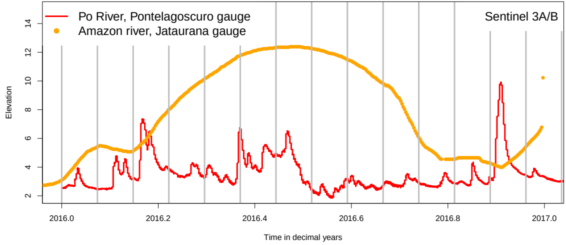
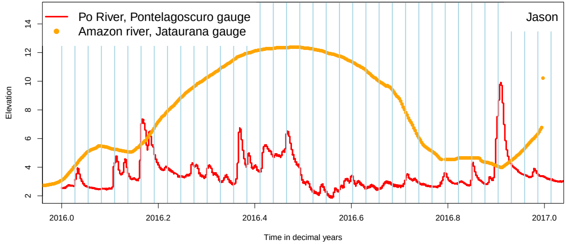
Reconstructing river water level time series from multi-mission satellite altimetry – reach based methods

Karina Nielsen, Elena Zakharova, Angelica Tarpanelli, Ole B. Andersen, Jérôme Benveniste and Luciana Fenoglio-Marc

DTU Space
National Space Institute

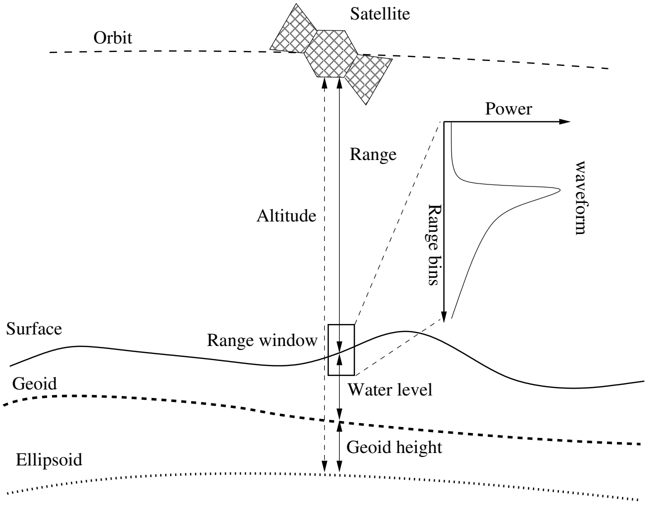
Motivation





- Single missions have a limited temporal resolution and might miss the signal
- How can we exploit missions in a geodetic orbit like CryoSat-2
- To increase the temporal resolution we must combine more missions
- Important for discharge estimation to have a good temporal resolution

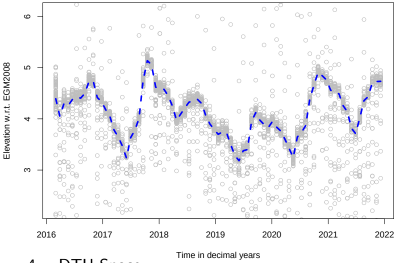
Satellite altimetry

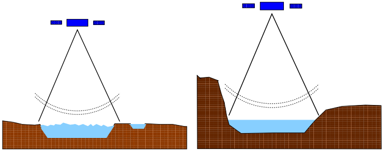


- Give us along-track point measurements of surface elevation H
- H = Altitude − Range   (−Geoid height)

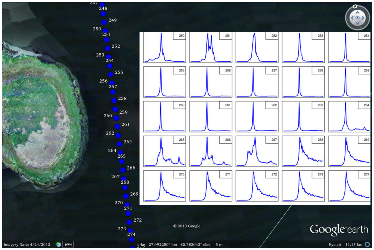
Satellite altimetry over inland water

- Challenges
  - The radar altimeter does not necessarily capture the nadir surface
  - This results in erroneous observations
- Satellite altimetry for inland water is a rapidly evolving field, many teams are working to improve.



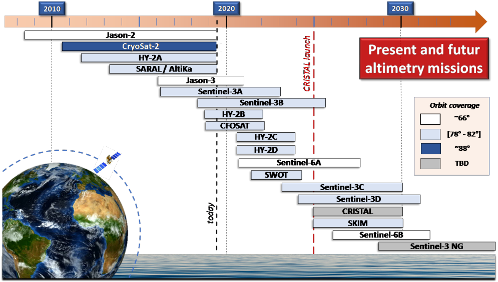


- Left: snagging, Right: topography



- The waveforms are influenced by the surroundings

Overview of satellite altimetry missions

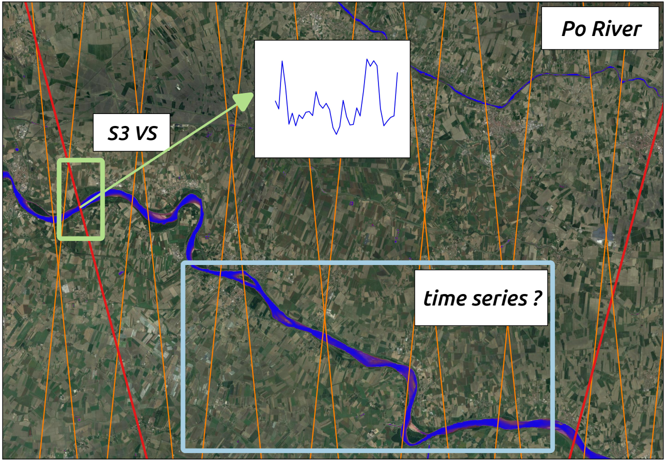


ICESat-2 ICE, CLOUD, AND LAND ELEVATION SATELLITE

- We are in a unique situation with several missions operation

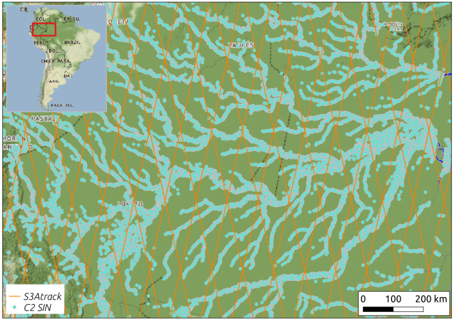
- S3A+S3B: repeat 27 days
- CryoSat-2: repeat 369 days
- SARAL: repeat 35 days, drifting after June 2016
- ICESat-2: repeat 90 days
- Jason: repeat 10 days

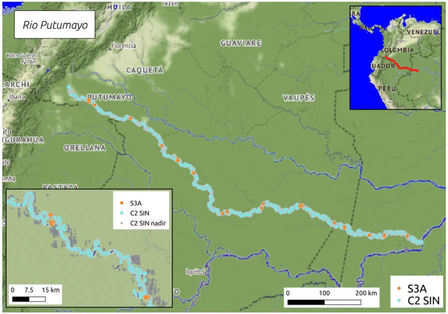
Deriving Water level time series from satellite altimetry



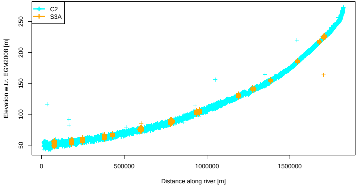
- For repeat missions water level time series can be derived at virtual stations
- How can we exploit geodetic mission?
- How can we combine different missions?

Why CryoSat-2 (geodetic missions) is amazing for Hydrology

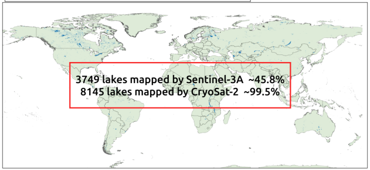




- Dense spatial coverage
- SARIn mode is highly beneficial for rivers
- Has fostered Several novel processing methods (new retrackers, FF-SAR,..)



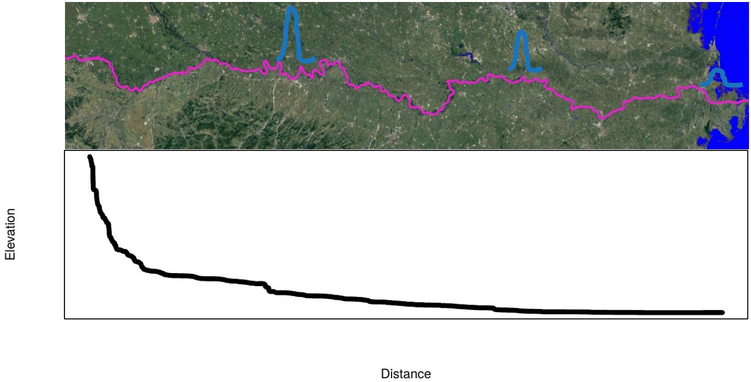
Subset of HydroLAKES (20-50000 km^2) - 8185 lakes



Similar studies by Others

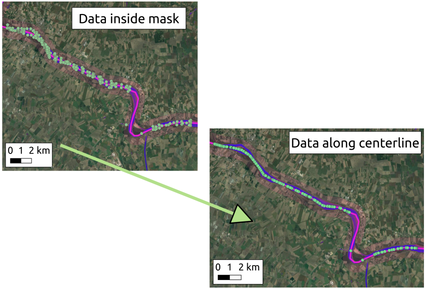
- Boergens, E., Buhl, S., Dettmering, D., Klüppelberg, C., & Seitz, F. (2017). Combination of multi-mission altimetry data along the Mekong River with spatio-temporal kriging. Journal of Geodesy, 91(5), 519–534. https://doi.org/10.1007/s00190-016-0980-z
- Tourian, M. J., Tarpanelli, A., Elmi, O., Qin, T., Brocca, L., Moramarco, T., & Sneeuw, N. (2016). Spatiotemporal densification of river water level time series by multimission satellite altimetry. Water Resources Research, 52(2), 1140–1159. https://doi.org/10.1002/2015WR017654
- Yoon, Y., Durand, M., Merry, C. J., & Rodriguez, E. (2013). Improving temporal coverage of the SWOT mission using spatiotemporal kriging. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 6(3), 1719–1729. https://doi.org/10.1109/JSTARS.2013.2257697
- ...

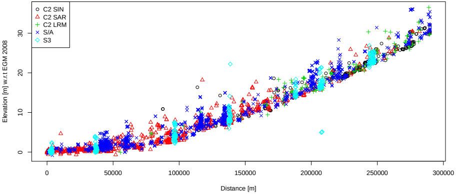
Challenges to consider



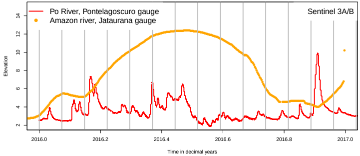
- Topography along the river
- The water level amplitude may change along the river
- How to deal with outliers?

View as a time-space problem





- By projecting the data to the center line of the river we can simply the problem
- The geodetic orbit of C2 is beneficial when mapping the river profile
- Add more missions to increase the amount of data
- More missions combined make it possible to achieve a better temporal resolution compared to the VS approach from repeat missions where details often are missed



River model, state-space model

River model observation part

$$H_i = \eta_{t_i}\alpha(x_i) + \tau(x_i) + \beta(sat_i) + \epsilon_i$$

- $\alpha(x_i)$ is a scaling factor given as a cubic spline assumed to be positive
- $\tau(x_i)$ is a cubic spline that describes the topography term assumed to be increasing as a function of distance.
- $\beta(sat_i)$ is a bias term depending on the satellite
- ϵ_i follows a normal distribution $\epsilon_i \sim \mathcal{N}(0, \sigma_\epsilon^2)$

River model process part (AR1)

$$\eta_{t_i} = \rho\eta_{t_{i-1}} + \xi_i, \quad -1 < \rho < 1, \quad \xi_i \sim \mathcal{N}(0, \sigma_\xi^2)$$

Implementation

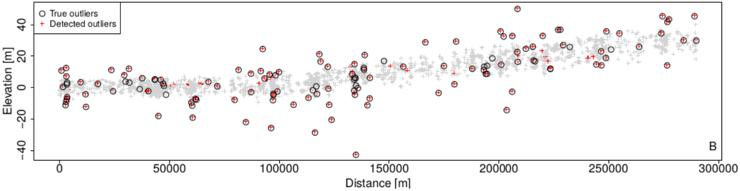
The model is implemented using the 'R' package TMB (Template Model Builder `https://github.com/kaskr/adcomp/wiki`)

Ensuring robustness

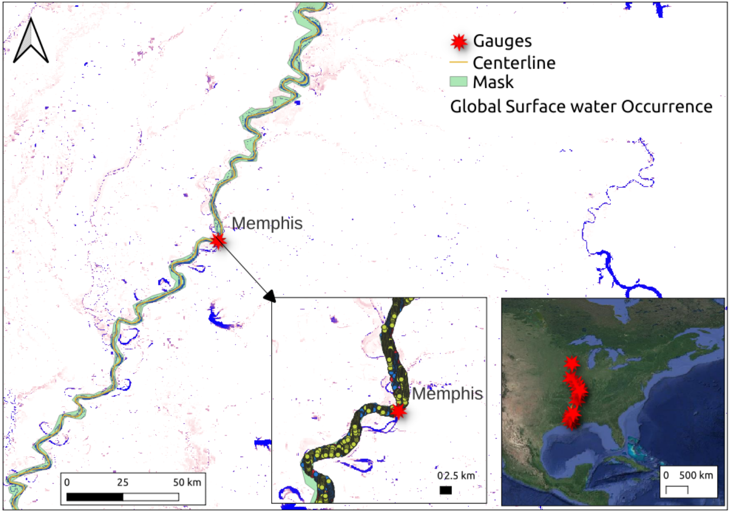
- Letting the error term ϵ_i follow a mixtures between a normal and a Cauchy distribution makes the solution more robust
- **Problem**: convergence problems (encounter ridge problem) :-(
- **Solution**: Apply weights iterative, convergence :-), and fast :-)

Weights:

- Compare predicted and observed river levels
- Down weight upper $p100$ percentile, where p is a small number below 0.1
- Estimate new river levels, ... repeat



Input data, Memphis example

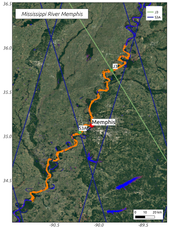


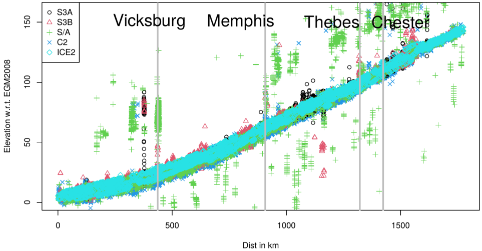
- Altimetry data, We apply water levels from CryoSat-2, Altika, and Sentinel-3A/B (now also ICESat-2)
- Auxiliary data, river centerline, mask to extract the data

Model input and data preparation

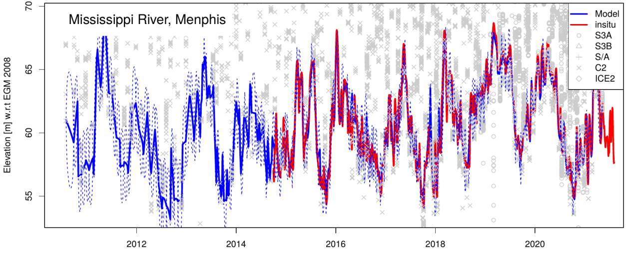
- Use a mask to extract observation over the river
- Water level positions are projected to the center line of the river
- Model input
  - Choose number of time steps for the joint solution N_t
  - Choose number of knots in the spline functions x_{knot}
  - The size of N_t and x_{knot} depends on the data
- The model allows evaluation of the water level at any given distance along the considered a reach of the river
- On the following result plots the water level is evaluated at the position of the gauge station

Example: Mississippi River around Memphis

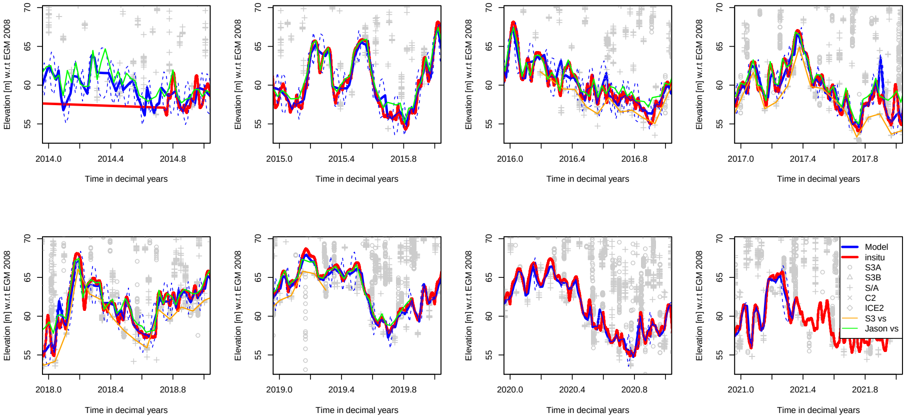




- Reach length=300km
- Model: $N_t = 700 \quad \sim$ 5 days and $x_{knot} = 5$

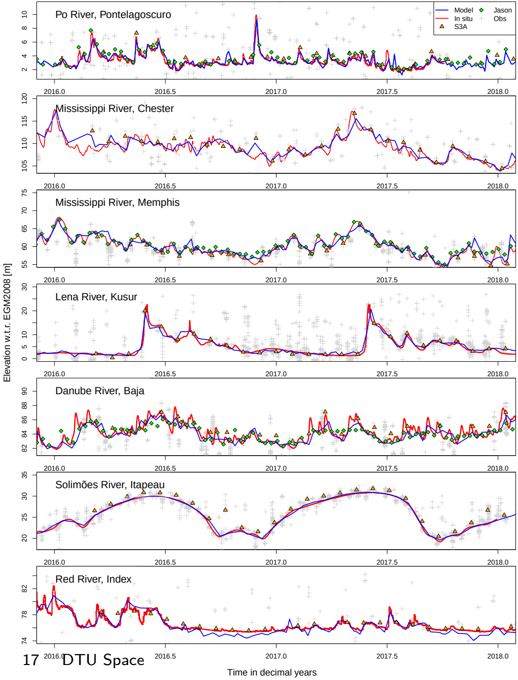


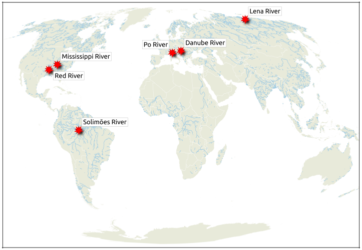
Reconstructed water level time series



- Similar detail as what is obtained with Jason
- More detail compared to Sentinel-3

Some additional results





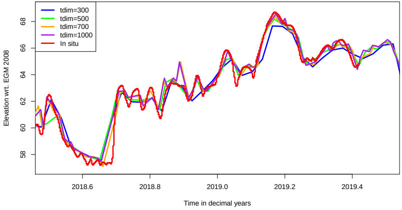
- The largest gain is seen for the smaller rivers
- If too little data is available the model might need to be simplified

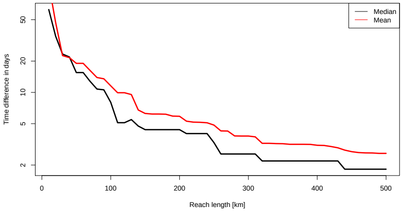
Validation statistics

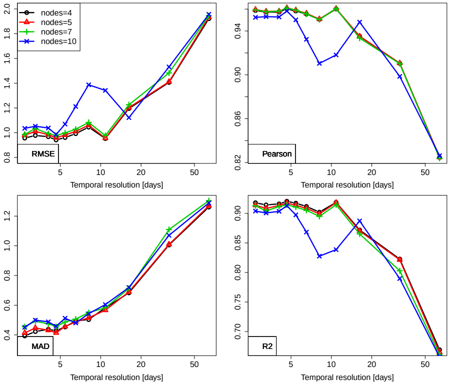
$River_{gauge}$	$t_{step,\ days}$	x_{dim}	RMSE [m]	MAD [m]	R	R^2	M
$Po_{Pontelagoscuro}$	1000, 3	10	0.70	0.33	0.91	0.80	3240
$Missisipi_{Chester}$	500, 7	4	1.10	0.45	0.94	0.87	3059
$Missisipi_{Memphis}$	700, 5	4	0.94	0.43	0.96	0.92	1706
$Lena_{Kusur}$	400, 5	7	2.53	0.66	0.88	0.77	2140
$Solimões_{Itapeau}$	200, 17	7	0.34	0.17	0.99	0.99	2555
$Danube_{Baja}$	600, 6	4	0.76	0.35	0.83	0.70	2891
Red_{Index}	500, 5	7	0.67	0.46	0.92	0.77	2415

- RMSE probably larger that other studies (see e.g. Scherer et al. 2020, `https://doi.org/10.3390/rs12172693`)

Model evaluation Memphis example







- Model output for different number of time steps
- Temporal resolution when reach length is varied

- Model performance as a function of the number of time steps and spline nodes
- The number of time steps and spline nodes depends on the available data. Experience show that the mean of the temporal resolution of the data is a good indicator

Benefits and limitations

Benefits

- We can exploit geodetic missions
- Combine missions to obtain an enhanced temporal resolution
- The water level time series can be constructed at any locations of the reach
- slope estimated can be derived

Limitations

- We need to take decisions regarding number of time steps and spline nodes, always better to be objective.
- The model assumes that the over all signal in the water level time series is the same
- Cannot handle dams and waterfalls along the reach
- The reconstructed time series can be view as a "scalable mean" for the given reach
- We do not account for any time lag

Future steps

- Create an R-packages to make the code more user friendly
- Apply the SWORD database, reaches and nodes, as input to make the workflow more dynamic
- Use other types of data in the model setup e.g. river width
- Estimate discharge

If you are interested in the details

- The code is available here `https://github.com/cavios/tsRiver`
    - Written in R via the R-package “TMB”
    - TMB is a tool to write non-standard models, fast minimization via automatic differentiation.
- Paper is available here `https://doi.org/10.1016/j.rse.2021.112876`

Thank you for your attention :-), questions?

An alternative reach approach

- We assume the observations follow

$$H_i \sim (\mu_i, \sigma^2(sat_i)), \quad \mu_i = u(t_i, x_i) + S(x_i) + \beta(sat_i)$$

- Here u is a Gaussian Markov random field, $S(x)$ is a cubic spline, β is a bias term

$$u \sim N(0, \sigma^2\Sigma), \Sigma = Q^{-1}$$

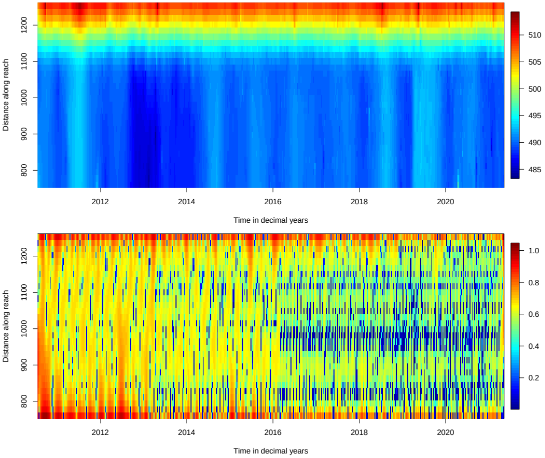
- The precision matrix Q can be defined as $Q = \phi_1 Q_{01} + \phi_2 Q_{02} + I$
- Here Q_0 specifies the neighbor structure and is given by

$$Q_{01}(i,j) = \begin{cases} \phi_1\#\text{neighbors in x direction}, & \text{if } i = j, \\ -\phi_1, & \text{if } i \sim j, \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

$$Q_{02}(i,j) = \begin{cases} \phi_2\#\text{neighbors in y direction}, & \text{if } i = j, \\ -\phi_2, & \text{if } i \sim j, \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

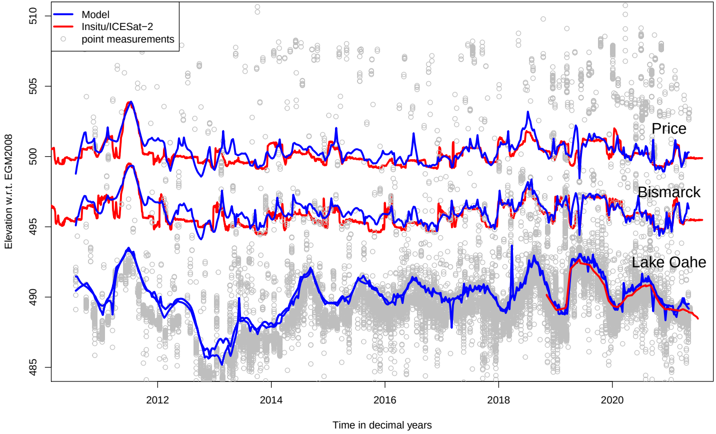
- Here ϕ_1 and ϕ_2 is the parameters that controls the correlation in the x and y direction, respectively

The estimated field and standard deviation



- Here vi have 500 time steps and 30 space steps – 15000 random effects

Validation of the model



- Comparison with gauges
- We can capture the signal of both the reservoir and river part of the reach
- We intend to test the approach on more reaches.