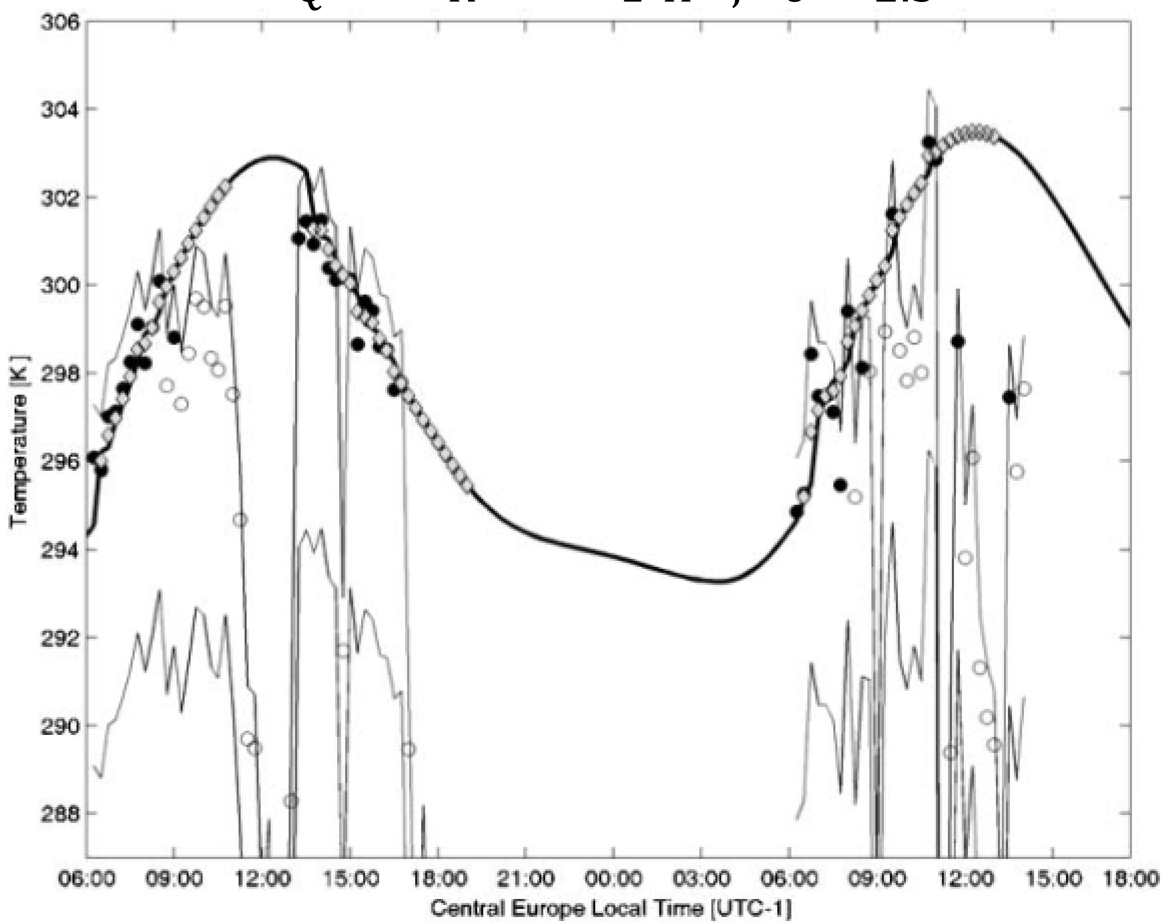
$$R_k = \begin{cases} R^* & if \quad T_k^- - LST_k \leq \delta(R^* + Q) \\ \infty & otherwise \end{cases}$$



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

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 ($P_k^+ \le R^* + Q$)

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

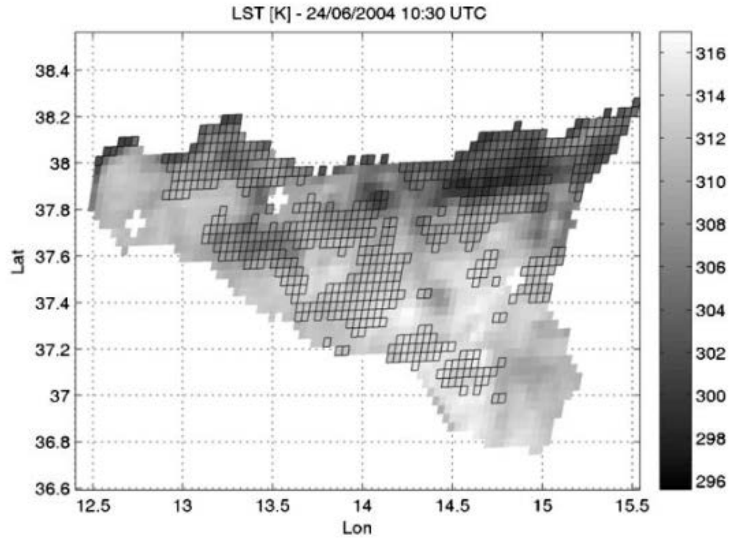
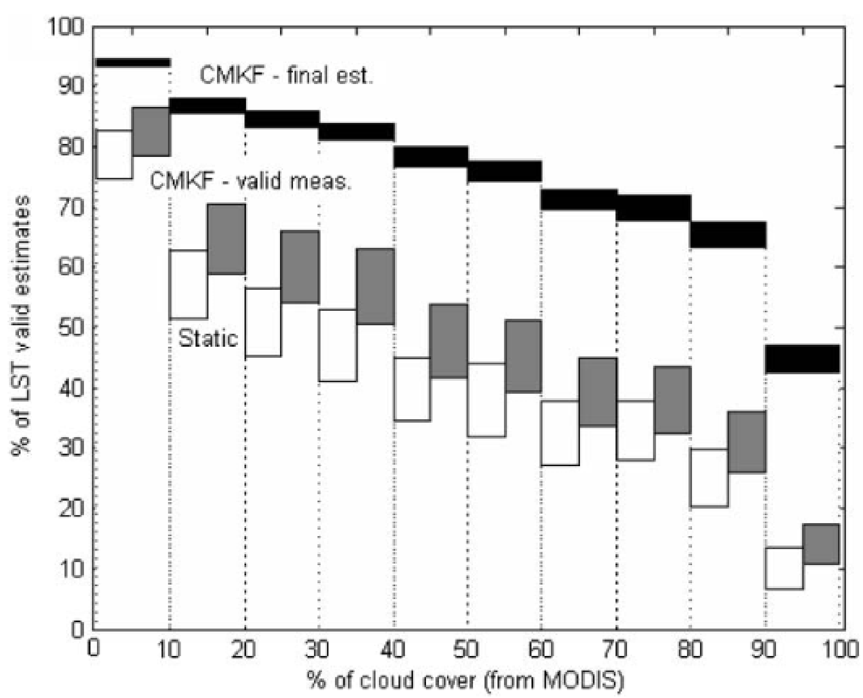


Figure 5. Enlargement of the map shown in figure 4 over Sicily. Boxed pixels identify regions where the LST retrieval was based on model prediction only, having recognized the GSW estimates as being too cloud contaminated.



Validation with MODIS

Extended Kalman Filter

Non-linear model and/or measurement equation:

$$\boldsymbol{x}_k = \boldsymbol{F}(\boldsymbol{x}_{k-1}, \boldsymbol{u}_k) + \boldsymbol{w}_{k-1}$$

$$\boldsymbol{z}_k = \boldsymbol{G}(\boldsymbol{x}_k) + \boldsymbol{v}_k$$

Linearization (Jacobians) to propagate covariances

$$\boldsymbol{P}_k^- = \boldsymbol{A}\boldsymbol{P}_{k-1}^+\boldsymbol{A}^T + \boldsymbol{Q}$$

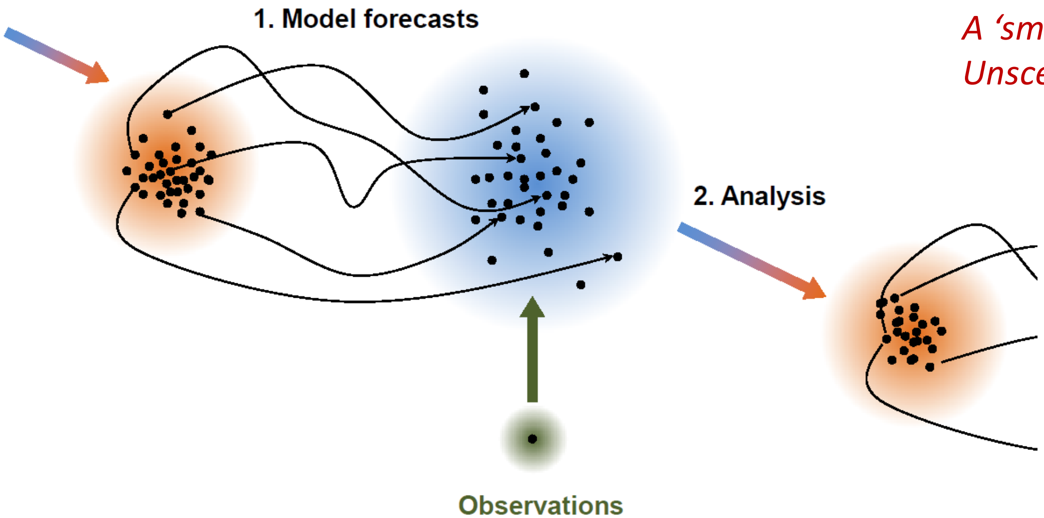
$$\boldsymbol{K}_k = \boldsymbol{P}_k^-\boldsymbol{H}^T(\boldsymbol{H}\boldsymbol{P}_k^-\boldsymbol{H}^T + \boldsymbol{R})^{-1}$$

$$\boldsymbol{P}_k^+ = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H})\boldsymbol{P}_k^-$$

$$\boldsymbol{A} = \mathbb{J}(\boldsymbol{F}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} \qquad \boldsymbol{H} = \mathbb{J}(\boldsymbol{G}) = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n}{\partial x_1} & \cdots & \frac{\partial g_n}{\partial x_n} \end{bmatrix}$$

May quickly diverge ($\boldsymbol{K}_k \to 0, \infty$) if the system is 'really' non-linear

Ensemble Kalman Filter



A 'smart' evolution of the Unscented Kalman Filter …

Non-linear properties of both model and measurement equations are fully retained, by:

- Heuristic sampling from a multidimensional Gaussian distribution;
- Not explicitly estimating the prior model error covariance $\boldsymbol{P}_k^-$ (the step that required linearization in the Ext. KF)

Initialization

M initial ensemble members $\{\hat{\boldsymbol{x}}_0^{1+}, \hat{\boldsymbol{x}}_0^{2+}, \dots, \hat{\boldsymbol{x}}_0^{M+}\}$

What choices can be made here? ... Based on what?

Ensemble Prediction

$$\hat{\boldsymbol{x}}_k^{i-} = \boldsymbol{F}\left(\hat{\boldsymbol{x}}_{k-1}^{i+}, \boldsymbol{u}_k\right) + \boldsymbol{w}_{k-1}^i$$

Explicit, sampled from $N(0, \boldsymbol{Q})$

$$\hat{\boldsymbol{x}}_k^- = \left\langle \hat{\boldsymbol{x}}_k^{i-} \right\rangle$$

Ensemble mean

$$\hat{\boldsymbol{z}}_k^i = \boldsymbol{G}\left(\hat{\boldsymbol{x}}_k^{i-}\right)$$

$$\hat{\boldsymbol{z}}_k = \left\langle \hat{\boldsymbol{z}}_k^i \right\rangle$$

Kalman Gain

Linear Kalman Filter

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

$$\boldsymbol{K}_k = \left\langle \left(\hat{\boldsymbol{x}}_k^{i-} - \hat{\boldsymbol{x}}_k^-\right)\left(\hat{\boldsymbol{z}}_k^i - \hat{\boldsymbol{z}}_k\right)^T \right\rangle \left\langle \left(\hat{\boldsymbol{z}}_k^i - \hat{\boldsymbol{z}}_k\right)\left(\hat{\boldsymbol{z}}_k^i - \hat{\boldsymbol{z}}_k\right)^T + \boldsymbol{R} \right\rangle^{-1}$$

Ensemble Analysis

$$\hat{\boldsymbol{x}}_k^{i+} = \hat{\boldsymbol{x}}_k^{i-} + \boldsymbol{K}_k\left(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k^i + \boldsymbol{v}_k^i\right)$$

$$\hat{\boldsymbol{x}}_k^+ = \left\langle \hat{\boldsymbol{x}}_k^{i+} \right\rangle$$

$\boldsymbol{P}_k^+$??

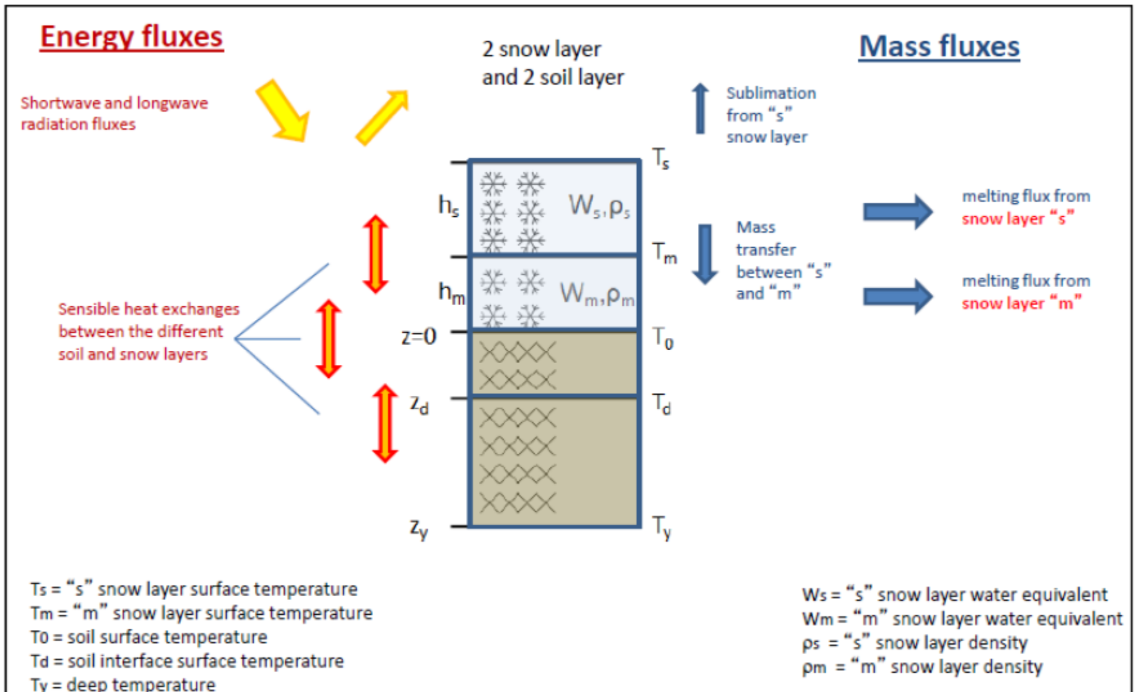
EnKF Assimilation of SEVIRI-LST & ground data for snowpack dynamics

J. Hydrol. Hydromech., 67, 2019, 1, 4–19
DOI: 10.2478/johh-2018-0013

An EnKF-based scheme for snow multivariable data assimilation at an Alpine site

Gaia Piazzi[1*], Lorenzo Campo[1], Simone Gabellani[1], Fabio Castelli[2], Edoardo Cremonese[3], Umberto Morra di Cella[3], Hervé Stevenin[4], Sara Maria Ratto[4]

[1] CIMA Research Foundation, via Armando Magliotto, 2 - 17100 Savona, Italy.
[2] Department of Civil and Environmental Engineering, University of Florence, Via Santa Marta, 350139 Florence, Italy.
[3] Environmental Protection Agency of Aosta Valley, Loc. Grande Charrière, 44 11020 Saint-Christophe, Aosta, Italy.
[4] Regional Center of Civil Protection, Aosta Valley Region, via Promis, 2/A - 11100 Aosta, Italy.
[*] Corresponding author. Tel.: +39 019230271. Fax: +39 01923027240. E-mail: gaia.piazzi@cimafoundation.org



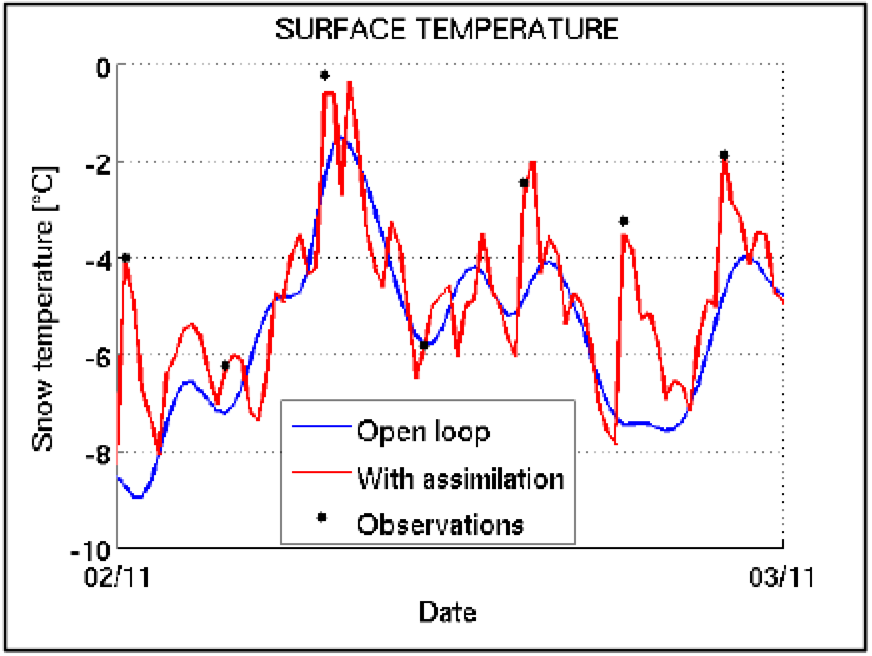
Torgnon, Val d'Aosta, 2160 m a.s.l.

Assimilated data

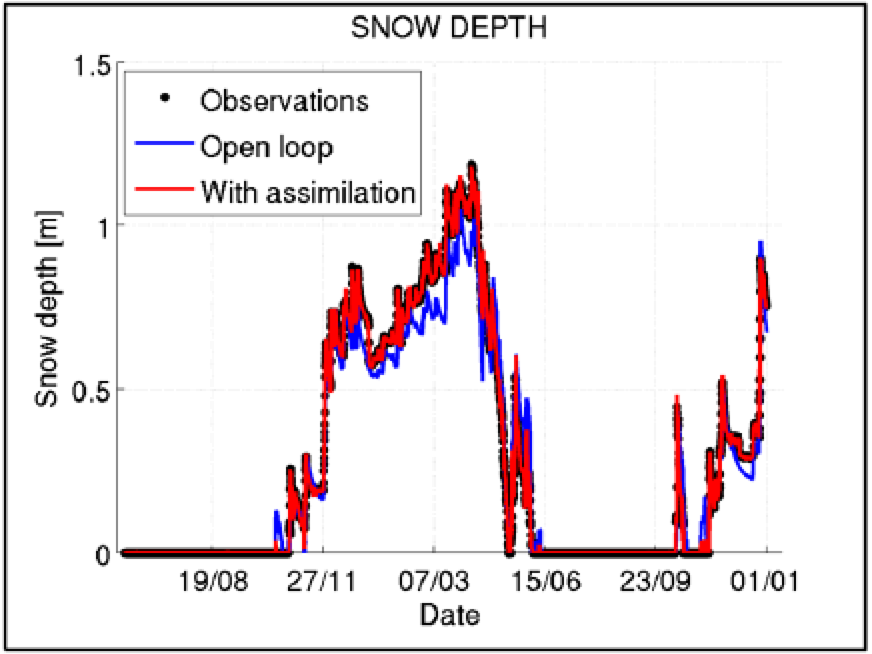
Different combinations of:

- MSG-SEVIRI LST, or
- Ground Station LST

- Snow depth
- Albedo

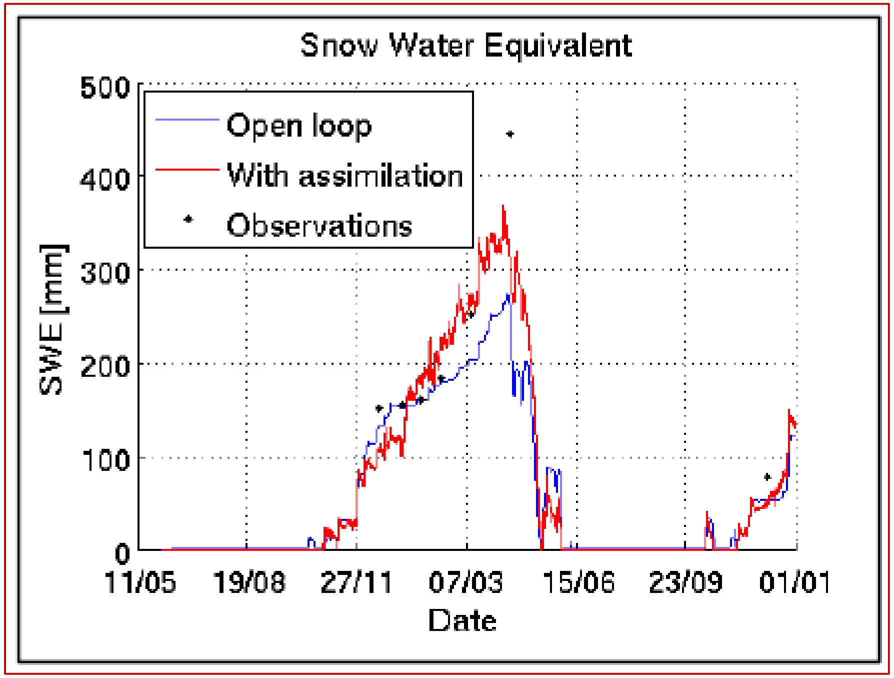


SURFACE TEMPERATURE
Snow temperature [°C]
0
-2
-4
-6
-8
-10
02/11
03/11
Date
Open loop
With assimilation
Observations

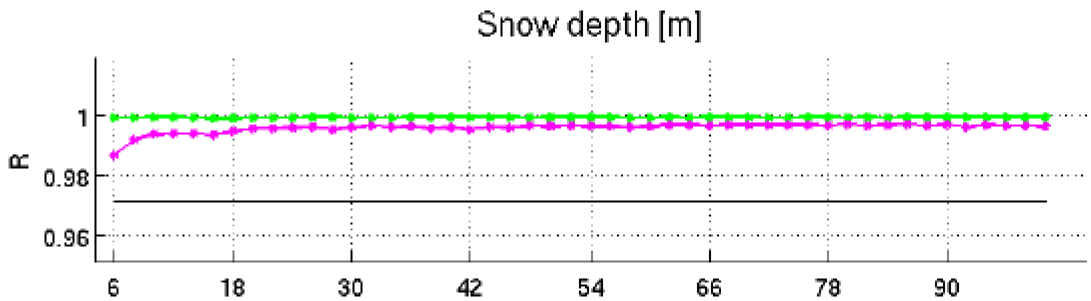


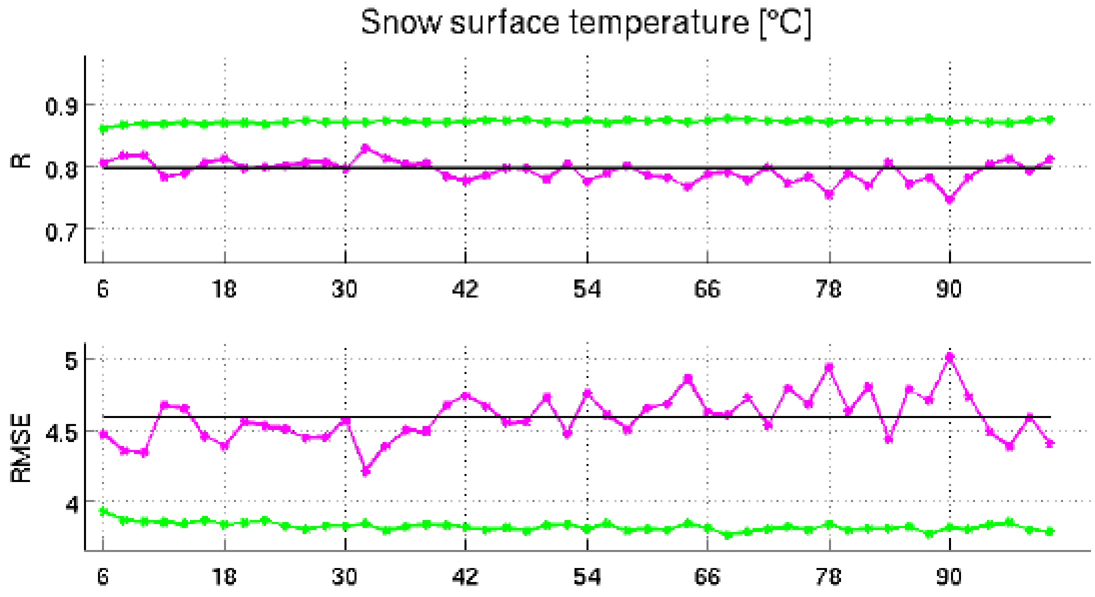
SNOW DEPTH
Snow depth [m]
1.5
1
0.5
0
19/08
27/11
07/03
15/06
23/09
01/01
Date
Observations
Open loop
With assimilation

Validation



Snow Water Equivalent
SWE [mm]
500
400
300
200
100
0
11/05
19/08
27/11
07/03
15/06
23/09
01/01
Date
Open loop
With assimilation
Observations





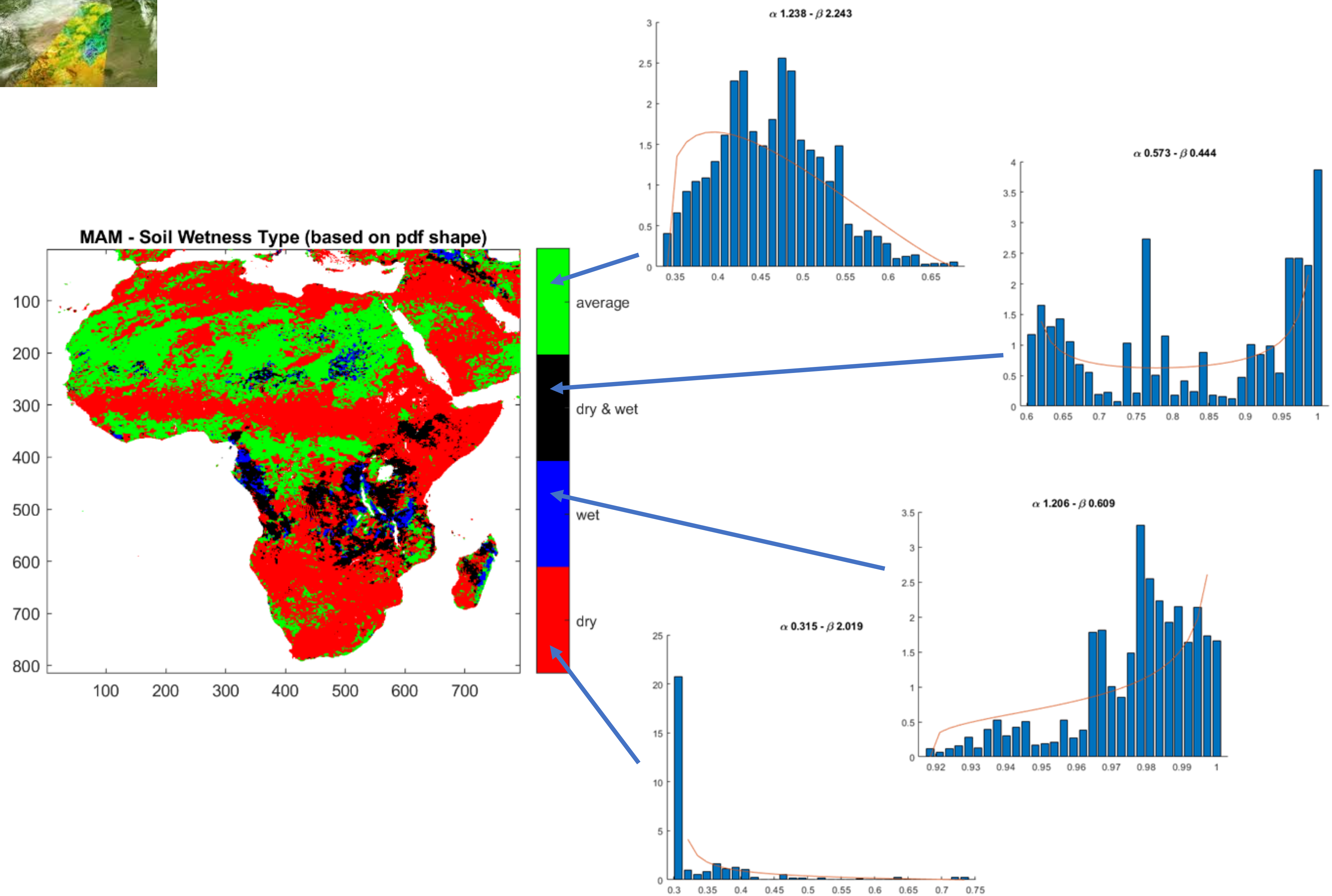
In practice:
Get reliable innovations from ground (very sparse) observations, use satellite LST to spatially interpolate them

Strengths and weaknesses in Kalman-based algorithms (LKF, EKF, UKF, EnKF …) come from the same key assumption:
Second-order (mean & covariance) approximation to probability distributions.

These two distributions have exactly the same mean an variance

Hypothesis of Gaussian distribution is not really the issue:
it is needed to prove 'optimality', but it works for LKF only anyway!

SMAP L4 Global 3-hourly 9 km Surface and Rootzone Soil Moisture, Version 4



Propagating and analyzing (.. an approximation of …) the entire proabability distribution

Sequential Monte Carlo / Particle Filter

Pros:

- Can accomodate non-linearity and non-Gaussianity
- Explicit implementation of the Recursive Bayesian State Estimation
- Very simple to code

Cons:

- Computationally expensive, especially for large dimensions
- Case-specific, euristic sampling techniques
- Risk of degeneracy of samples

Recursive Bayesian Estimation basics

Kalman Filter

$$\boldsymbol{x}_k = \boldsymbol{A}\boldsymbol{x}_{k-1} + \boldsymbol{B}\boldsymbol{u}_k + \boldsymbol{w}_{k-1}$$

$$\boldsymbol{z}_k = \boldsymbol{H}\boldsymbol{x}_k + \boldsymbol{v}_k$$

$$\hat{\boldsymbol{x}}_k^- = \boldsymbol{A}\hat{\boldsymbol{x}}_{k-1}^+ + \boldsymbol{B}\boldsymbol{u}_k$$

$$\boldsymbol{P}_k^- = \boldsymbol{A}\boldsymbol{P}_{k-1}^+\boldsymbol{A}^T + \boldsymbol{Q}$$

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

$$\boldsymbol{P}_k^+ = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H})\boldsymbol{P}_k^-$$

Probabilistic interpretation of model and observations

$$\boldsymbol{x}_k = \boldsymbol{F}(\boldsymbol{x}_{k-1}, \boldsymbol{u}_k, \boldsymbol{w}_{k-1}) \Rightarrow p(\boldsymbol{x}_k | \boldsymbol{x}_{k-1})$$

$$\boldsymbol{z}_k = \boldsymbol{G}(\boldsymbol{x}_k, \boldsymbol{v}_k) \Rightarrow p(\boldsymbol{z}_k | \boldsymbol{x}_k)$$

Forecast step (Chapman-Kolmogorov eq.)

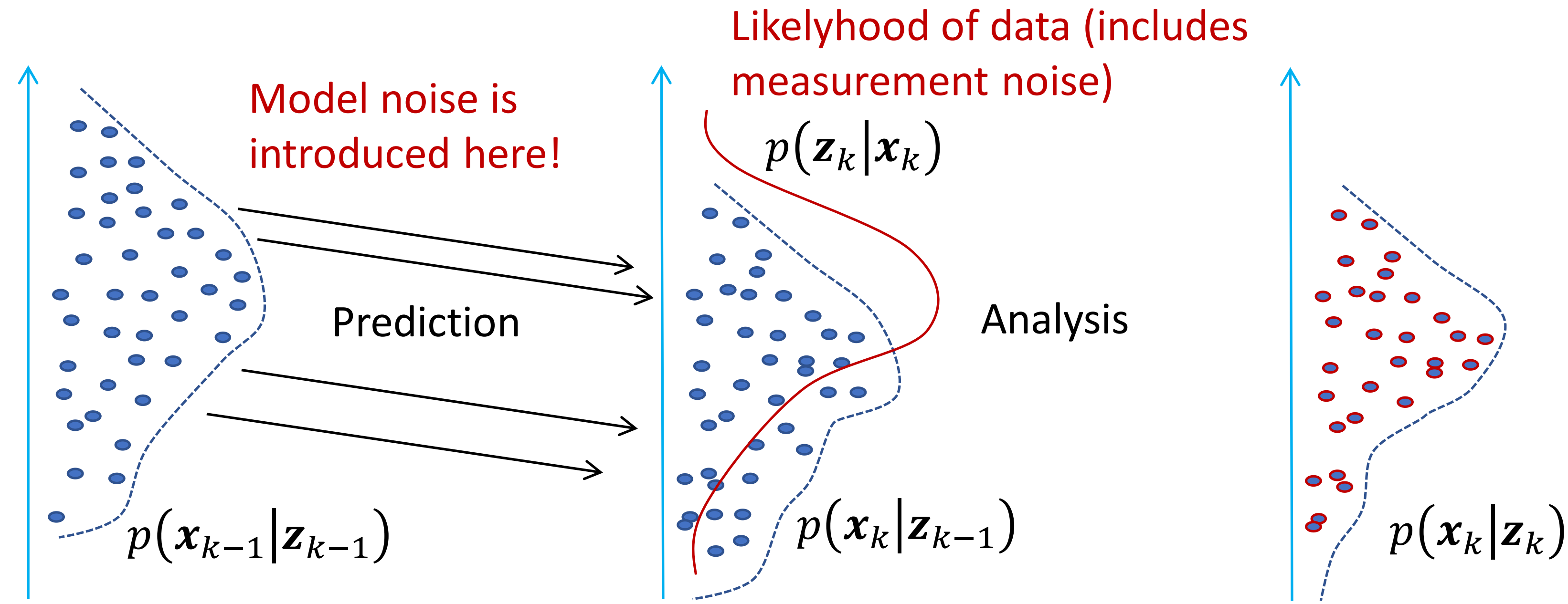
$$p(\boldsymbol{x}_k | \boldsymbol{z}_{k-1}) = \int p(\boldsymbol{x}_k | \boldsymbol{x}_{k-1}) p(\boldsymbol{x}_{k-1} | \boldsymbol{z}_{k-1}) d\boldsymbol{x}_{k-1}$$

Analysis step (Bayes rule)

$$p(\boldsymbol{x}_k | \boldsymbol{z}_k) \propto p(\boldsymbol{z}_k | \boldsymbol{x}_k) p(\boldsymbol{x}_k | \boldsymbol{z}_{k-1})$$

Particle Filter

Sequential Importance Sampling



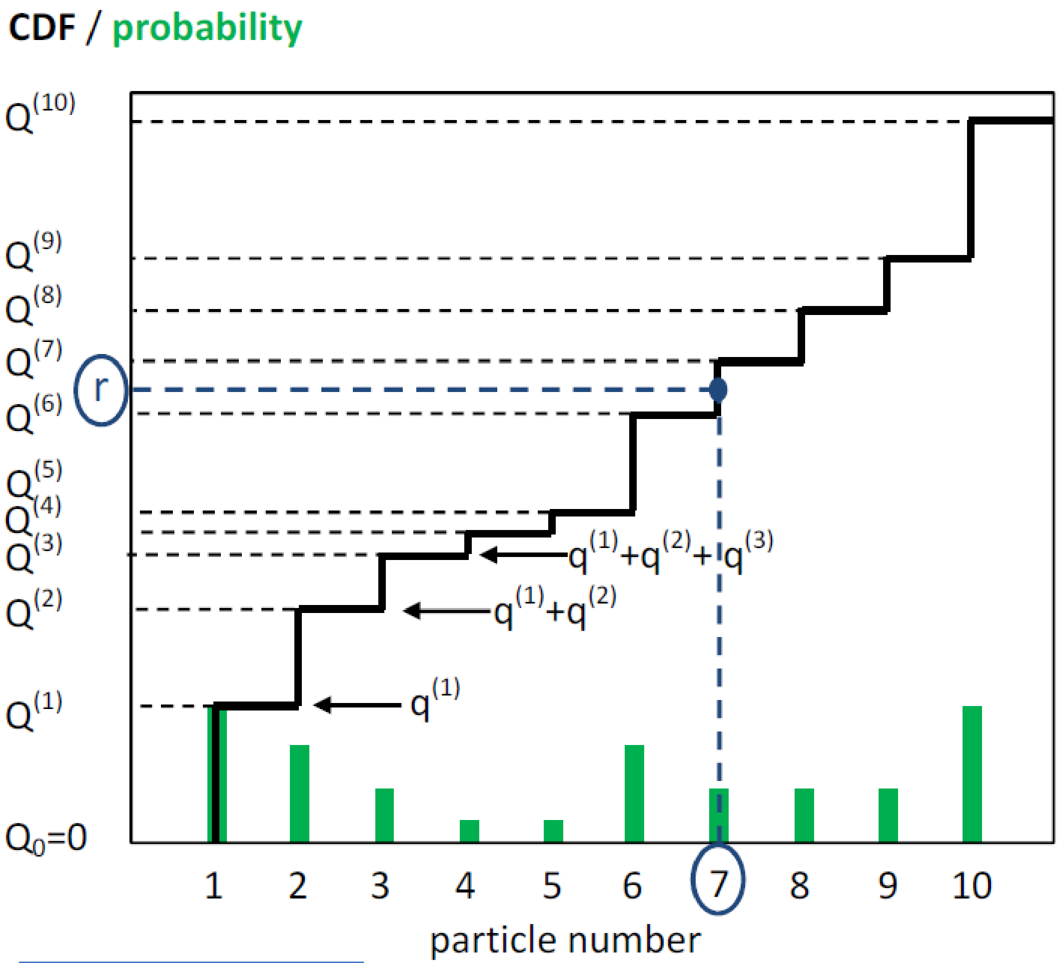
The analysis step attributes weights q^i (likelyhoods, summing to 1) to the $i = 1, \ldots, M$ particles, which are used to compute the posterior pdf

The problem of degeneracy: the likelihood of most particles becomes close to zero after a few iterations. **Need to re-sample!**

Kish's Effective Sample Size

$$M_{eff} = 1 \Big/ \sum_{i=1}^{M} {q^i}^2$$

Various resampling algorithms in the literature, variants of a similar basic concept



$$Q^i = \sum_{k=1}^{i} q^k$$

Select M new particles $\hat{\boldsymbol{x}}_{\boldsymbol{k}}^{j+}$ from the previus $\hat{\boldsymbol{x}}_{\boldsymbol{k}}^{i-}$ particles such that

$$P\left[\hat{\boldsymbol{x}}_{\boldsymbol{k}}^{j+} = \hat{\boldsymbol{x}}_{\boldsymbol{k}}^{i-}\right] = q^i$$

This is done repeating M times these two steps:

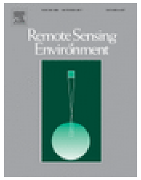
1. Generate a random number r sampling from $\mathcal{U}[0,1]$.
2. Assign the value $\hat{\boldsymbol{x}}_{\boldsymbol{k}}^{i-}$ to $\hat{\boldsymbol{x}}_{\boldsymbol{k}}^{j+}$ according to $Q^{i-1} < r \leq Q^i$

Advances in Water Resources
Volume 94, August 2016, Pages 364-378

Combined assimilation of streamflow and satellite soil moisture with the particle filter and geostatistical modeling

Hongxiang Yan, Hamid Moradkhani

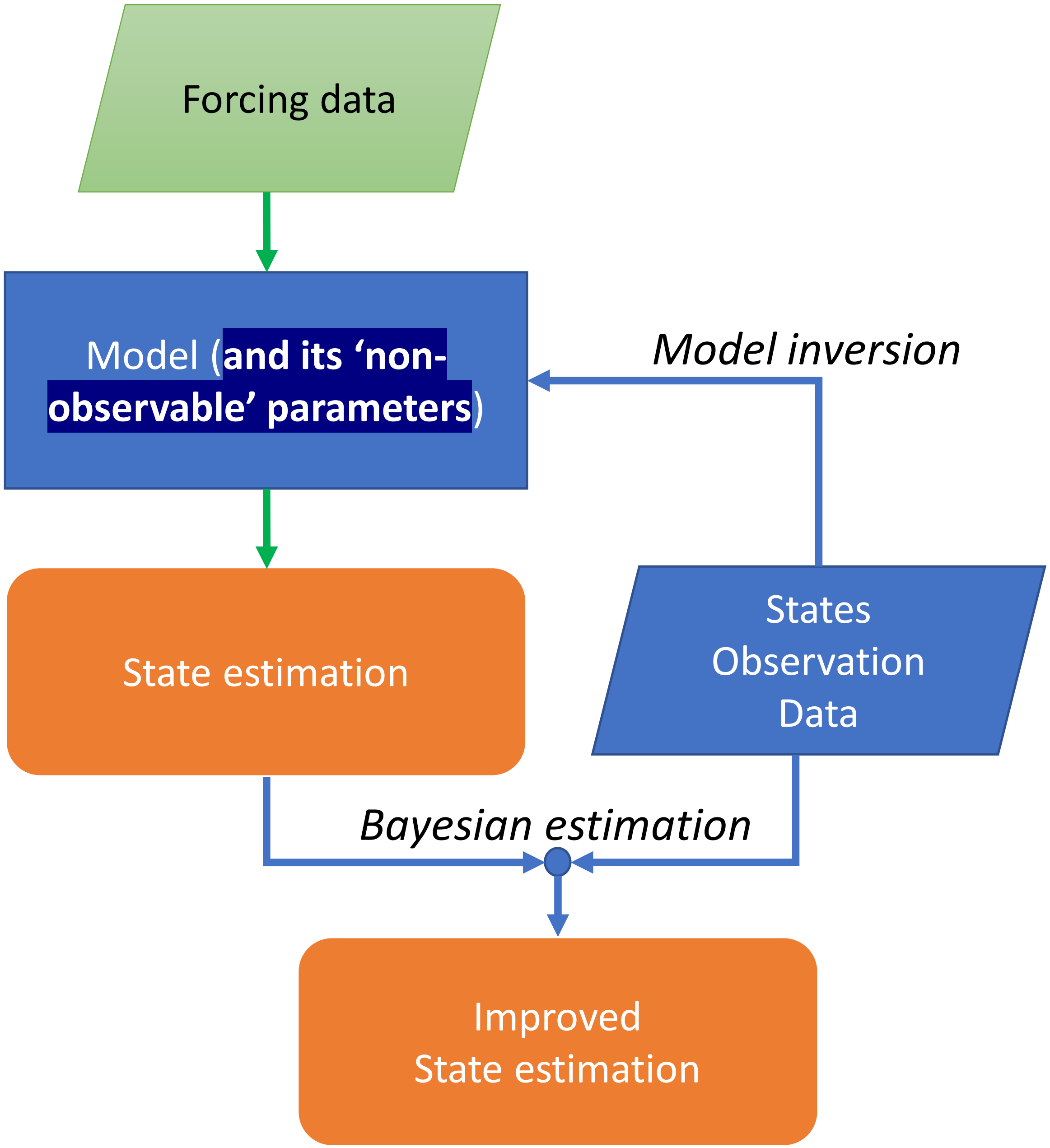
Remote Sensing of Environment
Volume 200, October 2017, Pages 295-310



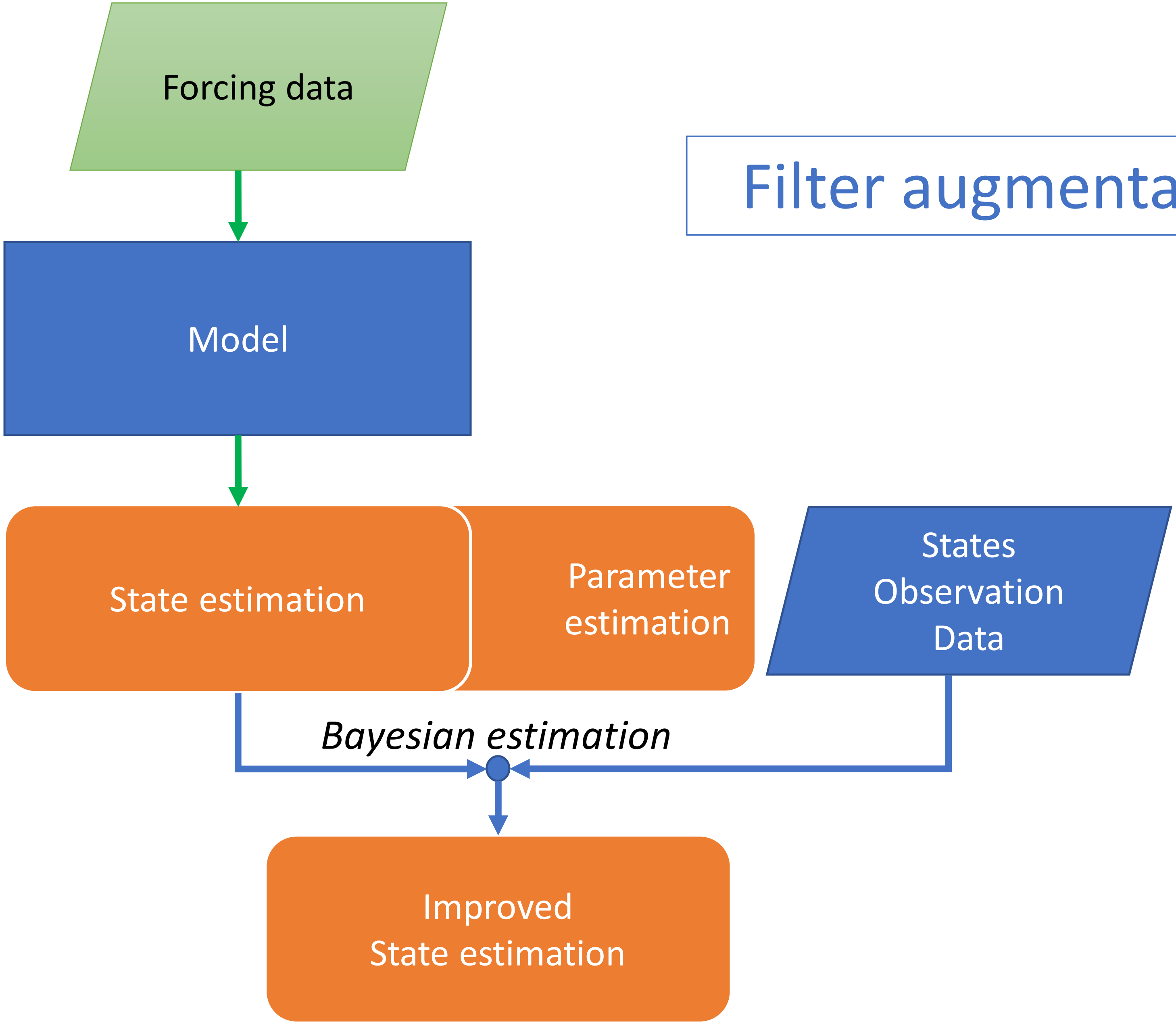
Correcting satellite-based precipitation products through SMOS soil moisture data assimilation in two land-surface models of different complexity: API and SURFEX

Carlos Román-Cascón [a], Thierry Pellarin [a], François Gibon [a], Luca Brocca [c], Emmanuel Cosme [a], Wade Crow [d], Diego Fernández-Prieto [e], Yann H. Kerr [b], Christian Massari [c]

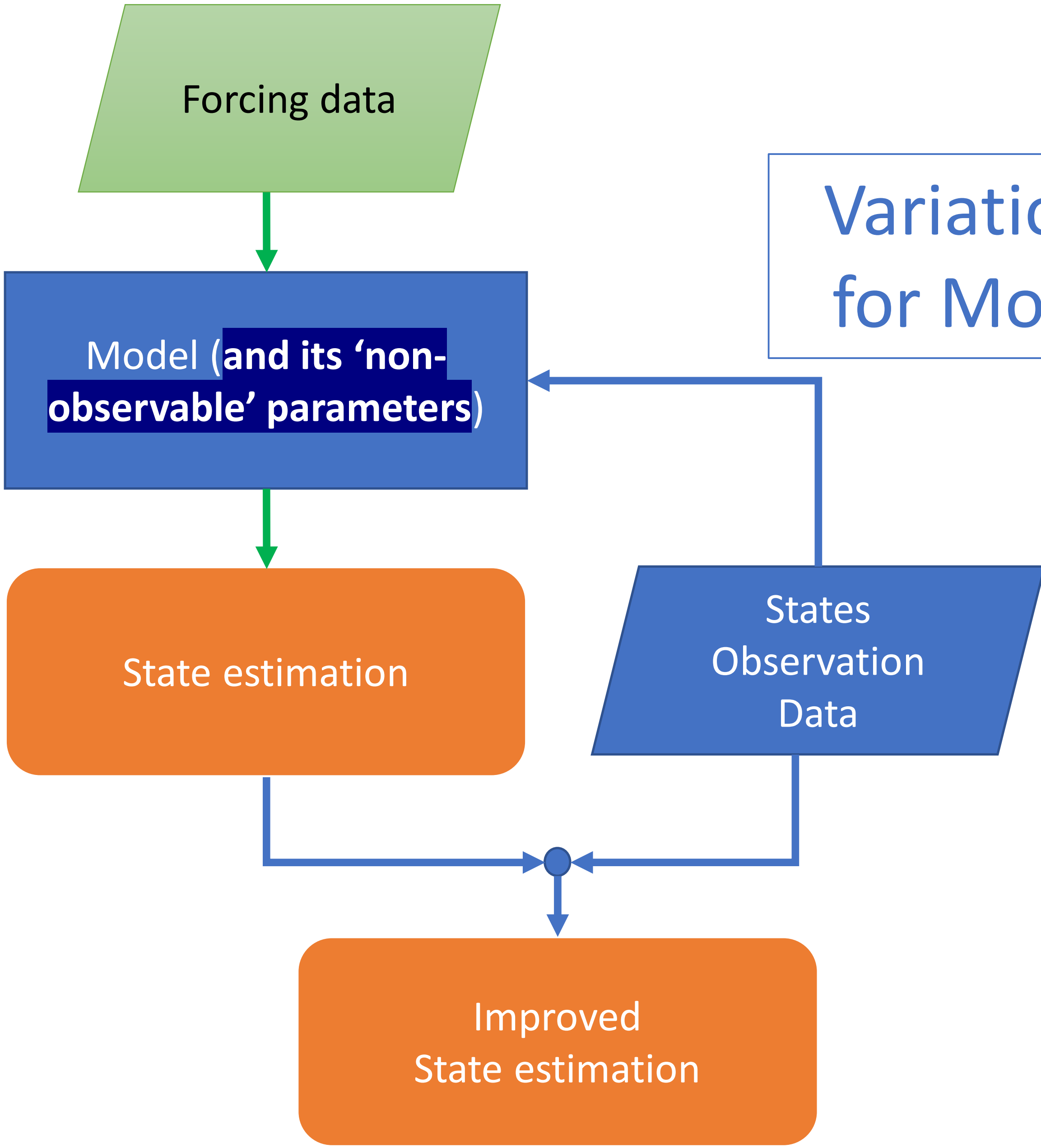
What about parameter estimation?



Filter augmentation



Variational methods for Model Inversion



Multivariate 1D-Var

Multidimensional (e.g. 4D) Var is mostly used in meteorology and oceanography)

State estimation as a time-continuous initial value problem

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{F}(\boldsymbol{x}, \boldsymbol{\theta}, \boldsymbol{u}) + \boldsymbol{w} \; t \in (t_0, t_1) \qquad \boldsymbol{x}(t_0) = \boldsymbol{x}_0$$

$$\boldsymbol{z} = \boldsymbol{G}(\boldsymbol{x}) + \boldsymbol{v}$$

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

Global penalty function with adjoined model constraint through Lagrange multipliers

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

$$\boldsymbol{J}(\boldsymbol{x}, \boldsymbol{\theta}, \boldsymbol{\lambda} | \boldsymbol{z}_k) = (\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}})\boldsymbol{\Gamma}_\theta(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}})^T + \sum_k (\boldsymbol{G}(\boldsymbol{x}) - \boldsymbol{z}_k)\boldsymbol{\Gamma}_z(\boldsymbol{G}(\boldsymbol{x}) - \boldsymbol{z}_k)^T +$$

$$+ \int_{t_0}^{t_1} \boldsymbol{\lambda} \left[\frac{dx}{dt} - \boldsymbol{F}(\boldsymbol{x}, \boldsymbol{\theta}, \boldsymbol{u})\right] dt + i.c.$$

Meaning of the three terms?

Global minimization by setting independent variates to zero

$$\boldsymbol{\delta J}(\boldsymbol{x}, \boldsymbol{\theta}, \boldsymbol{\lambda} | \boldsymbol{z}_k) = 0 \Rightarrow \begin{cases} \partial \boldsymbol{J} / \partial \boldsymbol{x} = 0 \\ \partial \boldsymbol{J} / \partial \boldsymbol{\theta} = 0 \\ \partial \boldsymbol{J} / \partial \boldsymbol{\lambda} = 0 \end{cases}$$

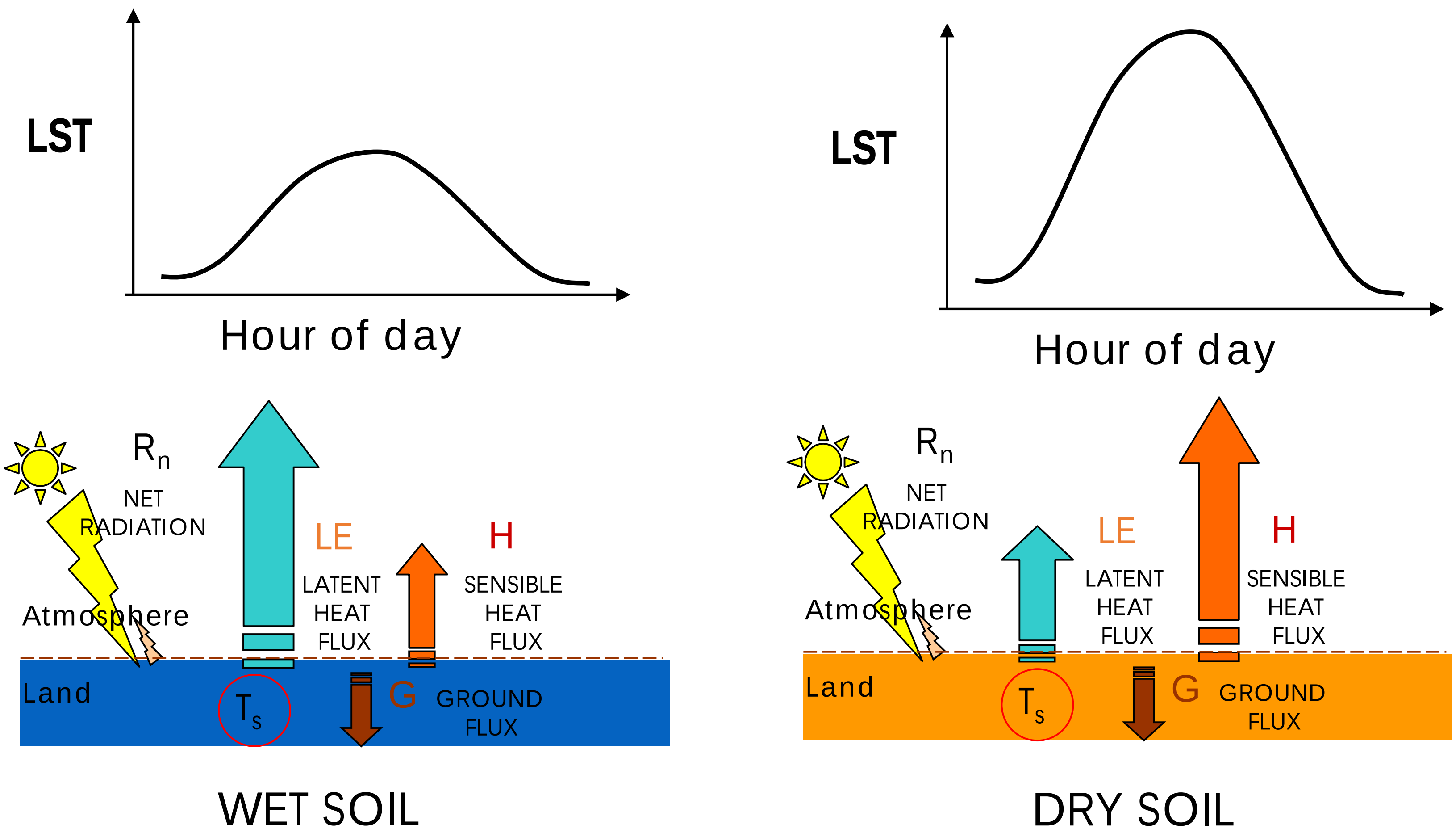
Forward Model $\dfrac{d\boldsymbol{x}}{dt} = \boldsymbol{F}(\boldsymbol{x}, \boldsymbol{\theta}, \boldsymbol{u})$

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

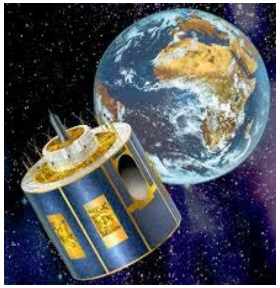
Parameters Update $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}} + \boldsymbol{\Gamma}_\theta^{-1} \int_{t_0}^{t_1} \boldsymbol{\lambda} \dfrac{\partial \boldsymbol{F}}{\partial \boldsymbol{\theta}} dt$

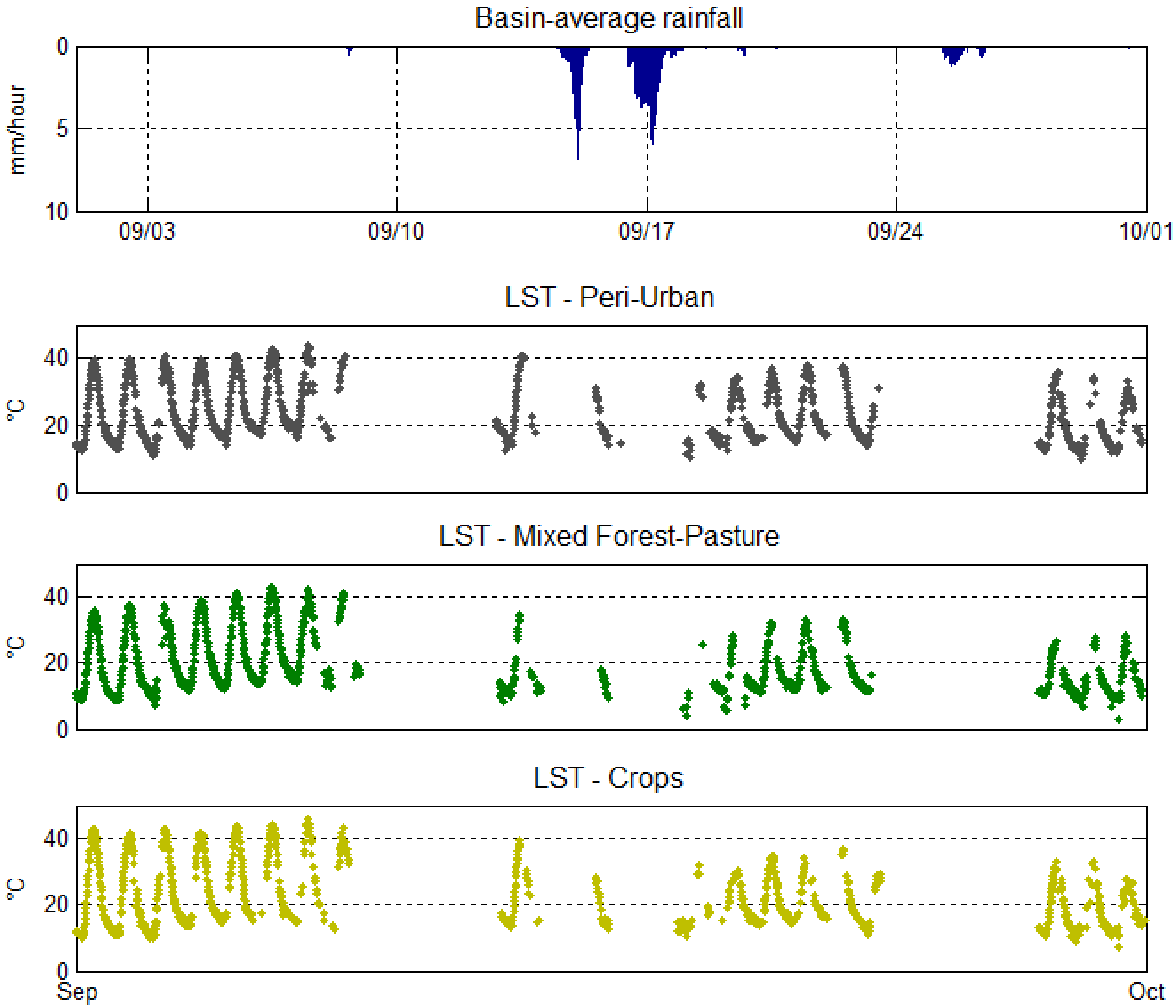
Iterate until $\boldsymbol{\lambda} \to 0$

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



Caparrini *et al.*, 2003, *BLM*, 107, 605-633
Bateni *et al.*, 2013, *WRR*, 49, 950–968

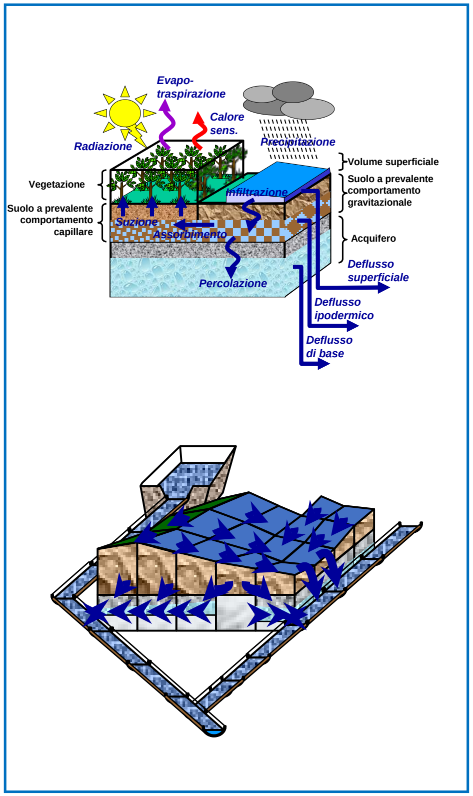




Basin-average rainfall
mm/hour
0
5
10
09/03
09/10
09/17
09/24
10/01
LST - Peri-Urban
°C
40
20
0
LST - Mixed Forest-Pasture
°C
40
20
0
LST - Crops
°C
40
20
0
Sep
Oct

Variational assimilation approach

Forward model



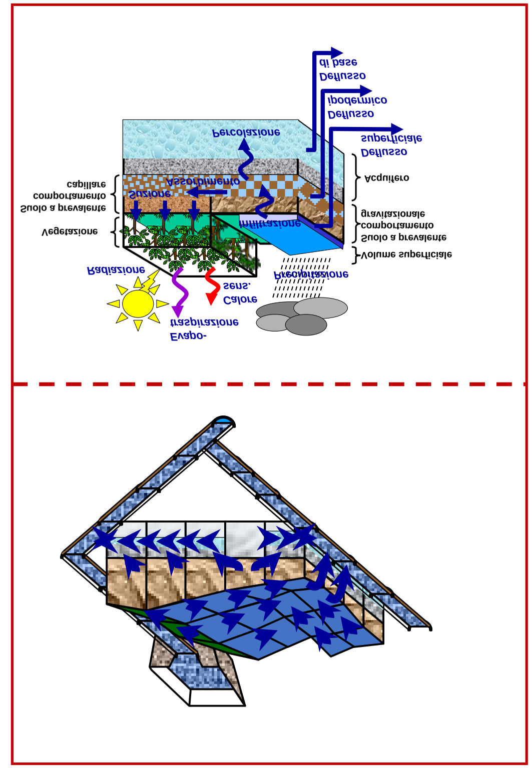
Precipitation

Runoff-formation
(**Soil Moisture state**)

Flood wave dynamics
(River hydraulics)

Streamflow

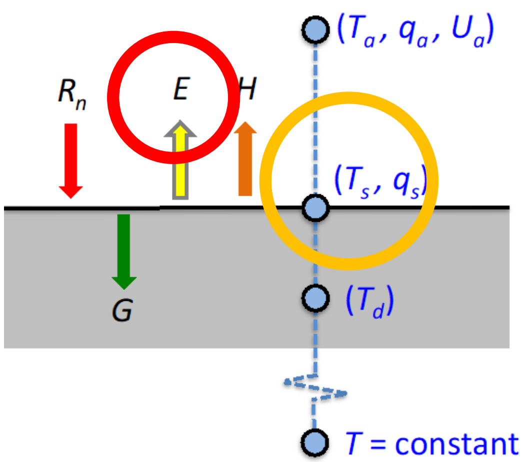
Adjoint sub-models

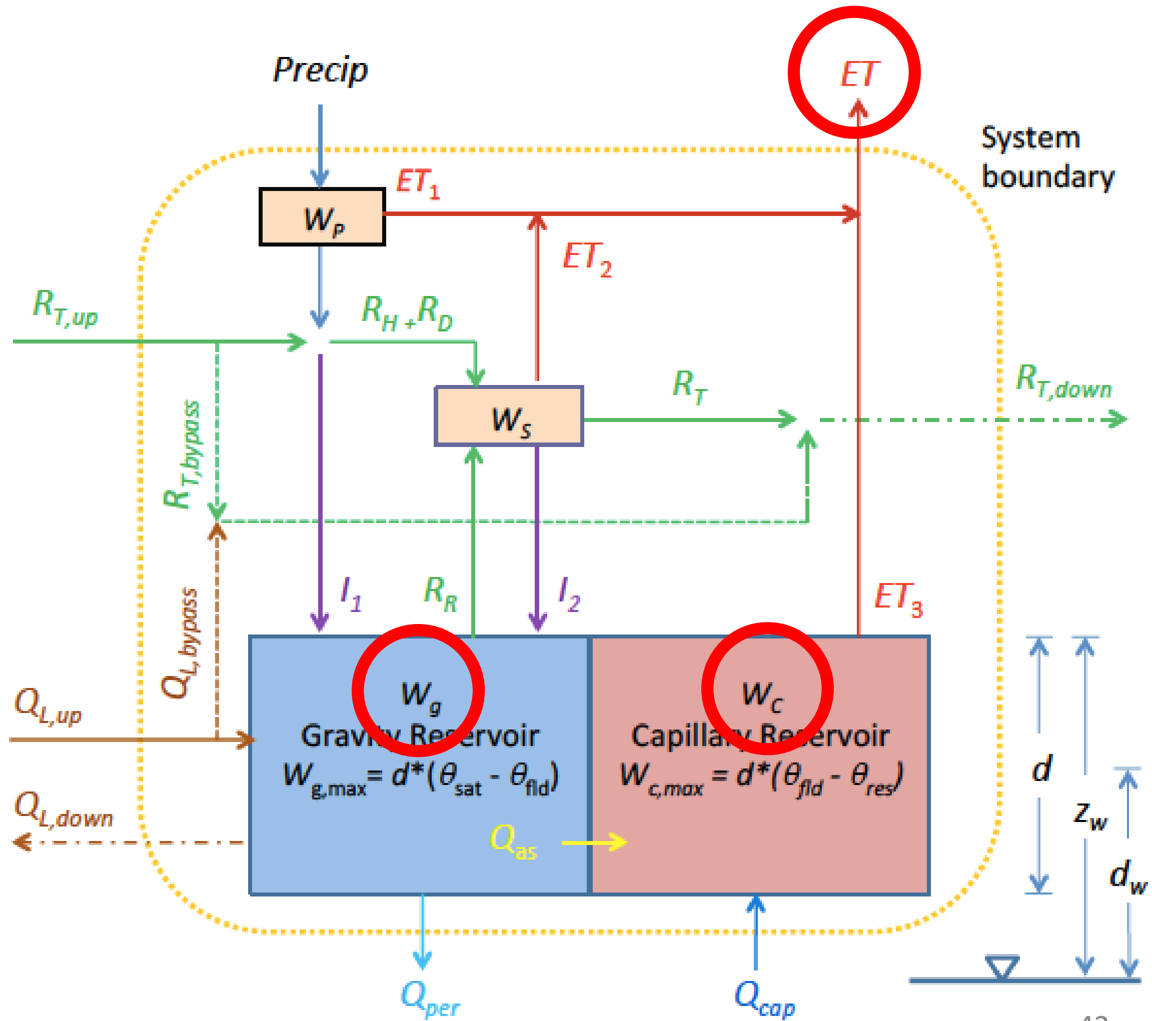


LST

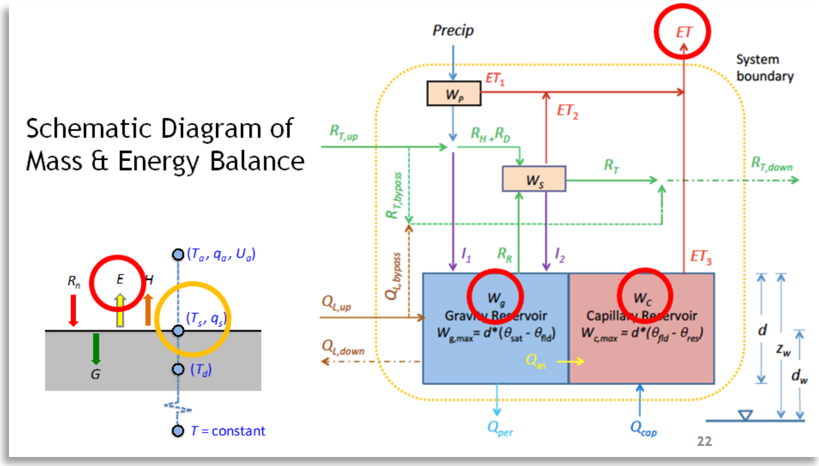
Variational assimilation of LST in surface water-energy balance

Schematic Diagram of Mass & Energy Balance





Reduction to a system of coupled ODE's (1D-VAR in time)



Castillo *et al.*, 2015, *WRR*, in press

$$\frac{dT_s}{dt} = F_1(\underline{T_s}, T_d, H(T_s, \dots), LE(T_s, W_c, \dots), \dots)$$

Observed & analyzed

$$\frac{dT_d}{dt} = F_2(T_s, T_d, \dots)$$

Coupling

$$\frac{dW_c}{dt} = F_3(\underline{W_c, W_g}, LE(T_s, W_c, \dots), \dots)$$

Analyzed

Cost function

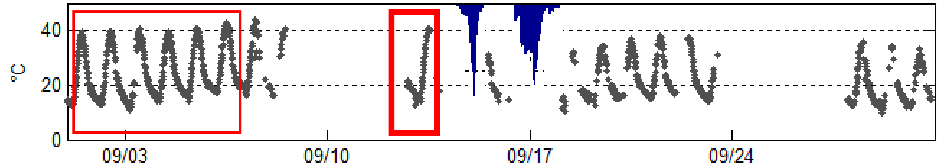
$$J = \int_{t_0}^{t_1} K_{Ts}\left(T_s - T_s^{obs}\right)^2 dt + K_{Wc}\left(W_c - W_c^{bg}\right)^2\Big|_{t_0} +$$

$$K_{Ts0}\left(T_s - T_s^{bg}\right)^2\Big|_{t_0} + K_{Td}\left(T_d - T_d^{bg}\right)^2\Big|_{t_0} + \int_{t_0}^{t_1} K_{Wg}\left(W_g - W_g^{bg}\right)^2 dt +$$

$$\int_{t_0}^{t_1}\left(\lambda_1\left(\frac{dT_s}{dt} - F_1\right) + \lambda_2\left(\frac{dT_d}{dt} - F_2\right) + \lambda_3\left(\frac{dW_c}{dt} - F_3\right) + \lambda_4\left(\frac{dW_g}{dt} - F_4\right)\right) dt$$

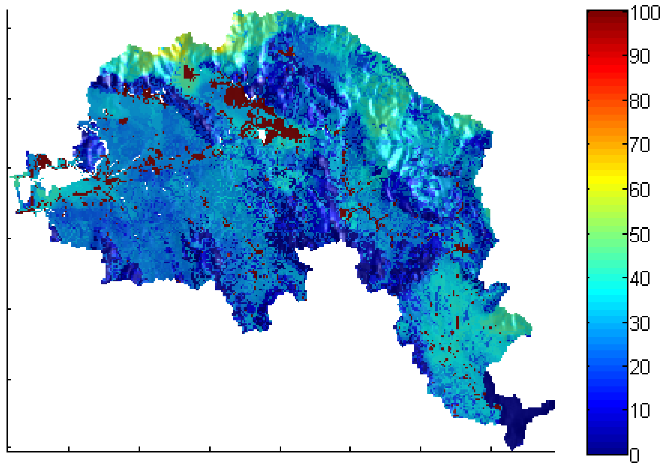
Weakly coupled
(soil moisture above
field capacity)

Assimilation of LST up to sept. 13th

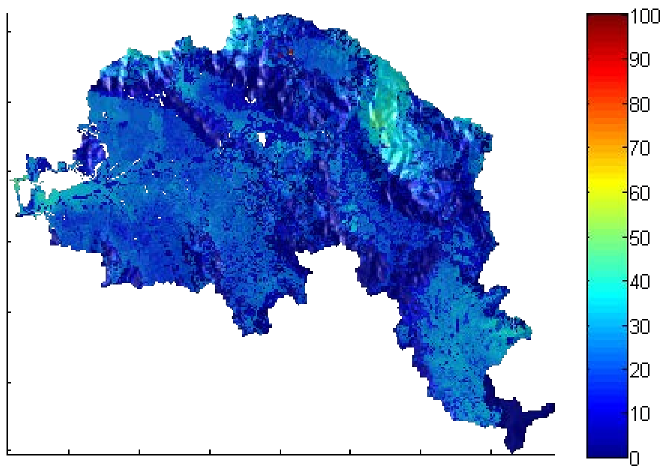


Soil saturation (%) (last day of assimil.)

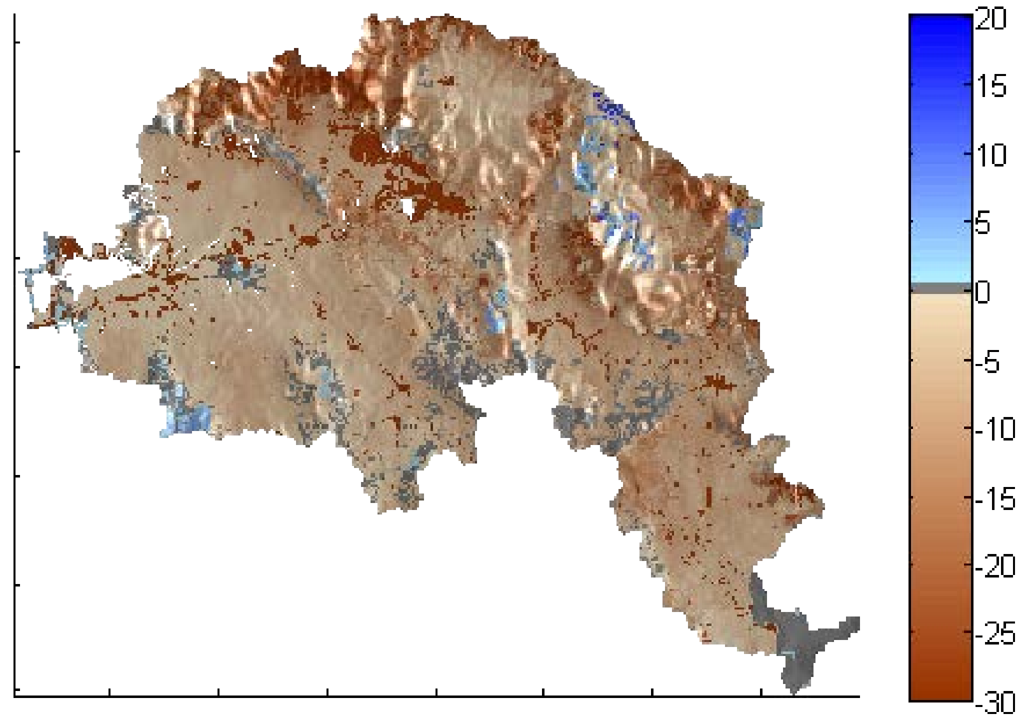
Background



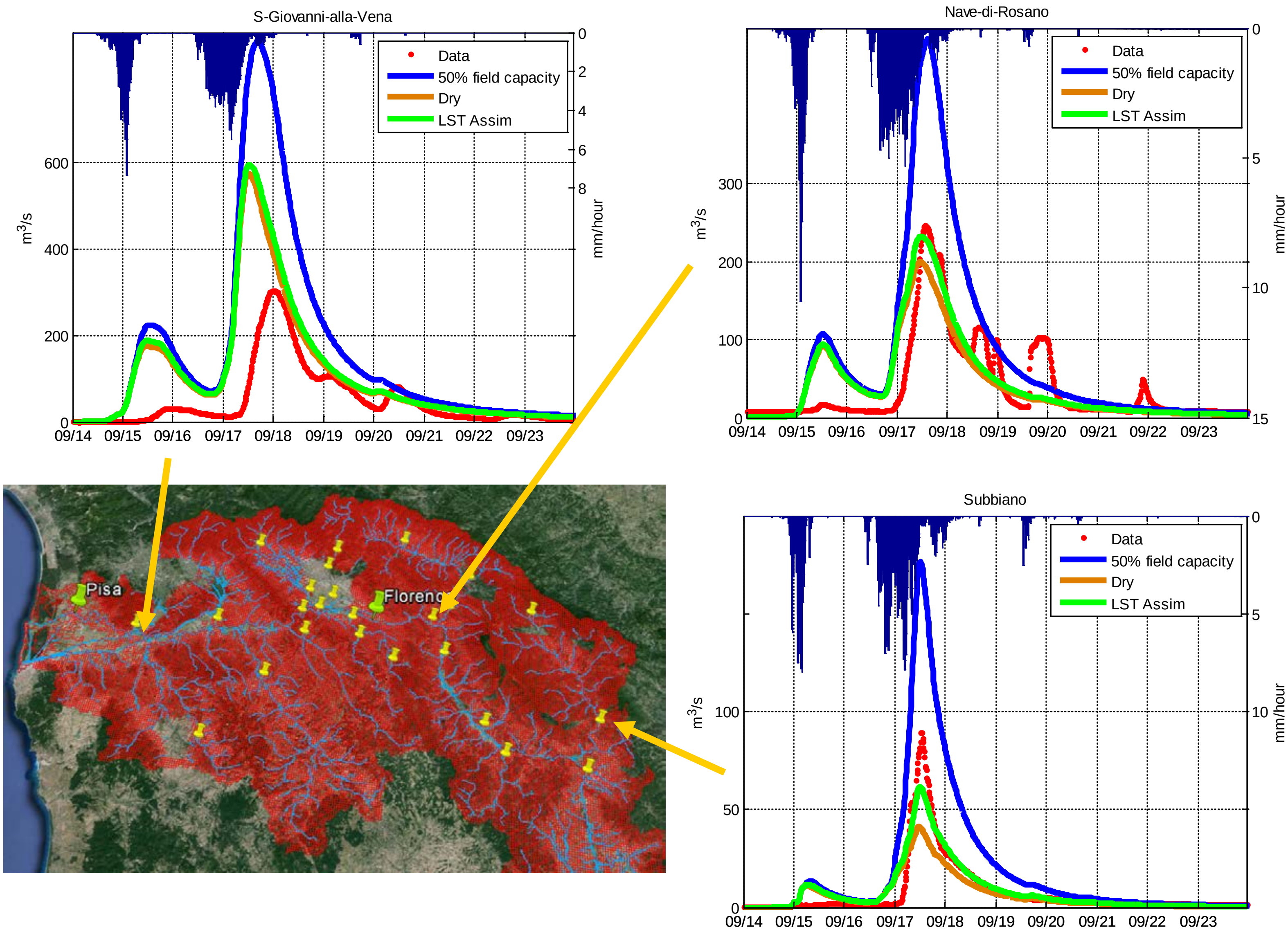
Analysis



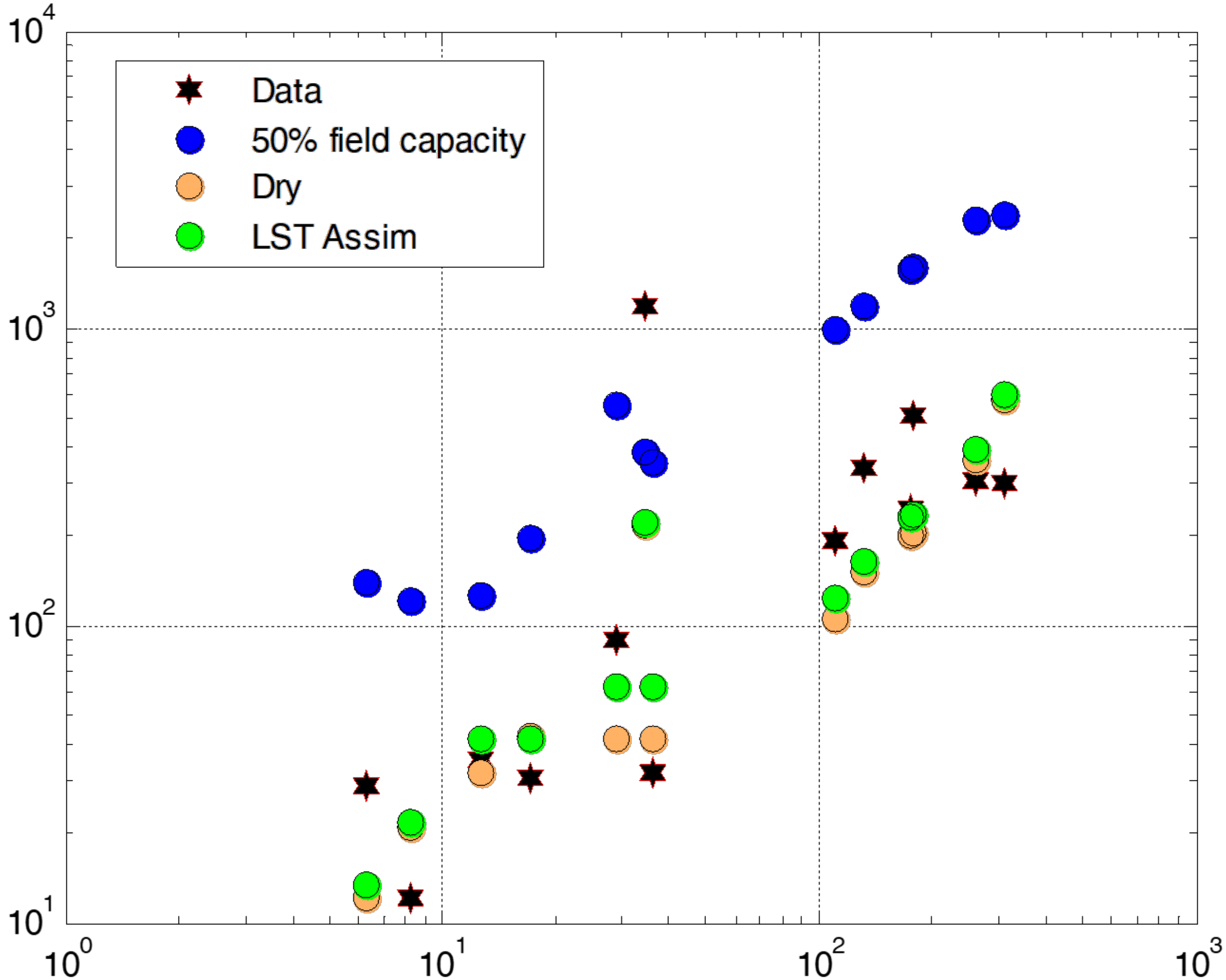
Analysis increment



Predictions at stremaflow stations, different initial soil moisture



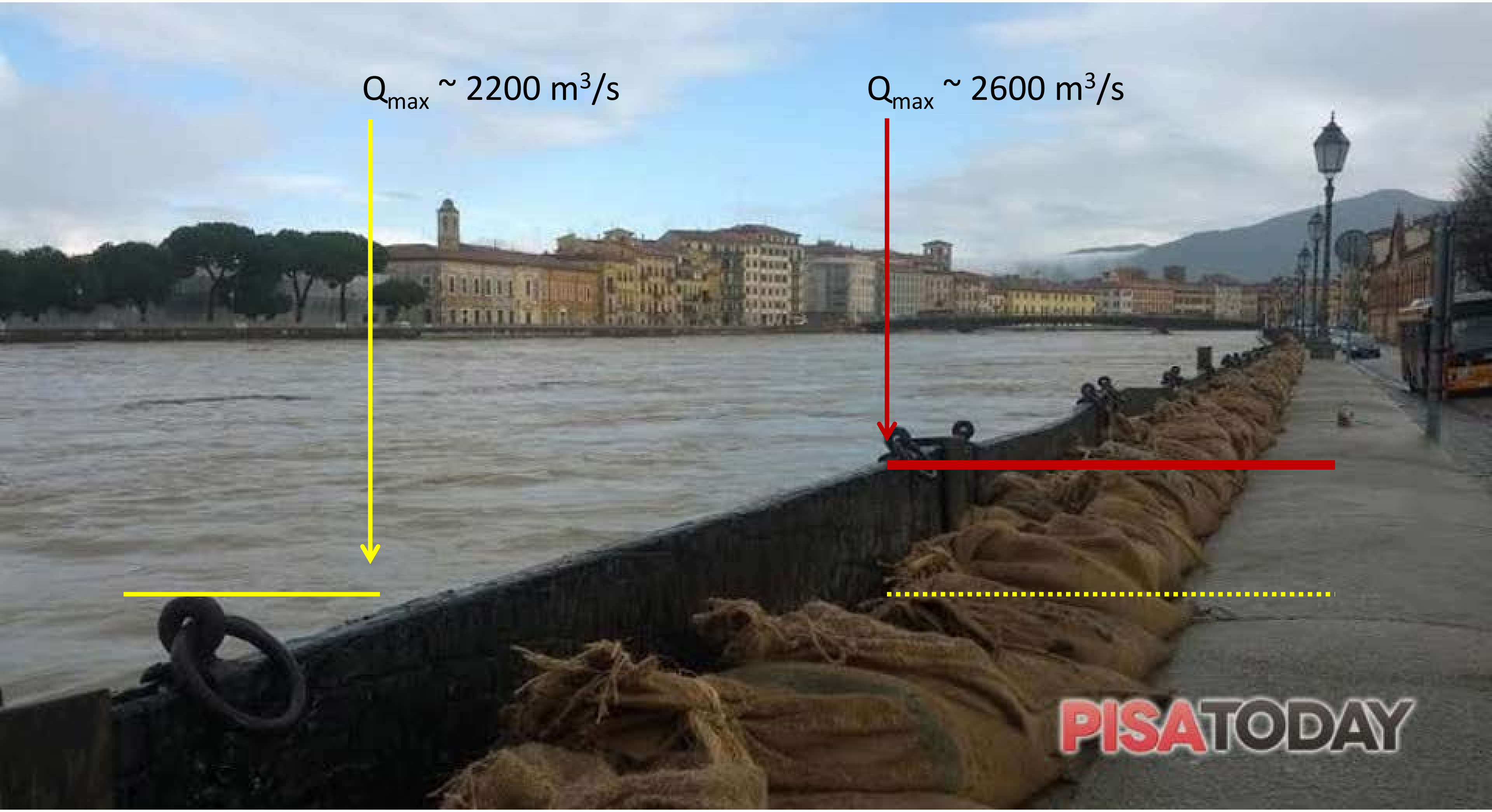
Predictions at stremaflow stations, different initial soil moisture conditions

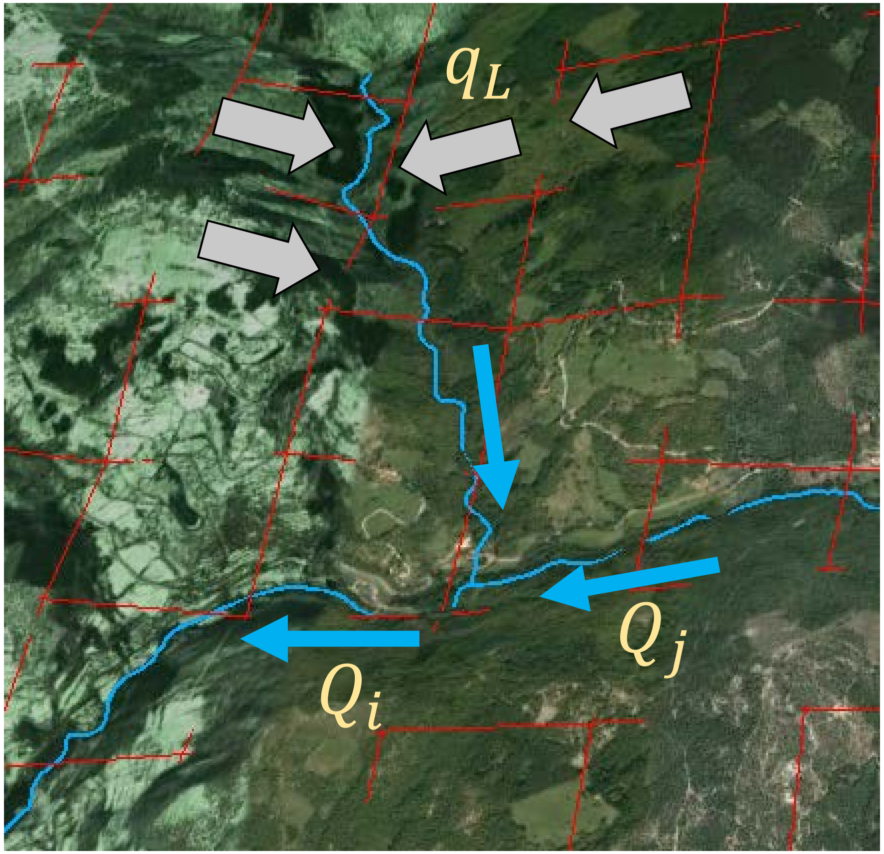


Peak discharge [m^3/s]

Precipitation volume[$10^6 m^3$]

‘Near-flooding’ event of february 2014





Assimilation of streamflow data for the analysis of hillslope runoff and river flow

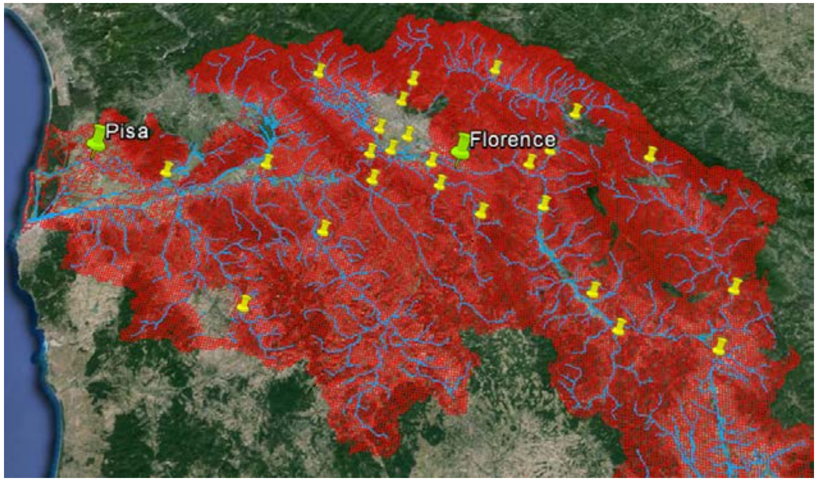
One key simplification with respect to other 'fluid flow' problems:

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

Cost function

$$J = \sum_{h=i,j,k} \left\{ \int_{t_2}^{t_1} \left[K_{Q,h}\left(\left(Q_h - {Q_h}^{obs}\right)^2 \right) + K_{qL,h}\left(\left(q_{L,h} - {q_{L,h}}^{bg}\right)^2 \right) + K_{c,h}\left(\left(c_h - {c_h}^{bg}\right)^2 \right) \right] dt + K_{Q_0,h}\left(\left(Q_{0,h} - {Q_{0,h}}^{bg}\right)^2 \right) \right\}$$

$$+ \int_{t_0}^{t_1} \lambda_j \left(\frac{dQ_j}{dt} - F_j\left(Q_j, q_{Lj}, \dots\right) \right) + \lambda_k \left(\frac{dQ_k}{dt} - F_k\left(Q_k, q_{Lk}, \dots\right) \right) + \lambda_i \left(\frac{dQ_i}{dt} - F_i\left(Q_i, Q_j, Q_k, q_{Li}, \dots\right) \right)$$

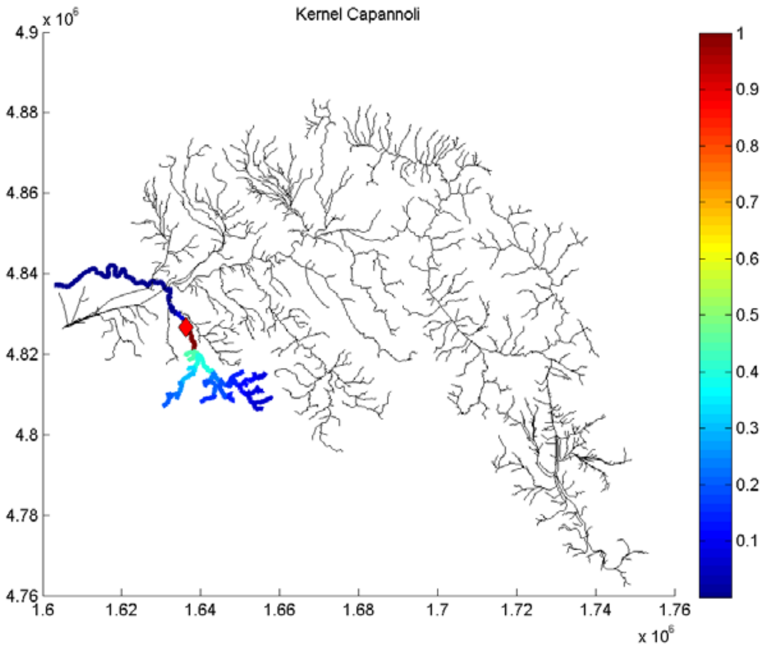


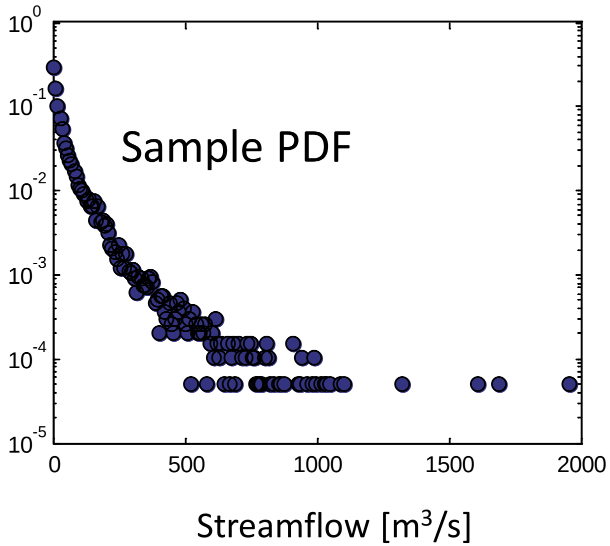
Assimilation of streamflow data for the analysis of hillslope runoff and river flow

Dealing with sparse observations along the river network

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

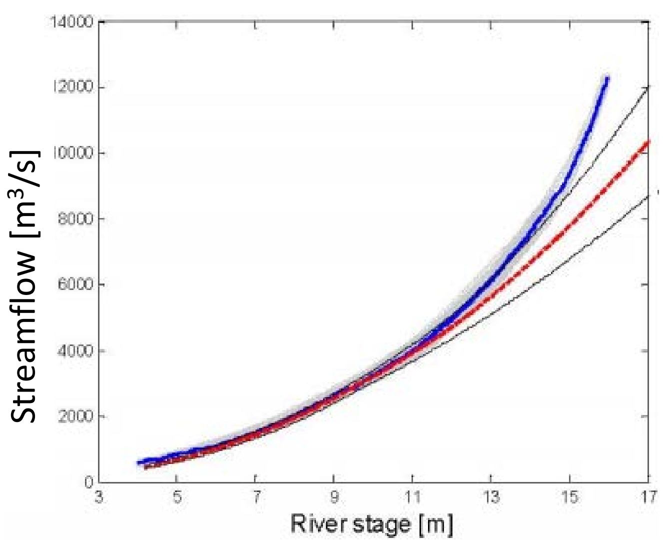
Dendritic Assimilation Kernel multiplying the error covariance



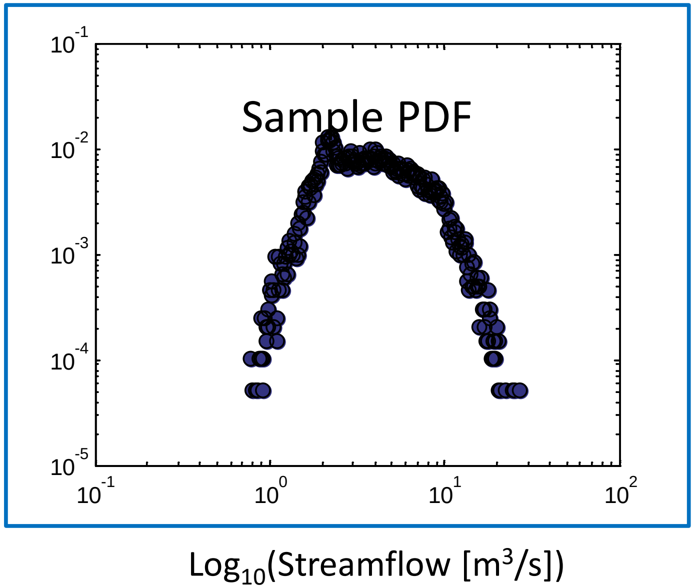


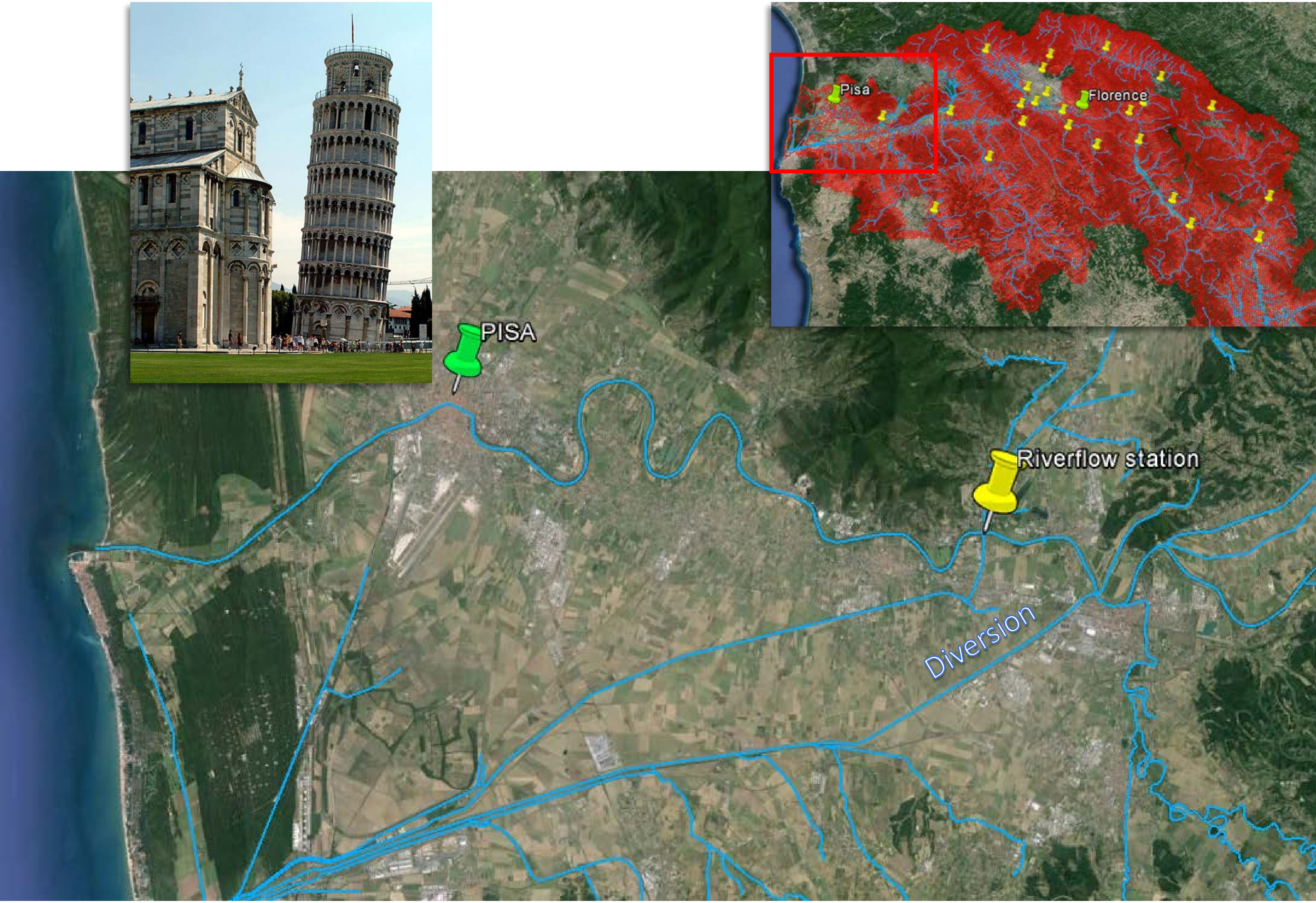
Strongly non-gaussian likelihood and multiplicative measurement errors

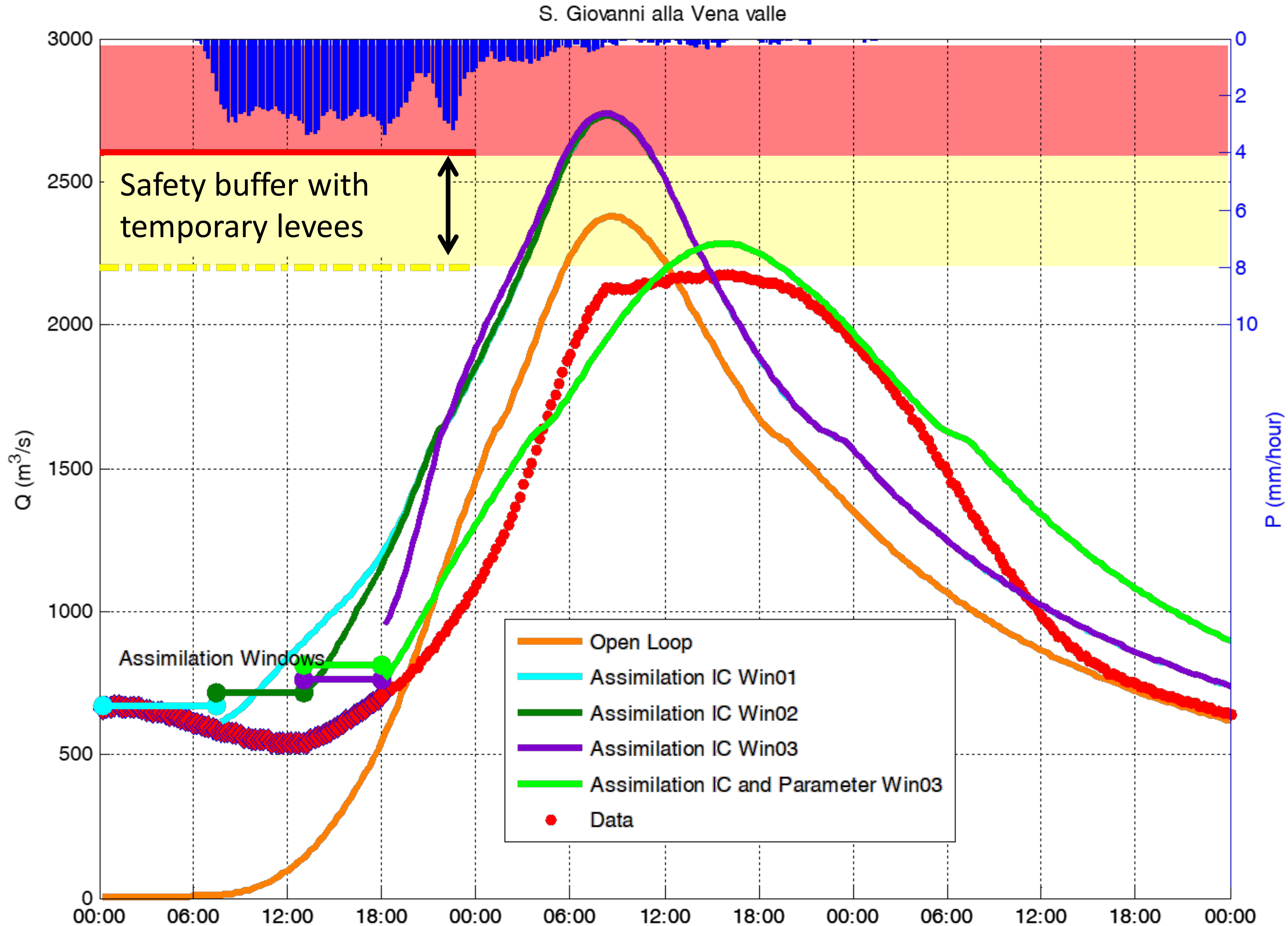
Assimilate and analyze the logarithm



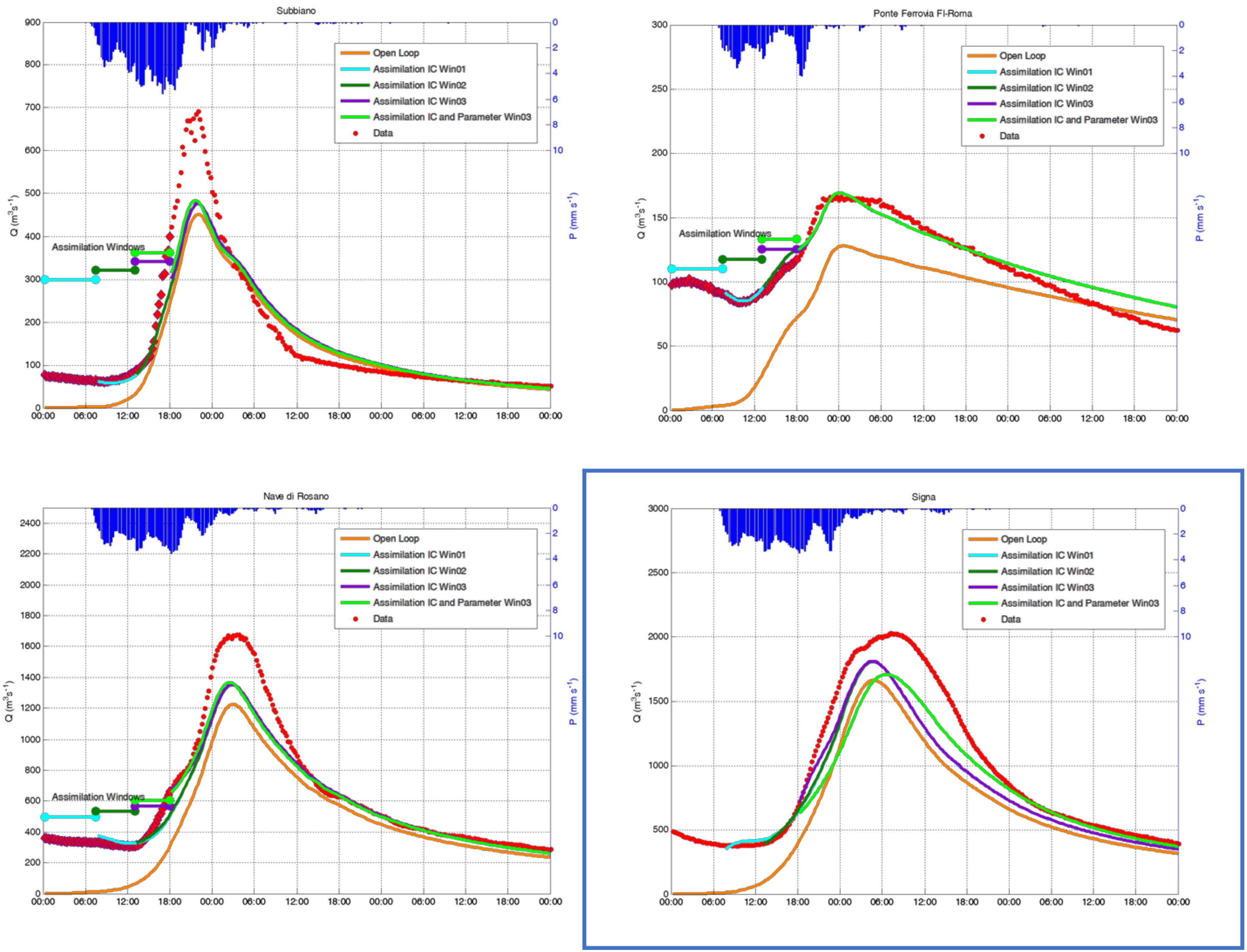
Domeneghetti *et al., Hydrol. Earth Syst. Sci.*, **16**, 2012




Pisa
Florence
PISA
Riverflow station
Diversion

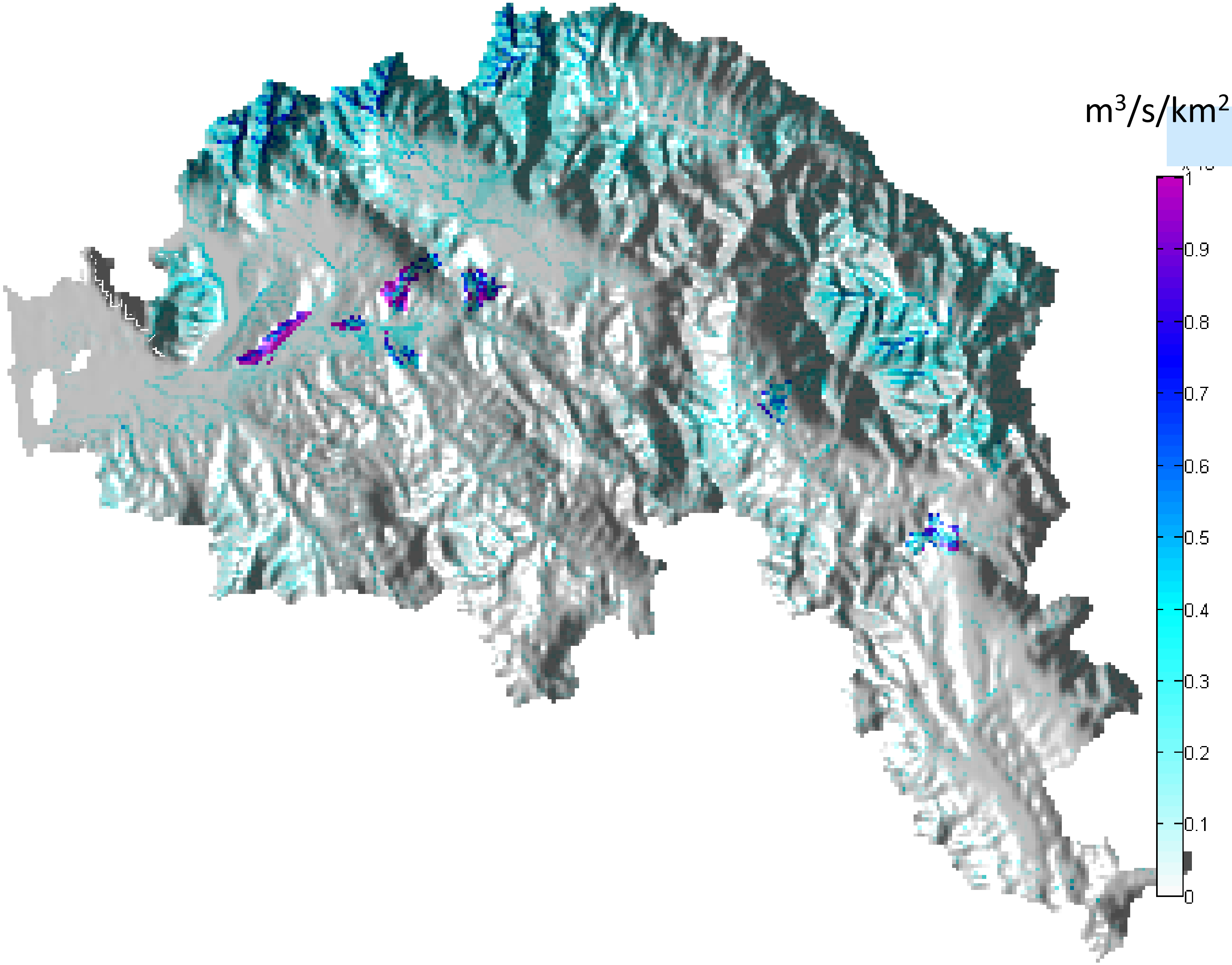


S. Giovanni alla Vena valle
Safety buffer with
temporary levees
Assimilation Windows
Open Loop
Assimilation IC Win01
Assimilation IC Win02
Assimilation IC Win03
Assimilation IC and Parameter Win03
Data
Q (m3/s)
P (mm/hour)
3000
2500
2000
1500
1000
500
0
0
2
4
6
8
10
00:00
06:00
12:00
18:00

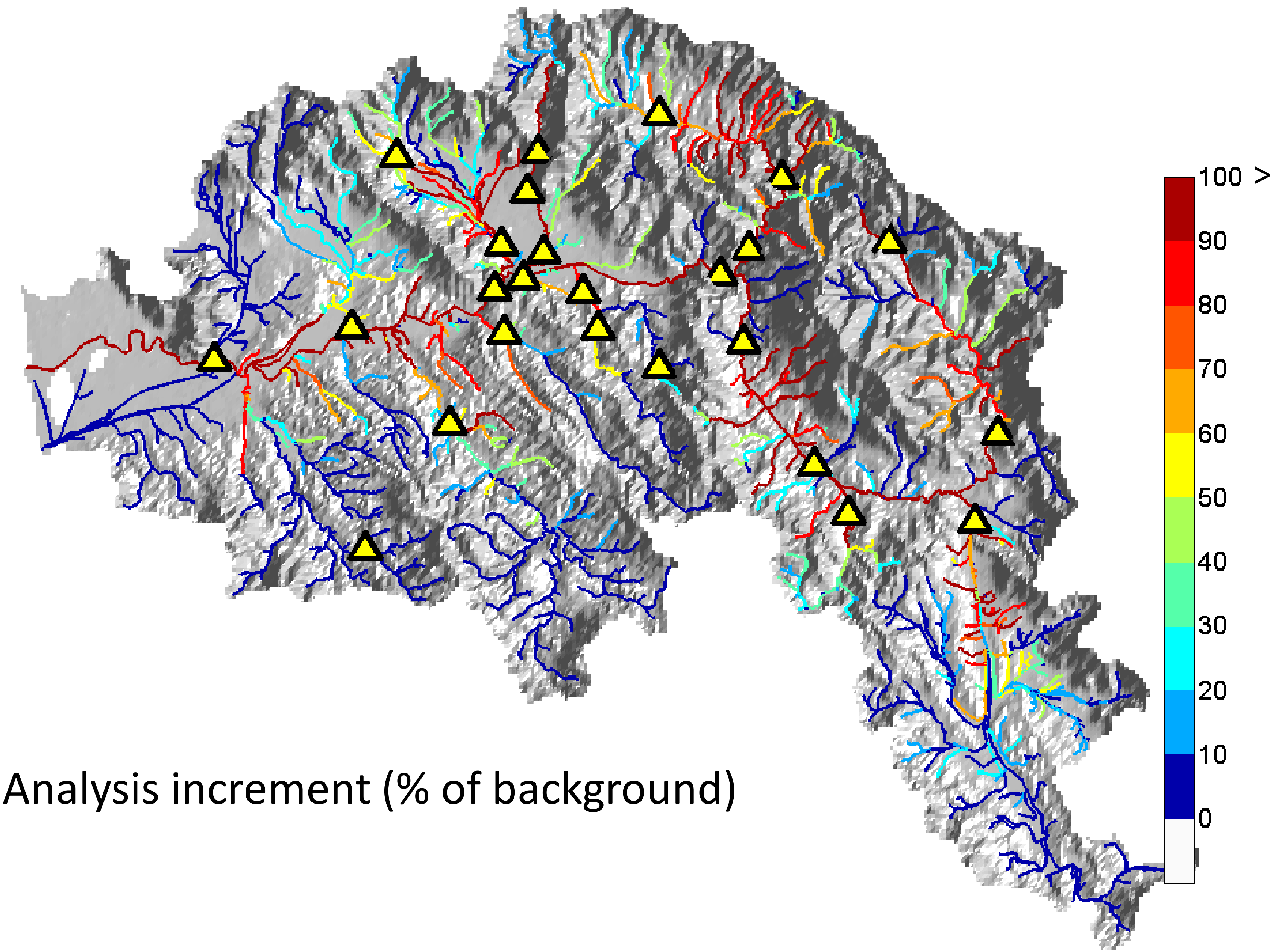


Subbiano
Ponte Ferrovia FI-Roma
Nave di Rosano
Signa
Open Loop
Assimilation IC Win01
Assimilation IC Win02
Assimilation IC Win03
Assimilation IC and Parameter Win03
Data
Assimilation Windows
Q (m^3s^-1)
P (mm s^-1)
00:00
06:00
12:00
18:00

Analysis of hillslope runoff at time of raising limb

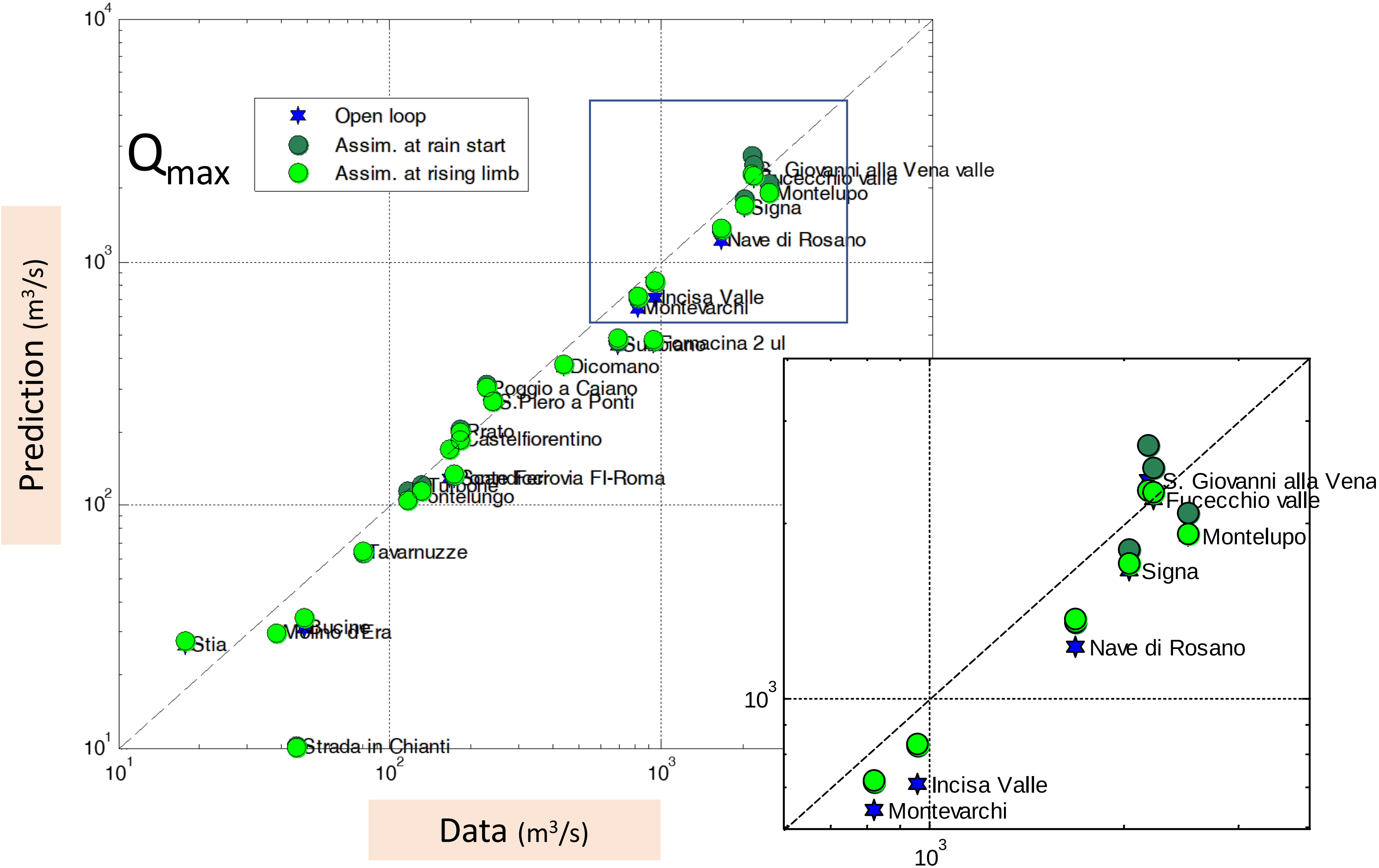


Analysis of streamflow at time of rainfall start

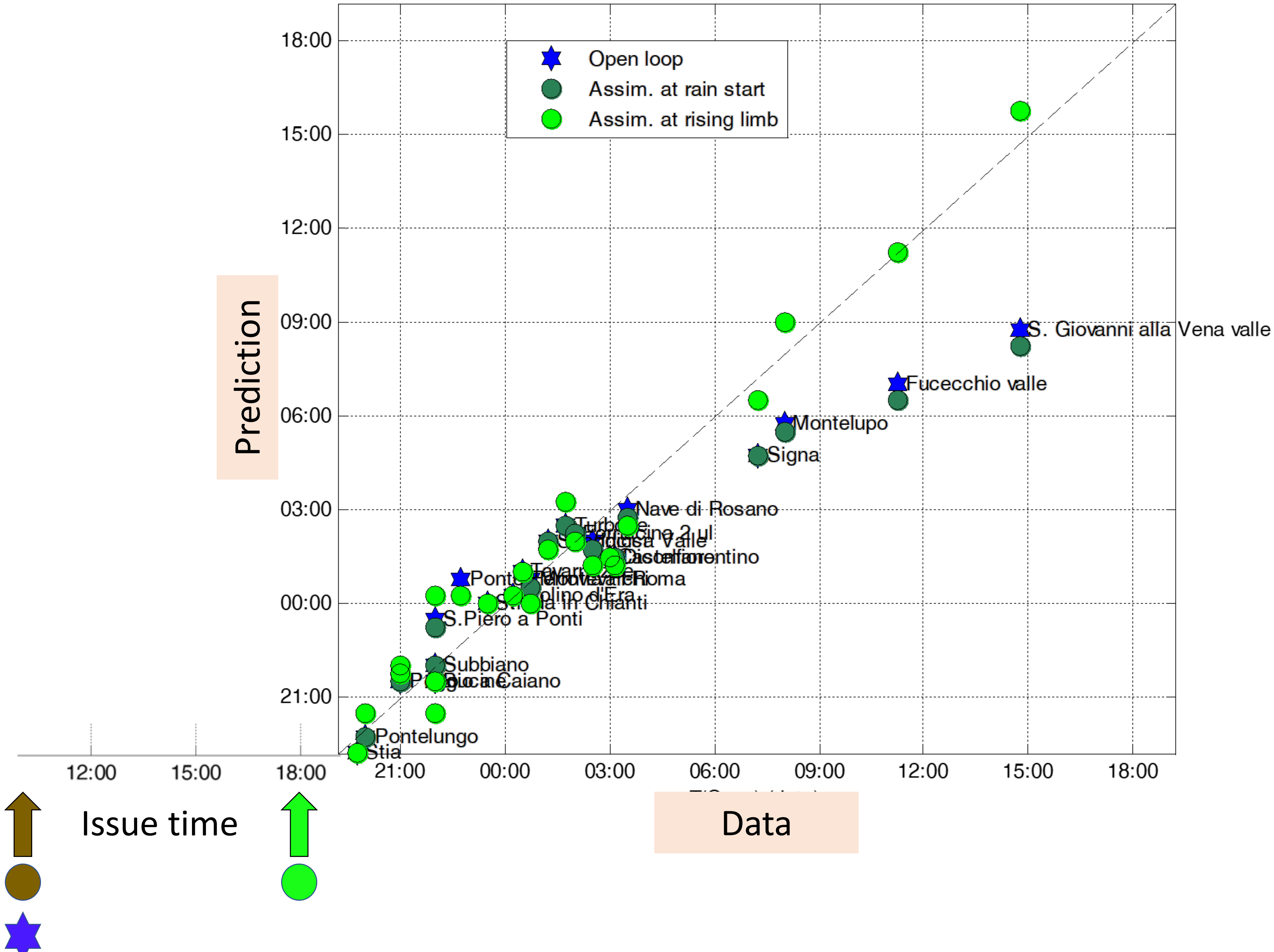


Analysis increment (% of background)

Prediction of flow peak value



Prediction of flow peak time



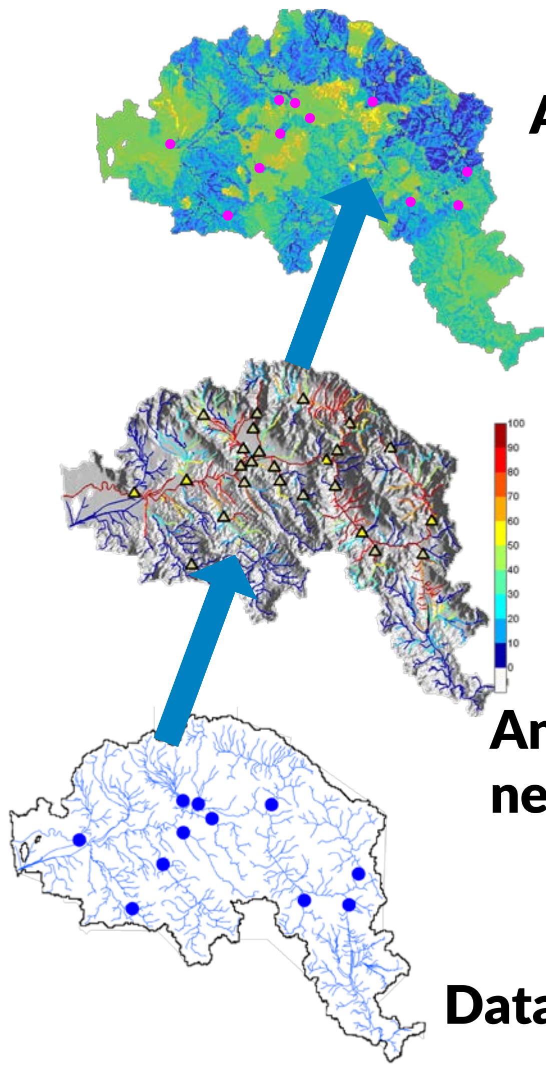
Water Resources Research

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



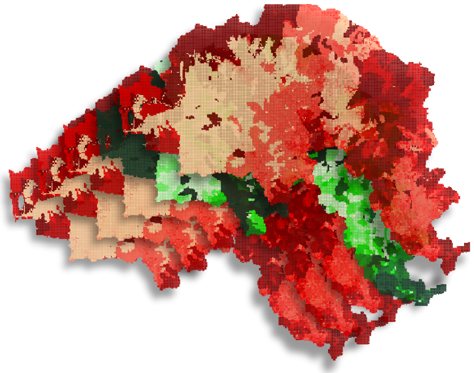
Analysis increment of hillslope runoff

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

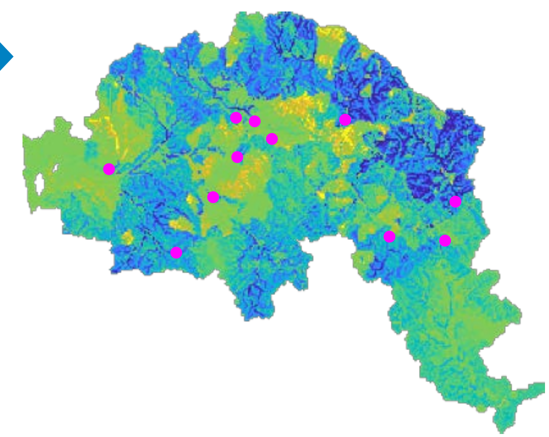
Assimilation scheme

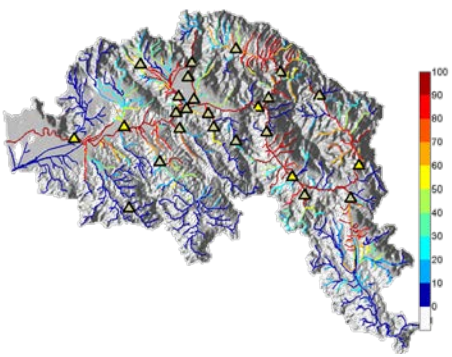


Montecarlo sampling of Antecedent Soil Moisture, rainfall interpolation

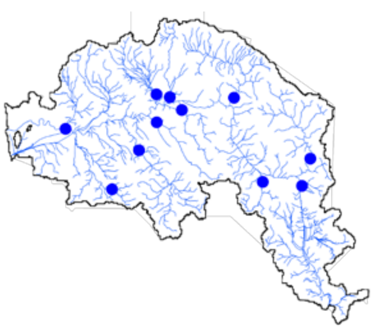
Likelihood

Analysis increment of hillslope runoff





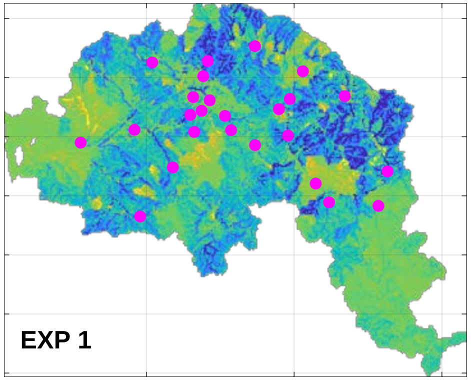
Analysis increment of river flows

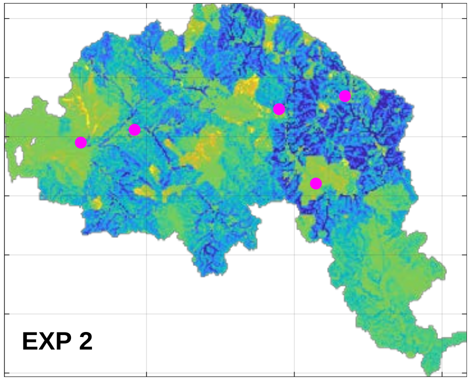


Variational assimilation with an adjoint

Hydrometric data

Hillslope runoff analysis increment (mm)



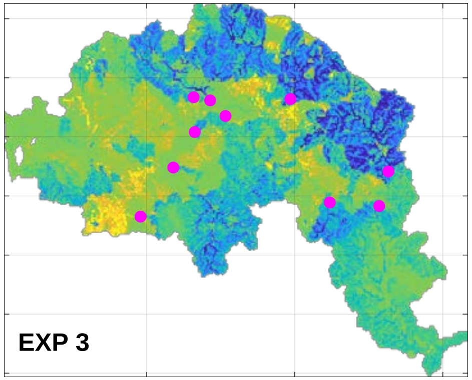


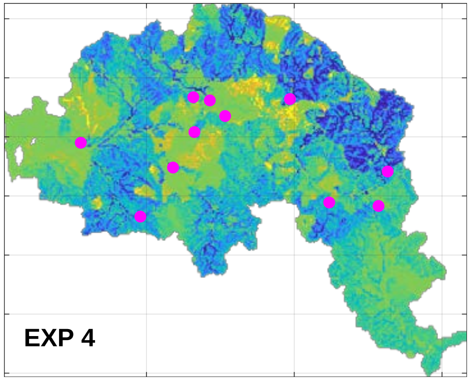
Exp1: assimilating all the synthetic stations

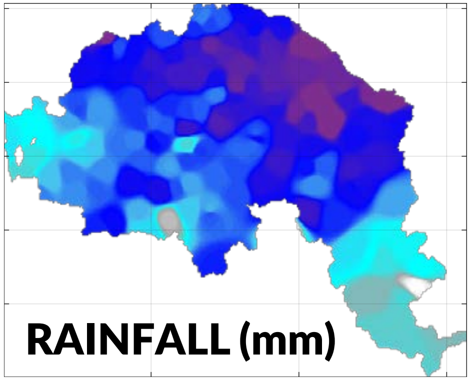
Exp2: assimilating along the mainstream, included basin outlet.

Exp3: assimilating close to main tributaries outlet

Exp4: assimilating close to main tributaries outlet & basin outlet.







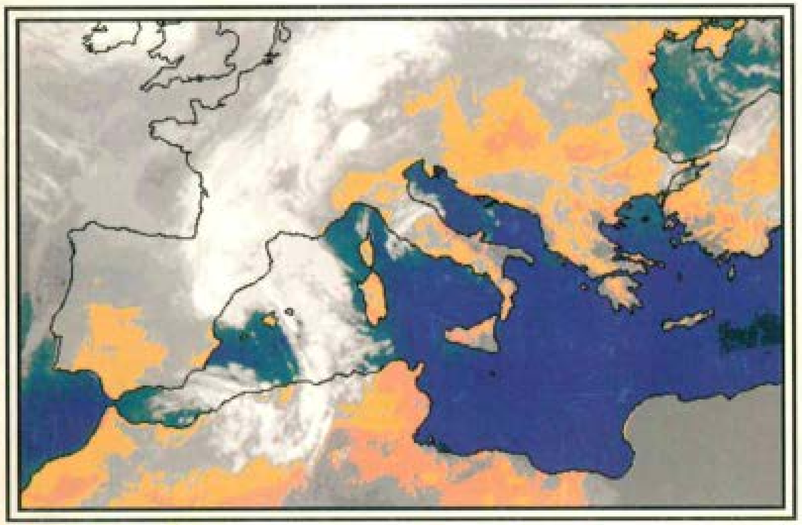
Ingredients
Technique
Tricks
....

... make you own recipe!

Selected Papers from Workshop
on

CLIMATE CHANGE AND HYDROGEOLOGICAL HAZARDS IN THE MEDITERRANEAN AREA

Colombella, Perugia, Italy June 27th-28th, 1994



Co-Conveners

Lucio Ubertini **Fabio Castelli** **Rafael L. Bras**

Organized by

Hydraulic Institute, University of Perugia

Water Resources Research and Documentation Center, University for Foreigners, Perugia

With the sponsorship of

National Research Council of Italy

Under the auspices of

Group for Prevention from Hydrogeological Disasters - National Research Council of Italy

Department of Civil and Environmental Engineering - Massachusetts Institute of Technology

Center for Global Chance Science - Massachusetts Institute of Technology