Responses to Referee #2

We thank referee #2 for showing interest in our work and the constructive comments. The issues raised are interesting and we (the authors) will address them in the revised version of this manuscript. Here we provide our response to the comments (in italics). Please note that all citations herein refer to the papers in the discussion paper.

Major comments:

1. *Lack of scientific innovation as a methodology paper. I couldn't consider the proposed complementary modelling framework as a new approach because inflow forecasting has been done by applying error models to base hydrologic model simulations for more than 20 years. There is nothing new on error model structure, hydrologic model calibration or the way to combine two models. I am aware there is a paragraph on Page 12067 attempting to describe two innovations of this work: forecasting with a lead-time up to 24 hr and enabled probabilistic forecasting. The length of lead-time depends on the need of the application, and it is not part of innovation. The probabilistic forecasting directly derived from error models have been already considered intensively in most previous work.*

   We agree with referee #2 that the point raised was not clear in the discussion paper. We elaborated methodological contribution of this work in response to referee #1 (first comment). This will be clearly spelled out in the revised paper.

2. *Lack of assumption validation as an application paper. To warrant a successful application, the model assumption should be examined under scrutiny. For example, the ACF and PACF plots based on the forecast error in the transformed space (instead of in the original space) should be provided. […] The normality of the residuals (after appropriate transformation) in the AR(1) model should be also validated.*

   We thank referee #2 for raising these issues. We employed techniques of visual inspection (of the residual, ACF and PACF plots) and statistical test (Kolmogorov-Smirov test) for validating the model assumptions. Omission of the residual, ACF and PACF plots corresponding to the residuals in the transformed space was in an effort to shorten the discussion paper to the present length. Yet, in line with the above comments, we believe the discussions on the Kolmogorov-Smirov test provided on P12076 (L24-L26) and P12078 (L6-L8) and the remark therein on the normality of the residuals are noteworthy. As can be seen in Fig. B1(a), the residuals show better variability over the entire range of predicted inflow in the transformed space. Similarly, comparison of the ACF and PACF plots of Fig. A1 and Fig. 4 (discussion paper) reveals the extent to which the serial correlation in the residual series reduced.

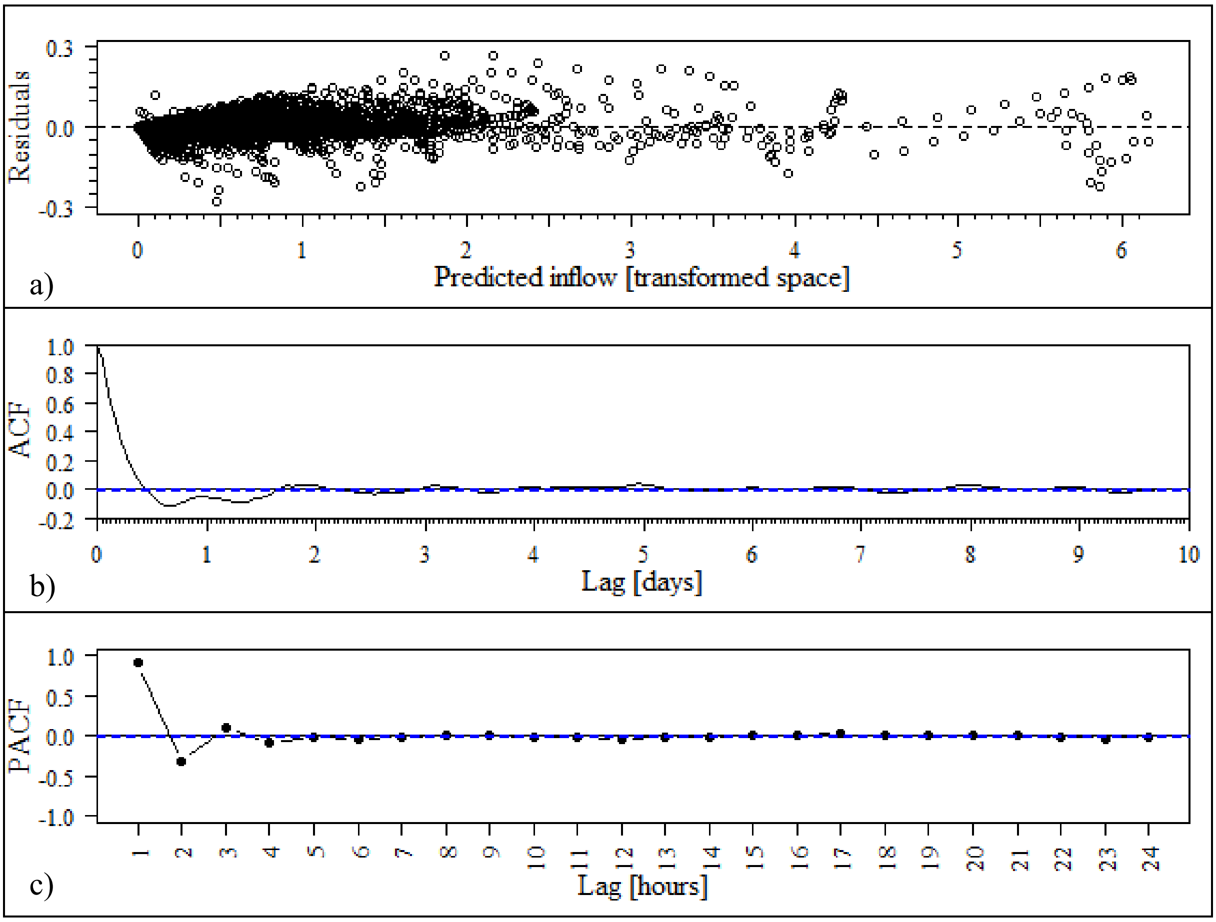

Figure B1. Plots of (a) the residuals as a function of predicted inflow (in the transformed space), (b) autocorrelation function of the residuals, and (c) partial autocorrelation functions of the residuals.

I doubt that an AR(1) model is sufficient to account for the strong persistence in the hourly time series.

We agree with referee #2 that the ACF and PACF plots of Fig. 4 (discussion paper) suggest AR model of order higher than one. Though not described in the discussion paper, the selection of AR(1) model was based on thorough assessment of AR(1), AR(2) and AR(3) models. The selection of the error model is intrinsic element of the error-model calibration process outlined in the “Parameter estimation” subsection of section 2.1.2. In accordance with step 3 (P12072 L1), first and foremost we calibrated several AR models of up to order $p = 3$ by minimizing the sum of the squares of the offsets between the inflows (observed and predicted) in the transformed space. Subsequently, we assessed whether the residuals of the complementary modelling framework appear homoscedastic and exhibited correlation. This assessment was carried out using the Kolmogorov-Smirov (KS) statistic followed by visual inspection of the residual plots. The KS statistic served as a relative measure of the difference between the distributions of the residuals from a number of AR model setups (see Table B1).

These issues will be made clearer in the revised paper.

Table B1: Example of comparison made to AR models of different orders

	AR(1)	AR(2)	AR(3)
Box-Cox	$\lambda = 0.9$	$\lambda = 0.2417$	$\lambda = 0.0013$
	$\beta = 41.4$	$\beta = 40.89$	$\beta = 70.47$
AR coefficients	$a_1 = 0.97$	$a_1 = 0.586$	$a_1 = 2.15$
		$a_2 = 0.406$	$a_2 = -1.26$
			$a_3 = 0.087$
KS statistic	0.1000	0.2578	0.2092

3. *(a) I can't see whether the AR model is applied to transformed or original data. From Equations (2) and (3), it seems to apply to the inflow without transformation. If so, I don't know why the Box-Cox transform is mentioned in the section related to "Parameter estimation".*

   The AR model is applied to the transformed data. This will be described better in the revised manuscript.

3. *(b) Some notations are not used consistently and cause confusion. For example, ε_t is differently defined in Equation (2) and in the last line of Page 12071.*

   As rightly pointed out by reviewer #2, ε_t denotes the error between the observed and predicted inflows before and after transformation (Eq. 2 and P12071 L22, respectively). We will address this in the revised manuscript.

   I am not sure why $\hat{\varepsilon}_t$ instead of e_t is used in Equation (5).

   Equation 5 provides the simulated error designated as $\hat{\varepsilon}_t$. We agree with referee #2 that this simulation can be expressed in terms of the bias free error (e_t). This will be corrected in the revised manuscript.

4. *The estimation of the transformation parameters described on Pages 12071-12072 is incorrect. My understanding is that the authors attempt to minimise the sum of forecast error in the transformed space (not really sure because of unclear notations). I suggest that the transformation parameters are estimated by a likelihood approach.*

   We agree with reviewer #2 that estimation of the transformation parameters can be carried out by a likelihood approach. However, we do not concur the opinion that the procedure outlined in the discussion paper is incorrect. As demonstrated by Beven et al. (2008), the procedure we adopted provides another way for selecting and estimating parameters of an

AR model while dealing with the heteroscedasticity the data exhibits at the same time. We accept that clearing the confusion related to the mathematical notations will benefit the manuscript very well, and will be done in the revised version.