Response to RC1 for hess-2024-3989: Matthews, G., et al. Error-correction across gauged and ungauged locations: A data assimilation-inspired approach to post-processing river discharge forecasts

We thank the reviewer for their comments and suggestions which we believe will greatly improve the manuscript and strengthen the motivation for the method. The reviewer's comments have been summarised and numbered for clarity. The authors' responses are in blue. Line number, sections, and figures refer to revised manuscript.

1. The paper is difficult to follow due to dense mathematical notation and long, complex sentences.
   We have revised some of the notation in the paper. We note that, where possible, we already used standard mathematical notation for data assimilation following Ide et al. (1997) and have added this reference to Section 2 (line 93). We have reduced the number of symbols where possible, such as the symbol for the augmented observation operator and some superscripts. We have also removed the additional notation used to describe the approach for dealing with non-negative values and estimating the initial error ensemble (and instead describe these processes with words) as this notation was only used in Sections 4.1 and 5.3, respectively.

   We have shortened particularly complex sentences throughout.

2. There is an overuse of jargon without sufficient introductory explanation for a broader hydrology audience.
   We have removed unnecessary jargon, such as "spatiotemporal consistency". Where data assimilation technical terms were deemed by us to aid the description of the method, such as state augmentation, we have added a clear description of the term (see also comment 4).

3. The structure could be more concise, with a clearer division between methodology and results.
   Thank you for this comment. We have restructured Section 7 to make a clearer division between the results that show how the method works (Section 7.1) and the results that show the skill of the resulting ensemble (Section 7.2). Where possible we have removed repetition and unnecessary detail (see comment 12).

4. The state augmentation approach is described in a way that makes the approach seem unnecessarily complex.
   We have added the following sentences (lines 133-134): "*State augmentation is a technique used for online bias-correction in data assimilation that allows the simultaneous estimation of the system state and biases. An augmented state is defined by appending the biases to the state vector, allowing both to be updated by the data assimilation method.*" We have also moved the definition of the error ensemble to Section 2, streamlining the state augmentation description.

5. The assumption of constant error propagation is not well justified.
   We have added the following to Section 3.1 (lines 146-148): "As the true evolution of the error vectors at all grid-boxes is unknown, we assume a simple persistence model, such that $\mathbf{b}_{\mathrm{k}}^{(\mathrm{i})} = \mathbf{b}_{\mathrm{k}-1}^{(\mathrm{i})}$. This is a common assumption used in state augmentation (Pauwels et al., 2020; Rasmussen et al., 2016; Ridler et al., 2018; Martin, 2001)." We have also added a discussion of this assumption to Section 7 (lines 671-678).

6. Also, related to this, the use of precomputed model outputs instead of an evolving state might introduce additional errors, which are not sufficiently discussed.
   We have extended the discussion of this assumption on lines 214-217: "The assumptions made in Eqs. (18) and (19) make our system sub-optimal from a data assimilation perspective but are necessary to avoid rerunning the hydrological model. Importantly, we aim to estimate the error of the precomputed model output at each lead time. Therefore, while the lack of state evolution makes the hindcast component update sub-optimal, the update of the error ensemble remains mathematically consistent".

7. Inflation:
   We have restructured Section 5.2 to address the following comments.
   i I find the inflation method to be heuristic with little to no mathematical rigor. For instance, the assumption that the hindcast variance is a proxy for error growth does not account for potential biases in the raw ensemble itself.
   The reviewer is correct that the covariance inflation technique used is a heuristic method. We initially tried a simpler multiplicative inflation but found that this was not suitable due to the large variations in hindcast spread as a function of lead-time.
   Our new approach is a practical solution to the issue of filter divergence that is inspired by blending techniques such as the RTPP method (Zhang et al., 2004). We have added the following sentence to Section 5.2 to make this clearer (lines 278-281): "We implement a heuristic covariance inflation method inspired by the *relaxation-to-prior perturbations* technique (Zhang et al., 2004; Kotsuki et al., 2017). However, as we are working within a post-processing context, we adapt the method for use with predefined ensembles (i.e., without evolving the inflated perturbations between timesteps)."

   The limitations of using the hindcast uncertainty as a proxy for the uncertainty in the error estimate are discussed in the discussion section (lines 636-647).

   ii Unlike RTPP, the proposed inflation blends analysis and "estimated" perturbation information without explicitly evolving them. What motivates such an approach?
   Our goal is to use this approach for post-processing with pre-computed ensemble hindcasts. Explicit online evolution of the inflated perturbations (as is required in the traditional RTPP method) is not feasible in this scenario. Blending the ensemble perturbation matrix with an estimated perturbation matrix is inspired by palaeoclimatological reanalysis work such as Valler et al., (2019), where a climatological error-covariance matrix is blended with the ensemble error-covariance matrix. We add the following justification to Section 5.2 for our choice of estimated matrices (lines 290-292): "During development, it was found that the estimated matrices must have spatial structures consistent with the river network and be forecast and lead-time-dependent. For simplicity, and as the raw hindcast perturbations satisfy these requirements, we set both $\mathbf{X}_{k+1}^{\mathrm{est}}$ and $\mathbf{B}_{k+1}^{\mathrm{est}}$ equal to the raw hindcast perturbation matrix (Dee, 2005; Martin et al., 2002)."

   iii The inflation parameter, alpha, is computed from a 3 steps-average of the hindcast (eq. 28). Why this choice is appropriate? I recommend testing with different alpha values through sensitivity experiments.
   Sensitivity experiments were conducted in the development of this method, but we did not include the results in the original manuscript for brevity. Our analysis indicated that

1) a constant alpha value was not suitable at different lead-times due to the change in hindcast spread, and 2) a lead-time dependent constant alpha value was not suitable for different flow situations. We therefore selected a method that is forecast dependent. An average across 3 steps was selected to ensure a smoothly changing alpha mitigating instabilities. We have added this justification to Section 5.2 (lines 301-303).

iv If inflation is not localized along the network, that should be clarified and justified.
The inflation factor does not vary in space (although it does vary in time). We add the following sentence for clarity (lines 303-305): "While α is not spatially varying, it is applied to perturbation matrices with spatial structures consistent with the river network, ensuring physically plausible ensemble perturbations."

8. Spread: It's clear that the method tends to overcorrect at short lead times but yields underconfident ensembles at longer lead times (as shown in Figs. 4, 5). In general, one expects the ensemble spread to accurately represent the forecast uncertainty but the issues the authors face could be related to the ad-hoc inflation.
We agree with the reviewer that the inflation method is a reason for limited reliability of the ensemble-spread at longer lead-times. However, we also note that it is very rare that the ensemble spread accurately represents the forecast uncertainty for all lead times and locations (e.g., see Kotsuki et al, (2017)'s comparison of inflation approaches). For our proof-of-concept study, we have not carried out extensive tuning experiments, instead making some pragmatic choices. A study comparing different inflation approaches for the context of post-processing ungauged locations is left for future work. The limitations of using the inflation method are discussed on lines 636-647. We have extended the discussion on the spread in Section 7.1.2.

9. I would also note that real hydrological errors are dynamic, but the paper assumes the errors to remain constant between cycles. A flow-dependent error propagation model and perhaps an adaptive inflation approach could address these issues.
The error covariance propagation is flow dependent (based on an ensemble of precomputed hindcasts) and the inflation factor is adaptive in time. Please see our response to comments 5 and 7.iii.

10. Localization: The choice of the length scale (262 km) should be better justified. There is no sensitivity analysis to determine whether this choice is optimal or whether smaller/broader radius would improve the results.
The length scale is defined as the maximum distance between any grid-box and its closest river gauge which for our case study is 262km. This definition ensures all grid-boxes are updated by the LETKF and allows the method to be transferred to other catchments and models without the need to perform computationally expensive tuning experiments. Sensitivity experiments conducted during the development of this method found that the optimal length scale varied by location, lead-time, and tuning metric of choice, but overall, the differences were small for length scales from 65km to 786km. This clarification is provided on lines 268-273.

11. Also tangential to this, the authors need to revisit the equal error correction assumption in upstream and downstream locations. Overall, upstream locations are less dependent on distant downstream observations. Obviously, downstream conditions are often affected by accumulating upstream flows.

The propagation of the error-correction along the river network is determined by the ensemble covariances and the localisation applied. This means that while the localisation length is equal upstream and downstream, the actual analysis increments are not. This can be seen in Figure 2 where we show the analysis increments for single observation experiments. We have added a discussion on the impact up- and downstream from the observation on lines 449-463.

We agree with the reviewer that the relationships upstream and downstream are different. This is shown in Figure 3a. The cross-correlations between the hindcast and error ensembles are strongest along the river stretch near the station and decrease at longer distances. The larger correlations downstream of the station are along the flow path of the river whereas upstream the correlations show a more branch like structure because the station location is impacted by accumulation of flows from all upstream tributaries. We have added comments to the discussion of Figure 3 to make this point clearly (lines 482-488)

The localisation also dampens the influence of distant observations. Emery et al., (2020) investigate the use of localisation along the river network. They found that an observation can be beneficial to both upstream and downstream locations particularly for distances for which the flow transit time is less than the time between analyses. We have extended the discussion regarding the localisation in Section 8 (lines 628-635).

12. Figures: The figures are well-intended but too dense and overloaded with information, making them difficult to interpret and extract keys findings. I suggest splitting the complex ones (e.g., Figs. 3, 4, 6) and definitely simplify the annotations
    Thank you for this comment. To address this comment, and comment 1 of RC2, we have reduced the content of some of the figures. We have made the following changes to the figures:
    - Figure 3: We have removed panels c and g from this figure. Some of the information from c and g was duplicated in the hydrographs in Figure 5. This has allowed more room for remaining panels of Figure 3 making key details clearer.
    - Figure 4: We have simplified the annotations as request by the reviewer. We have combined the legends making the comparison between panels easier and have more clearly indicated the difference between rows 1 and 2.
    - Figure 6: We have removed panels a, d, g, and j, and the related discussion as this information is shown in the remaining panels. We have also removed the river names, which are already shown in Figure 2. We have combined the legends of panels c, f, i, and l and clearly labelled the metric shown in each column.

13. Line 6: "Error vector for each ensemble members" seems vague and unclear.
    Thank you for this comment. We have changed this sentence (lines 4-6): "Our new method employs state augmentation within the framework of the Local Ensemble Transform Kalman Filter (LETKF). Using the LETKF, an error vector representing the forecast residual is estimated for each ensemble member."

14. Line 12: The term "proxy" could mean a lot of different things. Clarify the nature of updates, whether that's real data assimilation experiment or an OSSE.
    We have changed this statement to read "A spatial cross-validation strategy is used to assess the ability of the method to spread the correction along the river network to ungauged locations" (lines 12-13).

15. Line 160: I would use "cycled" instead of "iterated"
    Thank you. This has been changed.

16. Line 160: Replace "at each timestep" with "at each observation time"
    We use "timestep" rather than "observation time" as the analysis times are dictated by the availability of the precomputed hindcast data as well as by the availability of observations. For clarity, we have changed this to "at each hindcast timestep for which observations are available" (lines 161-162)

17. Line 178: Replace "weights" with "weighs"
    Thank you. We have changed this sentence to (lines 177-179): "The Kalman gain matrix determines the impact of the innovation vector in the update step. The respective uncertainties of the prior modelled state and the observations determine their weight within the LETKF."

18. There are too many "see section xxx". This made navigation frustrating; I kept going back and forth. Consider restructuring for better flow.
    We have removed cross-references where we deem them to add complexity rather than improve the clarity of the manuscript e.g., as a signpost for the reader for the more novel components of the method (see comment 8 in RC2).

19. The word "improved" is overused in my opinion. Consider other synonyms "enhanced", "refined", ...
    We have reworded sentences where appropriate to be more specific about the effect being described.

20. Explain technical terms more clearly, for instance "spatiotemporal consistency"
    Please see comment 3.

References

Emery, C. M., David, C. H., Andreadis, K. M., Turmon, M. J., Reager, J. T., Hobbs, J. M., ... & Rodell, M. (2020). Underlying fundamentals of Kalman filtering for river network modeling. *Journal of Hydrometeorology*, *21*(3), 453-474. https://doi.org/10.1175/JHM-D-19-0084.1

Ide, K., Courtier, P., Ghil, M., & Lorenc, A. C. (1997). Unified notation for data assimilation: Operational, sequential and variational (special issue data assimilation in meteorology and oceanography: Theory and practice). *Journal of the Meteorological Society of Japan. Ser. II*, *75*(1B), 181-189. https://doi.org/10.2151/jmsj1965.75.1B_181

Kotsuki, S., Ota, Y., & Miyoshi, T. (2017). Adaptive covariance relaxation methods for ensemble data assimilation: Experiments in the real atmosphere. *Quarterly Journal of the Royal Meteorological Society*, *143*(705), 2001-2015. https://doi.org/10.1002/qj.3060

Martin, M. J. (2001). *Data assimilation in ocean circulation models with systematic errors* (Doctoral dissertation, University of Reading).

Pauwels, V. R., Hendricks Franssen, H. J., & De Lannoy, G. J. (2020). Evaluation of State and Bias Estimates for Assimilation of SMOS Retrievals Into Conceptual Rainfall-Runoff Models. *Frontiers in Water*, *2*, 4. https://doi.org/10.3389/frwa.2020.00004

Rasmussen, J., Madsen, H., Jensen, K. H., & Refsgaard, J. C. (2016). Data assimilation in integrated hydrological modelling in the presence of observation bias. *Hydrology and earth system sciences*, *20*(5), 2103-2118. https://doi.org/10.5194/hess-20-2103-2016

Reichle, R. H., D. B. McLaughlin, and D. Entekhabi, 2002: Hydrologic Data Assimilation with the Ensemble Kalman Filter. *Mon. Wea. Rev.*, 130, 103–114, https://doi.org/10.1175/1520-0493(2002)130<0103:HDAWTE>2.0.CO;2.

Ridler, M. E., Zhang, D., Madsen, H., Kidmose, J., Refsgaard, J. C., & Jensen, K. H. (2018). Bias-aware data assimilation in integrated hydrological modelling. *Hydrology Research*, *49*(4), 989-1004. https://doi.org/10.2166/nh.2017.117

Valler, V., Franke, J., & Brönnimann, S. (2019). Impact of different estimations of the background-error covariance matrix on climate reconstructions based on data assimilation. *Climate of the Past*, *15*(4), 1427-1441. https://doi.org/10.5194/cp-15-1427-2019

Zhang, F., Snyder, C., & Sun, J. (2004). Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter. *Monthly Weather Review*, *132*(5), 1238-1253. https://doi.org/10.1175/1520-0493(2004)132<1238:IOIEAO>2.0.CO;2

Response to RC2 for hess-2024-3989: Matthews, G., et al. Error-correction across gauged and ungauged locations: A data assimilation-inspired approach to post-processing river discharge forecasts

We thank the reviewer for their insightful comments and helpful suggestions which we believe will greatly strengthen the evaluation and discussion of the new method. The reviewer's comments have been summarised and numbered for clarity. The authors' responses are in blue. Section and line numbers refer to the revised manuscript.

1. The paper is too long, in particular the description of the methods.
   We have shortened the descriptions of methods by restructuring and condensing the descriptions in Sections 3.1, 5.2, and 5.3. Addressing some of the comments from both reviewers such as comment 1, 3, and 12 from RC1 have also reduced sections of the manuscript.

2. I am a bit concerned about how applicable these methods are outside the case study attempted (see specific comments below). For example, the use of a very large catchment is likely to allow the authors to make simplifying assumptions such as that residuals will be normally distributed, or that errors can be characterised using a 10-day window. I would be interested in some discussion of how generalisable these methods are.
   Thank you for this comment. We have extended the discussion on the generalisability of the method to include the estimation of the initial error ensemble (where the 10-day window is used; lines 660-670) and the assumption of Gaussianity (648-659). See comments 5, 6, and 18.

3. L58 "ensemble Kalman Filters are common data assimilation methods for hydrological applications" this is true for hydrological research, but (to me at least) it remains a curiosity as to why data assimilation within hydrological models - including with ensemble Kalman Filters - remains to my knowledge quite rare in operational streamflow forecasting systems.
   This is a good point. We have specified in "hydrological research applications" and have added the following to the introduction (lines 62-68) "Whilst many studies have shown the benefits of data assimilation for hydrological forecasting (Tanguy et al., 2025; Valdez et al., 2022; Piazzi et al., 2021), the process is rare in operational systems (Pechlivanidis et al., 2025), particularly in large-scale systems (Wu et al., 2020). This limited uptake is partly due to data latency issues (WMO, 2024), time constraints, and the potential impact on the interpretation of the forecasts (e.g., thresholds based on model climatology may no longer be consistent; Emerton et al., 2016). Additionally, the benefit of data assimilation at longer lead-times is uncertain (e.g., Valdez et al., 2022). In this paper, we leverage key advantages of data assimilation—such as the ability to propagate observational information to ungauged locations—within a post-processing framework that is more readily integrated into operational systems."

4. L90 "Hydrological ensemble forecasts consist of N potential realizations referred to as ensemble members" I think it would be good to state explicitly which variable(s) you are discussing here, as it wasn't clear to me - I'm assuming streamflow (or runoff, as it's on a grid?)?
   The variable of interest is streamflow or river discharge. We have changed this to "The hydrological ensemble forecasts consist of N potential realizations of future river discharge, referred to as ensemble members" (lines 96-97).

5. L107 I would have thought with a strongly skewed (and potentially zero bounded) variable like streamflow, an additive error only generally holds after a normalising transformation has been applied (and, if applicable, zero values have been dealt with).
   We have added normalising transformations to the discussion (see comments 2 and 18). The reviewer is correct that the assumption of Gaussianity limits the applicability of an additive error by occasionally resulting in negative discharges. We deal with negative discharge values as described in Section 4.1. The impact and potential solutions are discussed in Sections 7.2 and 8.

6. L112 Similar to the above criticism at L107, Equation 6 appears to assume that errors are normal and homoscedastic. If my understanding of what is being assimilated is correct, this is highly unlikely to hold for streamflow, for which residuals are almost always non-normal and heteroscedastic. See e.g. Smith et al. 2015, among many others.
   We agree that streamflow residuals are often non-Gaussian and heteroscedastic. Our framework does employ updates based on Gaussian assumptions as we use the LETKF, but the resulting distribution is not necessarily Gaussian (Reichle et al., 2002). However, we do not assume homoscedasticity: the ensemble spread evolves dynamically and reflects state-dependent and lead-time-dependent error variability. We have changed the description of the forecast correction to make this more clear (lines 110-120).

7. L145 "we adopt the common assumption that the error is constant" I would not have said this is common. I would say it's much more common to use autoregressive models (often AR1) to describe the autocorrelation between residuals in streamflow. I understand why this is a pragmatic simplification, but errors often do change with lead time as the value of forecast information decays.
   Thank you for the comment. We have added the following to Section 3.1 (lines 146-148): "As the true evolution of the error vectors at all grid-boxes is unknown, we assume a simple persistence model, such that $\mathbf{b}_{\mathrm{k}}^{(\mathrm{i})} = \mathbf{b}_{\mathrm{k}-1}^{(\mathrm{i})}$. This is a common assumption used in state augmentation (Pauwels et al., 2020; Rasmussen et al., 2016; Ridler et al., 2018; Martin, 2001)." We have also added a discussion of this assumption to Section 8 (lines 671-678).

8. L149 "define the propagation" I'm not sure what 'propagation' means here, given the error is assumed constant in time. Can the authors clarify? Nevermind - the authors do this in Section 3.2! The authors may want to flag that the explanation for this is coming.
   We have signposted that the propagation equation defined on line 147 is for use in the LETKF described in Section 3.2

9. L180 "(see Eqs. (8) and (9) in Bell et al., 2004)" I feel that if the authors need to specify equations from another study to describe these methods, the equations should be present in the paper (in an appendix is fine) - especially Eq (9) of Bell et al. which the authors later describe as 'key' to the method. (Unless they are included later?)
   We have added the decomposition of the Kalman gain matrix as an appendix (Appendix A)

10. Figure 1 - this is a really nice, clarifying figure.
    Thank you!

11. L223 "We enforce non-negativity by further adjusting the error ensemble members after the LETKF update step (Fig 1)." This indicates that zero values are present in output state, indicating

that errors are not continuously distributed. I realise not everyone handles zeros, but it would be good to acknowledge the limitation of this assumption (as noted above).
We have extended the discussion regarding the assumption of Gaussianity for river-discharge. See comment 5.

12. L265 "Eq. 4.10 in Gaspari and Cohn (1999)" - I think the authors should include this equation, as well as discussing (briefly) why they thought the form of this equation appropriate for this task. The regionalisation of errors is in my view the major contribution of the paper.
The Gaspari and Cohn correlation function is a commonly used localisation function in data assimilation as it smoothly decreases to a definable radius. We have added the Gaspari and Cohn function as an appendix (Appendix B).

13. L272 "We propose instead for the localisation length scale to be defined as the maximum distance between any grid point and its closest observation." This seems like a sensible choice.
Thank you.

14. L325 "(here 10 days)" This is a long period over which to assess an error - some use periods of this length for bias correction (e.g. Bennett et al. 2021). I'm assuming this really only works for larger catchments where rivers have slower varying errors; I would have thought for small headwater gauges shorter periods would be more appropriate. It also explains why errors are assumed not to vary with lead time, above. This is all fine, but the authors may wish to mention this in their discussion.
This is a good point. The 10-day period is used to generate the initial error ensemble mean. This initial estimate is not used to correct the river discharge ensembles directly, but rather to provide a starting point for the LETKF. The LETKF then updates the error ensemble at each timestep. We have updated Figure 1 to better clarify this process.

We selected a 10-day period to capture the consistent biases of the hydrological model but also to allow for seasonal/dynamic variation in this bias. We agree that shorter periods may be more appropriate for smaller, fast-responding catchments and have added this to the discussion (lines 660-670).

15.
   a. L408 "we assume that the observation errors from different gauge stations are uncorrelated" I'm not suggesting a change here, and I think this is a reasonable suggestion without additional information. But I suspect the long-range nature of the errors (a 10-day period) may undermine the assumption somewhat.
   The observation errors arise due to instrument uncertainty, observation processing, observation operator error and scale mismatch between the observations and the model resolution. Observation errors are also assumed to be uncorrelated with the prior errors which is a standard assumption in data assimilation. We have clarified this in lines 379-381 and lines 383-384.

   b. I'm also curious what happens when errors are propagated in space: what happens when you get a point equidistant (or close to equidistant) from two gauges, and the errors from the two gauges interact in some way (e.g. cancel each other, or sum).
   The Kalman gain matrix governs the spreading of observation information in space. A weighted mean is calculated. The weight of an observation is determined by the cross-

covariances between the hindcast and error ensembles, the distance from the observation (via the localisation), and the uncertainty in the observation itself. The left and central panels in Figure 2 show single observation experiments, indicating how information from one observation is spread spatially. The panels on the right of Figure 2 show how observation information is propagated when all available observations are assimilated. In Section 8, we have added a discussion on the potential benefits of future work using a block cross-validation to explore the impact of observation density and location (lines 679-683).

16. L438 "forecast mean is decreased by the proposed method we use the Normalised Mean Absolute Error" It's preferable to apply measures of absolute error to the ensemble median. See, e.g., Taggart (2022).
    Thank you for this very helpful comment. We have changed Figure 6, and corresponding discussion (lines 584-588) to use the RMSE.

17. L526 "However, this assumption is necessary to propagate the hindcast to the next time step without the use of a hydrological model (Section 3.1)." Perhaps, but one application of ensemble predictions is to sum ensemble members through time (e.g. to assess cumulative inflows to reservoirs). From this figure, it seems this would result in highly unreliable accumulations. This may not be an application of EFAS (I don't know), but if the method is to have more general applicability this is a serious weakness.
    The hydrographs shown in Figures 4b and 4e are not the final hydrographs resulting from this method but instead are intermediate steps used to investigate the impact of the methodological assumptions made. We have restructured Section 7 to make a clearer division between the results that show how the method works and the results that show the skill of the resulting ensemble (see comment 3 from RC1).

18. L660 "Future work could look into applying anamorphosis to make the ensemble distribution more Gaussian-like" I'm not familiar with the concept of anamorphosis, but a conventional way of doing this is to use normalising transformations, of which many are available for hydrological variables.
    Anamorphosis is very similar to normalising transforms used in hydrology. We have added the use of normalising transformations to this paragraph.

19-23. Typos and grammatical errors.
    These typos have all been corrected. Many thanks to the reviewer for catching them!

References

Martin, M. J. (2001). *Data assimilation in ocean circulation models with systematic errors* (Doctoral dissertation, University of Reading).

Pauwels, V. R., Hendricks Franssen, H. J., & De Lannoy, G. J. (2020). Evaluation of State and Bias Estimates for Assimilation of SMOS Retrievals Into Conceptual Rainfall-Runoff Models. *Frontiers in Water*, *2*, 4. https://doi.org/10.3389/frwa.2020.00004

Rasmussen, J., Madsen, H., Jensen, K. H., & Refsgaard, J. C. (2016). Data assimilation in integrated hydrological modelling in the presence of observation bias. *Hydrology and earth system sciences*, *20*(5), 2103-2118. https://doi.org/10.5194/hess-20-2103-2016

Reichle, R. H., D. B. McLaughlin, and D. Entekhabi, 2002: Hydrologic Data Assimilation with the Ensemble Kalman Filter. *Mon. Wea. Rev.*, **130**, 103–114, https://doi.org/10.1175/1520-0493(2002)130<0103:HDAWTE>2.0.CO;2.

Ridler, M. E., Zhang, D., Madsen, H., Kidmose, J., Refsgaard, J. C., & Jensen, K. H. (2018). Bias-aware data assimilation in integrated hydrological modelling. *Hydrology Research*, *49*(4), 989-1004. https://doi.org/10.2166/nh.2017.117