

Introduction

Time series models using predefined response functions are a fast, data-driven method to simulate head dynamics caused by various stresses such as recharge, pumping, or surface water fluctuations [1]. The explanatory stresses and the hydraulic heads contain errors originating in the observation process, which affect the simulated head dynamics and their uncertainty. Other factors that influence the performance of time series models are errors in parameter estimation, simplifications in the model concept, and numerical errors. All these factors affect the uncertainty of the model. The main objective of this study is to assess the impact of errors in the observations of the head and precipitation on the parameter estimates.

Noise in the precipitation series

The closest meteo station to Cabauw is De Bilt at ≈ 22 km distance. Figure 1 shows a density scatter plot of the observed values, the difference as a time series and the difference as a violin plot per month. Especially in summers large differences in observed precipitation occur. Over the period 2010 through 2022, the precipitation sum at De Bilt is slightly higher; 9436 mm (786 mm/year) for Cabauw and 10185 mm (849 mm/year) for De Bilt.

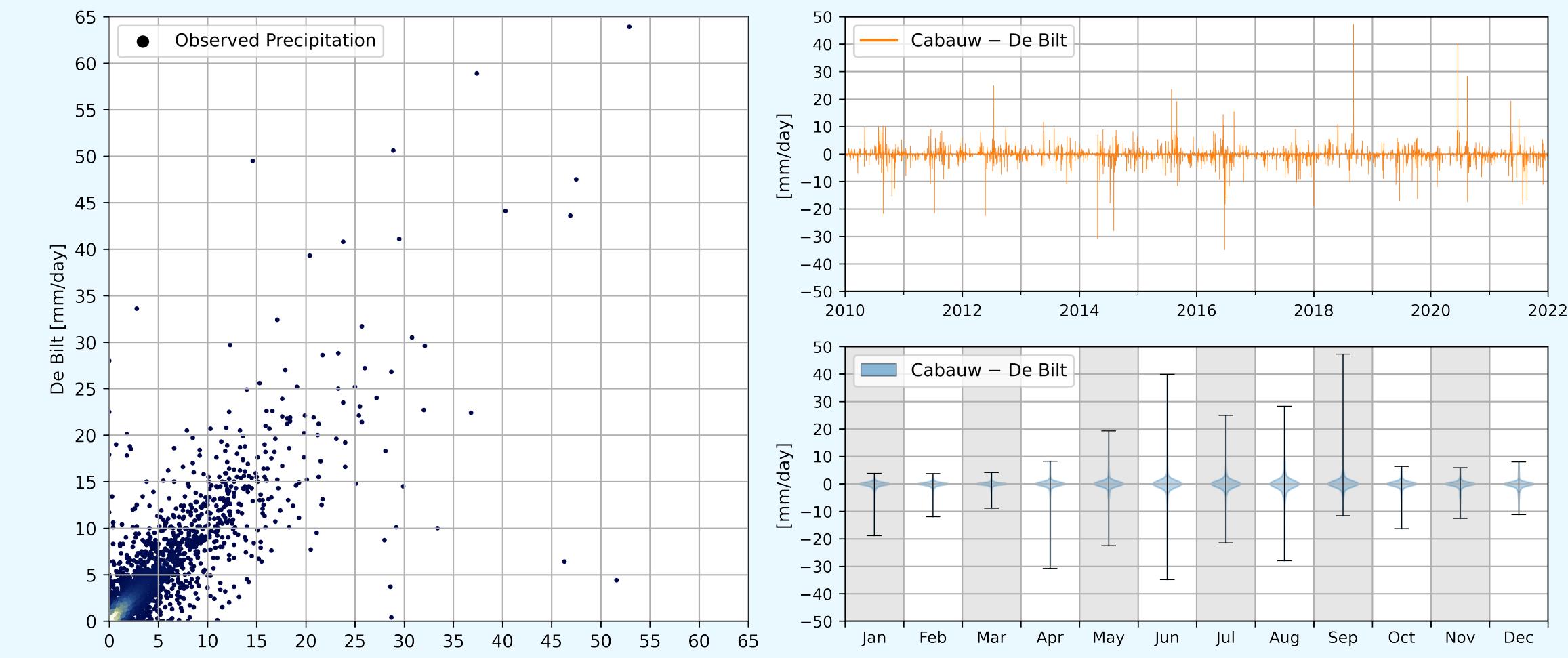


Fig. 1: Comparison between the observed precipitation from automatic weather station Cabauw and De Bilt (Royal Netherlands Meteorological Institute, KNMI)

Solving with PEST++ GLM examples

Correlation coefficients of the parameters from SciPy Least-Squares+AR(1) (upper triangle) and pestpp-glm (lower triangle) FOMS-approximate posterior covariance matrix.

Scenario 1: Cabauw without Noise		Scenario 2: Cabauw with Noise		Scenario 3: De Bilt without Noise		Scenario 4: De Bilt with Noise													
A	n	a	f	d	A	n	a	f	d	A	n	a	f	d	A	n	a	f	d
A 1.00 -0.47 0.84 0.77 -0.79	n -0.23 1.00 -0.81 -0.30 0.33	a 0.57 -0.76 1.00 0.57 -0.61	f 0.70 -0.07 0.30 1.00 -0.90	d -0.48 -0.33 0.20 -0.78 1.00	A 1.00 -0.35 0.65 0.84 -0.88	n -0.23 1.00 -0.85 -0.19 0.23	a 0.57 -0.76 1.00 0.45 -0.50	f 0.70 -0.06 0.30 1.00 -0.96	d -0.49 -0.33 0.20 -0.78 1.00	A 1.00 -0.42 0.83 0.80 -0.81	n -0.22 1.00 -0.82 -0.31 0.33	a 0.55 -0.75 1.00 0.60 -0.62	f 0.70 -0.07 0.29 1.00 -0.92	d -0.44 -0.37 0.28 -0.74 1.00	A 1.00 -0.42 0.78 0.84 -0.87	n -0.22 1.00 -0.82 -0.29 0.32	a 0.55 -0.75 1.00 0.60 -0.63	f 0.70 -0.07 0.30 1.00 -0.95	d -0.46 -0.36 0.26 -0.75 1.00

Solving with PEST++ iES example

The pestpp-ies algorithm runs 1000 ensembles with the prior parameters sampled within the provided parameter bounds. Ensembles with a bad fit, approximately 2.5%, are discarded. A total of 25 iterations is allowed to calculate a posterior distribution but less iterations are allowed if the objective function does not change much anymore between iterations. Noise with a standard deviation of 5 cm is added to the head observations to account for measurement error and model discrepancy [6]. The prior and posterior parameter ensembles, including the base realization and true parameter are shown in the upper row of figure 4. The lower row shows the 95% confidence interval of the parameter ensemble for each iteration.

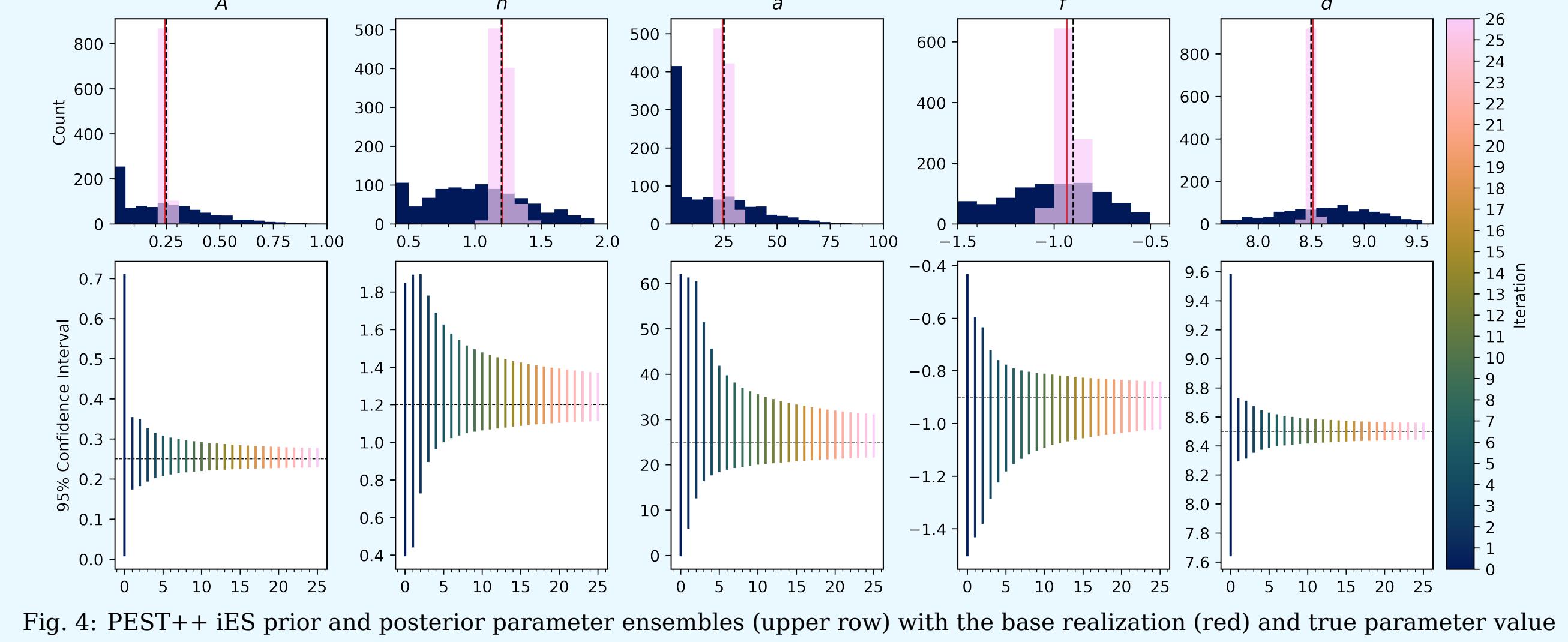
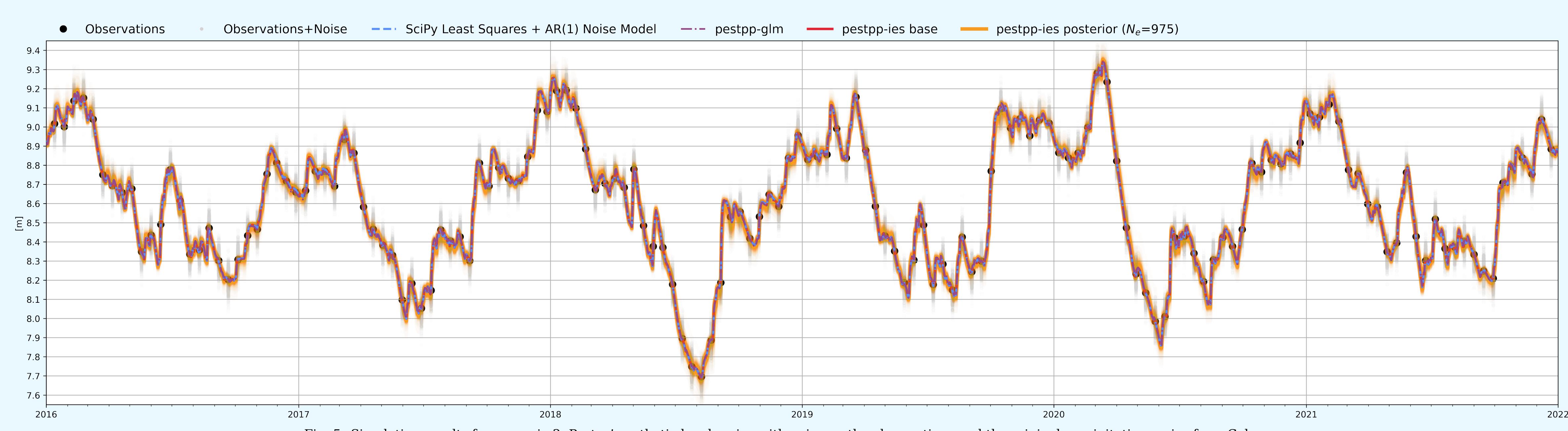


Fig. 4: PEST++ iES prior and posterior parameter ensembles (upper row) with the base realization (red) and true parameter value (black dashed) and the 95% confidence interval per iteration based on the quantiles.

Comparing parameter estimates



Four scenarios are tested where the parameters are estimated with SciPy+AR(1), pestpp-glm and pestpp-ies:

1. No noise on the head observations and the original precipitation series from Cabauw
2. Noise on the head observations and the original precipitation series from Cabauw (figures 4 and 5)
3. No noise on head the observations and the precipitation series from De Bilt
4. Noise on the head observations and the precipitation series from De Bilt

The synthetic head time series from figure 3 with a 14 day frequency is used for a period of twenty years, 2003 through 2022. If noise is added on the head observations, pestpp-glm and SciPy LeastSquares+AR(1) solve only for one noise realization while pestpp-ies has a noise realization for each ensemble.

Testing with realistic synthetic head series

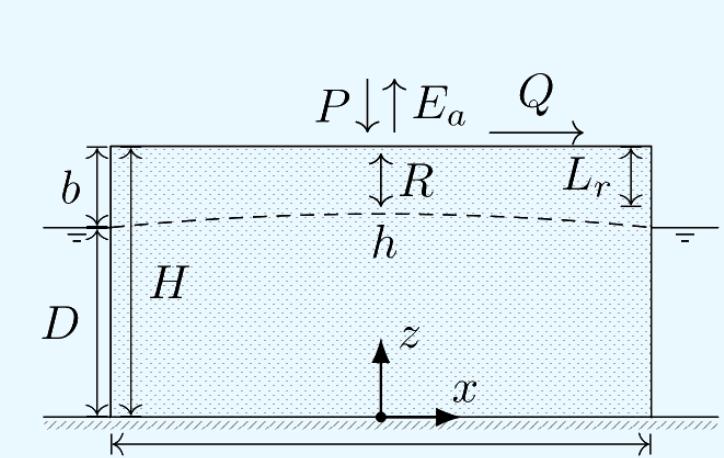


Fig. 7: Homogeneous aquifer between two parallel canals

It can be helpful to know the true Pastas parameters when analyzing synthetic head series as showed in this poster. However, it might be more useful to create realistic synthetic head series where physical processes that are deemed important are incorporated. MODFLOW USG-Transport [8] is an example of a groundwater model used that can be used to generate realistic synthetic heads [9]. It numerically solves Richards' equation for two- (and three-) dimensional flow in variably saturated porous me-

dia. The van Genuchten soil water retention curve, combined with Brooks-Corey's relative permeability function are used to define the soil hydraulic properties which results in remarkable numerical stability. Due to unsaturated zone processes the head response to precipitation can be highly nonlinear which has been tested for a homogeneous aquifer as seen in figure 7. Read more on this topic in Vonk et al. [9] by scanning the QR code!



doi: 10.1111/gwat.13403

Discussion

- Please note that the results on this poster are **preliminary**! We have more questions than answers at the moment.
- Possibilities for interaction between PEST++ and SciPy LeastSquares. SciPy calculates a jacobian for Pastas much faster than PEST++ after which the jacobian can be parsed to PEST++ as prior for iES or posterior to GLM for FOSM-based Monte Carlo).
- Many opportunities possible within the PEST++ framework for uncertainty estimation. Further exploration of the applicability of PEST++ on time series models is a topic for future research. For instance, PEST++ iES deals efficiently with large parameter samples. This allows for parameterization of the input stresses, i.e., estimating the precipitation error each day.

References

- Contact: m.vonk@artesia-water.nl
 Download: Download poster as pdf: <https://bit.ly/IAH2024>
- J.R. von Asmuth, M.F.P. Bierkens, and K. Maas. Transfer function-noise modeling in continuous time using predefined impulse response functions. *Water Resources Research*, 2002. doi: 10.1029/2001WR001136.
 - R.A. Collenteur, M. Bakker, R. Calié, S.A. Klop, and F. Schaars. Pastas: Open source software for the analysis of groundwater time series. *Groundwater*, 2019. doi: 10.1111/gwat.12925.
 - P. Virtanen, R. Gommers, T.E. Oliphant et al. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, 2020. doi: 10.1038/s41592-019-0686-2.
 - J.T. White, R.J. Hunt, M.N. Fienen, and J.E. Doherty. Approaches to highly parameterized inversion: PEST++ version 5, a software suite for parameter estimation, uncertainty analysis, management optimization and sensitivity analysis. *U.S. Geological Survey Techniques and Methods* 7-C2, 2020. doi: 10.3133/tm7C26.
 - J.T. White. A model-independent iterative ensemble smoother for efficient history-matching and uncertainty quantification in very high dimensions. *Environmental Modelling & Software*, 2018. doi: 10.1016/j.envsoft.2018.06.009.
 - M.N. Fienen, J.T. White, and M. Hayek. Parameter ESTimation with the Gauss-Levenberg-Marquardt algorithm: An intuitive guide. *Groundwater*, 2024. doi: 10.1111/gwat.13433.
 - R.A. Collenteur, E. Haaf, M. Bakker, T. Liesch, A. Wunsch et al. Data-driven modeling of hydraulic head time series: results and lessons learned from the 2022 groundwater modeling challenge. *Hydrology and Earth System Sciences*, 2024. doi: 10.5194/hess-2024-111.
 - S. Panday. *USG-Transport: Transport and other Enhancements to MODFLOW-USG*. GSI Environmental, 2024.
 - M.A. Vonk, R.A. Collenteur, S. Panday, F. Schaars, and M. Bakker. Time series analysis of nonlinear head response using synthetic data generated with a variably saturated model. *Groundwater*, 2024. doi: 10.1111/gwat.13403.

$$h(t) = \int_{-\infty}^t R(\tau) \vartheta(t-\tau) d\tau + d + r(t) \quad (1)$$

$$\vartheta(t) = \frac{A}{a^n \Gamma(n)} t^{n-1} \exp(-t/a), \quad (2)$$

$$R = P + f E_p \quad (3)$$

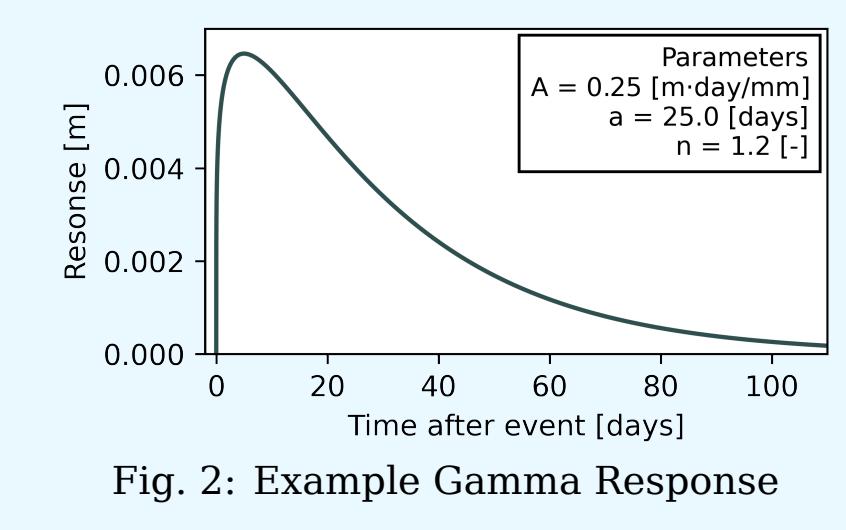


Fig. 2: Example Gamma Response

Pastas time series models

The Python package Pastas [2] uses predefined impulse responses to estimate the contribution of stresses to the hydraulic head fluctuations [1]. An important stress is the groundwater recharge, which can sometimes be estimated using a linear recharge model. The linear recharge model estimates the recharge as a linear combination of the precipitation and potential evaporation multiplied with a scale factor f (eq. 3). The time series model estimates both the recharge flux and the response of the groundwater table to the recharge. Different response functions can be applicable but in general the Gamma response function (eq. 2) is a versatile function that gives a good fit in many circumstances. Parameters of Pastas models are often estimated with the nonlinear least squares algorithm from SciPy [3]. To obtain an accurate estimate of the parameter uncertainty, a first order autoregressive (AR(1)) noise model is applied on the model residuals: $r(t_i) = u(t_i) + r(t_{i+1}) \exp\left(-\frac{\Delta t_i}{\alpha}\right)$. This results in one extra parameter α that is estimated along with the other parameters by minimizing the sum of squared innovations u . On top of that, the head observations are resampled to a lower frequency (generally 14 days) in order to remove autocorrelation in the innovations and obtain so called *white noise*.

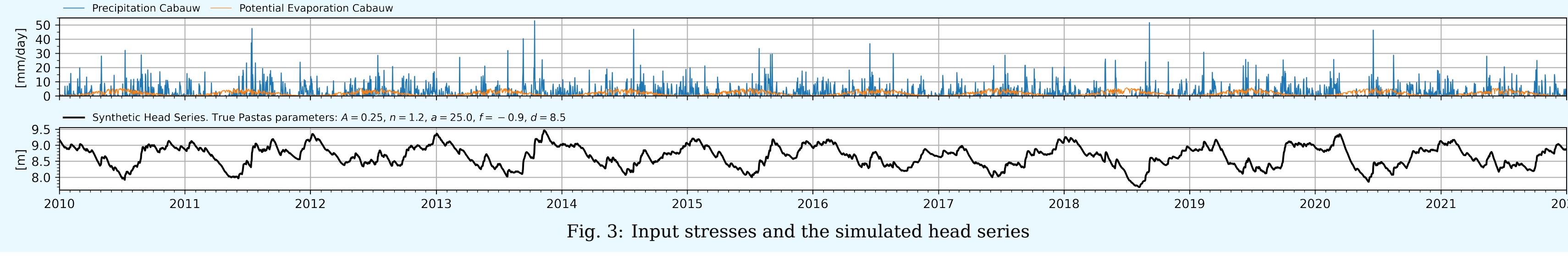


Fig. 3: Input stresses and the simulated head series