Figure S1– The flowchart of the WRF-Hydro calibration experiments.

S2. WRF-Hydro

WRF-Hydro is one of such advanced hydrological models developed at the National Center for Atmospheric Research (NCAR), USA. It has five components (atmospheric model, land-atmosphere parameterization, groundwater, hillslope and/or channel routing, and water management) that are connected through a coupler (Figure S1).
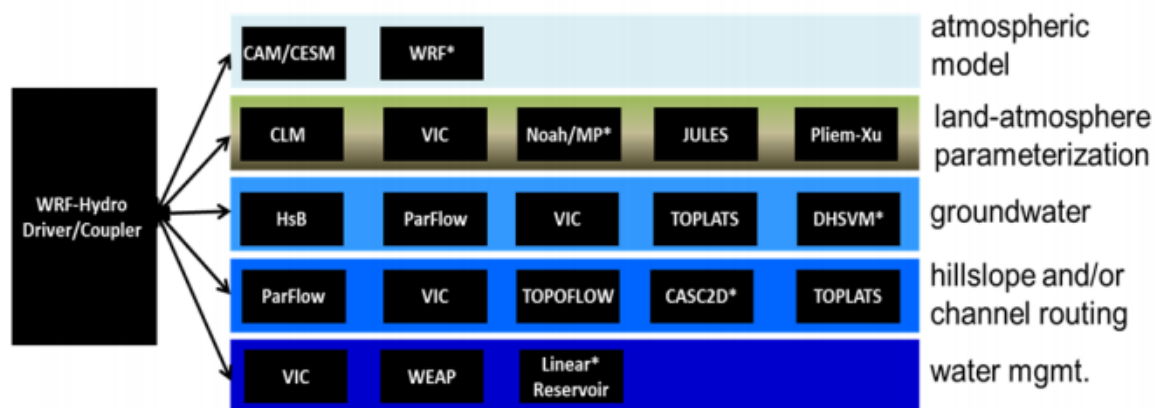

Figure S2 – Schematic diagrams of the WRF-Hydro coupler and components. From the technical description of WRF-Hydro (adopted from Gochis et al., 2018).

The atmospheric component can be coupled through an online/offline mode. In the case of online coupling, it feeds back to the atmospheric component. In this model, various modules integrate detailed hydrological processes such as vertical land surface parameterization, surface overland flow, saturated subsurface flow, channel routing, reservoir routing, and conceptual base flow processes (Gochis et al., 2018). The model allows the routing of water that is excess to the soil infiltration capacity and the saturated subsurface water. The infiltration of excess water, together with possible exfiltration from the saturated soil, is subsequently allowed to pond on the land surface, move laterally, and reinfiltrate if conditions are feasible (Senatore et al., 2015).

The hydrological processes are more sensitive to the grid size than to the atmospheric processes. Therefore WRF-Hydro allows the aggregation-disaggregation from the atmospheric grid to the routing grid. The disaggregation is performed after executing land surface processes and prior to routing exercise. Thereby, more accurate channel flow can be estimated with the more accurate representation of channel grids.

Subsurface flow is computed prior to the overland flow routing to allow the overland routing of the water from exfiltration. It uses the quasi 3-D flow to calculate the saturated soil moisture's lateral flow, including the effect of topography, saturated conductivity, and saturated soil depth. WRF-Hydro does support a fully unsteady, explicit, finite difference, diffusive wave approach for the overland flow routing (Julien et al., 1995; Senatore et al., 2015; Ogden, 1997) that also accounts for the backwater effects. However, the simple steepest descent approach (D8 approach) is also available for overland flow routing. The numerical stability (meeting the courant criteria), to solve the diffusive wave equations is achieved through the conservative time step that can be changed, if required, in the namelist file. The model performs channel flow routing by integrating the diffusive wave equations through a first-order Newton-Raphson solver. The initial time step of channel flow routing remains the same as overland flow routing. The channel routing is performed grids-by-grids on routing grids along with the channel network. The overland flow discharge into the channel is computed using the mass balance analysis.

S3. Parameter ESTimation (PEST) Tool

PEST is a model-independent parameter estimation tool. It requires three input files: template files, instruction files, and control files. The template files are cloned to the parameter files, just that it holds a placeholder in place of the involved parameter's value. The instruction files are created to collect the model output. For every observation, one instruction file must be created. Control file acts as a controller to define the configuration of the inversion process, including the mode of operation. PEST can be run into four modes, estimation, predictive analysis, regularisation, and pareto. PEST was utilized in the estimation and regularisation modes in this study.

PEST is iterative in behavior that means it runs in repetitive iterations. These iterations are required because of the non-linear behavior of the models. Due to this non-linearity, the parameter sensitivity changes by changing its value. Only one iteration would suffice for the inversion process in linear models (Doherty 2015). Each iteration computes the initial parameter sensitivity through the Jacobian matrix (Jacobian transformation). The initial sensitivity was then used to estimate the improved set of parameters through Marquardt lambda. The different lambda values were generated using the trial and error method.

The calculation of partial derivatives (of observation with respect to the adjustable parameters) is a key to the inversion process. These derivatives are the elements of Jacobian matrix and can be calculated through forward or central differences with two, three, or five finite points. PEST is capable of starting the inversion using forward differences and switching to a central difference using three or five finite points if the inversion process seems to be slower, identified through the relative reduction in the objective function below a certain

threshold. The runtime switch prevents the iterations' slowdown and fastens the inversion process.

The performance of PEST with the traditional methods could be compromised if the inversion problem is ill-posed. The Gauss-Marquardt-Levenberg method calculates the updated parameter set through the inversion of the normal matrix, added by the Marquardt lambda to its diagonals. If this matrix is singular, thereby non-invertible, then the problem is ill-posed and unable to compute the global minima of the objective function. However, Singular Value Decomposition (SVD) and Least Square (LSQR) guarantee numerical stability to solve the inverse problem. Hence, the global minima of the objective function are achievable. The addition of Tikhonov Regularisation ensures obtaining the parameter set with minimum error variance. In this study, the PEST was used in both with and without regularization mode. In case of regularization, both SVD and LSQR were used as a solver. The regularization was applied keeping in mind its superiority over the default classical mode.

S4. PEST configuration

PEST requires a few thresholds to perform model simulations iteratively to optimize its parameters to suggest as terminating conditions either for iteration or experiment. Details of PEST and its functioning can be found in Doherty (2016). Following are some of the important thresholds.

PEST uses objective function to minimize the total squared error between observed and simulated discharge:

$$\emptyset_m = \sum_{k=1}^{n} w_k \left(q_k^{obs} - q_k^{sim} \right)^2$$

Equation 1

ϕ_m represents the measurement objective function, w_k represents the weight, q_k^{obs} is the observed discharge and q_k^{sim} is the simulated discharge for kth timestamp. ϕ_m works as an objective function when PEST uses traditional methods to solve the inverse problem, however if regularisation is applied, the objective function is modified to:

$$\emptyset = \emptyset_m + \mu^2 \emptyset_r$$

Equation 2

where ϕ represents the total objective function, ϕ_m represents the measurement objective function and ϕ_r represents the regularised objective function.

PEST runs through iterations because of the non-linear behaviour of the model and variable composite sensitivity of the parameters. Composite sensitivity is computed using Eq. 3, where J stands for Jacobian matrix and Q denotes the weight matrix.

$$cs = \frac{[Q^{1/2} J^t J Q^{1/2}]_{jj}^{1/2}}{m}$$

<div align="right">Equation 3</div>

The multiple iterations are required because parameter sensitivity changes upon changing the parameter initial sensitivity using Jacobian derivatives, which depends on their initial values. PEST converges and optimizes the parameters over these iterations; however a certain threshold is necessary as termination condition. A few important thresholds that we used are as follows

PHIRATSUF (phi ratio sufficient) was defined to determine whether the phi reduction goal for a particular iteration is achieved. So, if threshold criteria met (Eq. 4), PEST moves on to the next iteration. This value was chosen to be 0.001.

$$\frac{\emptyset_i^j}{\emptyset_{i-1}} \leq \text{PHIRATSUF}$$

<div align="right">Equation 4</div>

PHIREDLAM (phi reduced lambda) determines when to move to the next iteration. However, the decision is made based on the different lambda values. If the ratio of objective function of successive lambdas (Eq. 5) fails to meet the threshold, PEST will not try with more lambda values and move on to the next iteration. This value was chosen almost zero to allow the PEST to try with the maximum number of lambdas.

$$\frac{\left(\emptyset_i^{j-1} - \emptyset_i^j\right)}{\emptyset_i^{j-1}} \leq PHIREDLAM$$

<div align="right">Equation 5</div>

PHIREDSWH (phi reduced switch) determines when to switch to the higher orders for derivative calculation (i.e., two point forward difference to three/five point central difference). If the ratio of objective function through two successive lambdas (Eq. 6) values fails to meet the threshold, the higher order derivative calculation activates. This threshold makes sure that the inversion process is not slowing significantly. This value was chosen to be 0.1.

$$\frac{\left(\emptyset_{i-1} - \emptyset_i\right)}{\emptyset_{i-1}} \leq \text{PHIREDSWH}$$

Equation 6

PHIREDSTP (phi reduced stop) puts the threshold to the objective function of an iteration. If the objective function does not meet this threshold, PEST will cease execution after the occurrence of such iteration NPHISTP times. PHIREDSTP was chosen to be 0.01 and NPHISTP to be 3.

$$\frac{(\emptyset_i - \emptyset_{min})}{\emptyset_i} \leq \text{PHIREDSTP}$$

Equation 7

We used specialized version of PEST i.e. BeoPEST (it will be referred as PEST hereafter) where a parallelization architecture is in form of master and slave nodes (Figure 3). A manager controls all the parallel ongoing simulations while doing nothing by itself but recording performance of all iterations and generate new parameter sets based on simulation configurations. The simulations are performed at slave nodes that do not interact with each other directly.
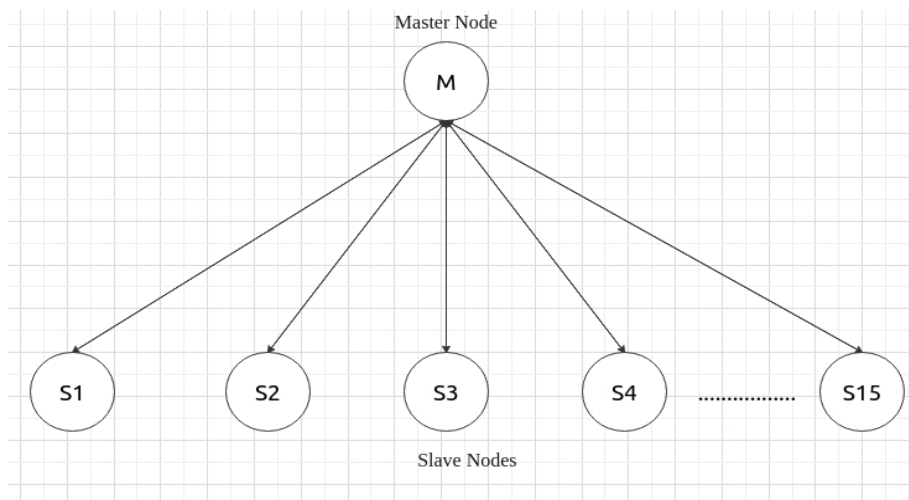

Figure 3– The schematic diagram represents the parallel processing during the model (WRF-Hydro) inversion under the PEST framework.

Experiment	Calibration Period	Optimization Iteration	Model Calls	Comment
MP8KF_em	ANN	9	743	(~ 743 model years)
MP8KF_reg_lsqr	ANN	9	737	(~ 737 model years)
MP8KF_reg_svd	ANN	22	1603	(~ 1603 model years)
WSM6BMJ_em	ANN	9	735	(~ 735 model years)
WSM6BMJ_reg_lsqr	ANN	15	1103	(~ 1103 model years)
WSM6BMJ_reg_svd	ANN	18	1348	(~ 1348 model years)
WSM6BMJ_reg_lsqr_jjas	JJAS	24	1687	(~ 562 model years)
WSM6BMJ_reg_svd_jjas	JJAS	14	987	(~ 329 model years)

Table 1 – Design experiments for the WRF-Hydro calibration under the PEST framework. The experiments include the traditional as well as more sophisticated (SVD or LSQR regularisation) numerical solutions to the inverse problem. The comment column shows the number of model years required to achieve the optimum solution.

Using above thresholds, two simulations were designed under parameter estimation mode, each with MP8KF and WSM6BMJ meteorological forcing. Under the parameter estimation mode, PEST provides a traditional numerical solution; hence it does not guarantee a solution to the inverse problem. To solve this problem, four more experiments were designed with a regularisation mode that uses SVD or LSQR (MP8KF_reg_lsqr, MP8KF_reg_svd, WSM6BMJ_reg_lsqr, and WSM6BMJ_reg_svd experiments in Table 3). The experiments designed so far were using meteorological forcing for an entire year. However, WSM6BMJ performs well just for JJAS that accounts for highest annual variability, hence two more experiments were designed focused just on JJAS using WSM6BMJ forcing to explore whether a season-specific calibration can improve performance (WSM6BMJ_reg_lsqr_jjas and WSM6BMJ_reg_svd_jjas experiments in Table 4.3). Hereafter, MP8KF* experiments will

refer to a set of hydrological experiments that uses MPKF meteorological forcing and similarly for WSM6BMJ* experiments.

S5. Regularisation

Regularisation refers to the mathematical measures taken to find a stable solution to otherwise ill-posed problems (Doherty and Skahill, 2006). It also helps finding the solution to inverse problems that are well-posed but ill-conditioned (Linden et al., 2005), besides being applied to the well-posed and well-conditioned problems. The over-parameterization of a model can lead to the ill-posedness of the numerical solution. The constraints or additional information provided in the form of mathematical relationships to stabilize the solution is known as regularization constraint optimization.

The regularisation can offer a robust optimization strategy while estimating the unique value of model parameters. There are two general approaches implementing the regularization, penalty function and dimensionality reduction. Tikhonov regularization is an example of a penalty function, and truncated singular value decomposition is an example of a dimensionality reduction approach.

PEST offers the use of Tikhonov regularization by aiding an additional objective function besides the one for the measurement, known as the regularization objective function. Therefore, the overall objective function will be $\Phi = \Phi_m + \mu^2 \Phi_r$ where Φ_m is a measurement objective function and Φ_r is a regularization objective function. μ is a regularization weight factor. PEST calculates μ following an iterative search process during every optimization iteration of the Gauss-Marquardt-Levenberg method (Doherty, 2003).

S6. Bias correction challenges

The bias correction of the simulated precipitation was attempted using the station observed precipitation. It is to be noted that the stations are spread within a small area, and cover only a very small fraction of the model domain. Supplementary Figure A2.1 shows the observed, WRF simulated, and bias corrected precipitation for ANN, JJAS, and DJF. No significant changes were found in the simulated precipitation after the bias correction, since the stations cover a very small fraction of the model domain. Also, to be noted is that large biases occur over higher altitudes where there are no stations. Bias corrected JJAS precipitation was found to underestimate the observation throughout the region, similar to the simulated (non bias-corrected) precipitation pattern. For JJAS precipitation, R-square was found to be slightly decreased, however, R to be slightly increased. The RMSE remains almost similar (Table A2.1).

Therefore, using the bias corrected precipitation for WRF-Hydro calibration may not add much value. Instead, the weighted ensemble was used to represent the JJAS and DJF precipitation with different parameterization schemes. The ensemble methodology appeared to overcome the underestimation/overestimation in the precipitation for different seasons. Thereby, a good accuracy was achieved when using a weighted ensemble. Moreover, the observations are available at daily frequency. However, the hourly precipitation was used to account for the more reliable snowmelt and discharge consequently. Notably, the important conclusion from this chapter is to have different parameterizations into consideration (for different seasons) to produce reliable precipitation/discharge.

	ANN		JJAS		DJF	
	WRF	**WRF-BC**	**WRF**	**WRF-BC**	**WRF**	**WRF-BC**
RMSE	4.25	4.15	9.5	9.3	2.8	2.82
R-square	0.51	0.46	0.58	0.49	~0	~0
R	0.56	0.66	0.41	0.45	0.32	0.47

Table S6 – The accuracy metrics of the WRF simulated precipitation with APHRODITE observation before and after the bias correction.
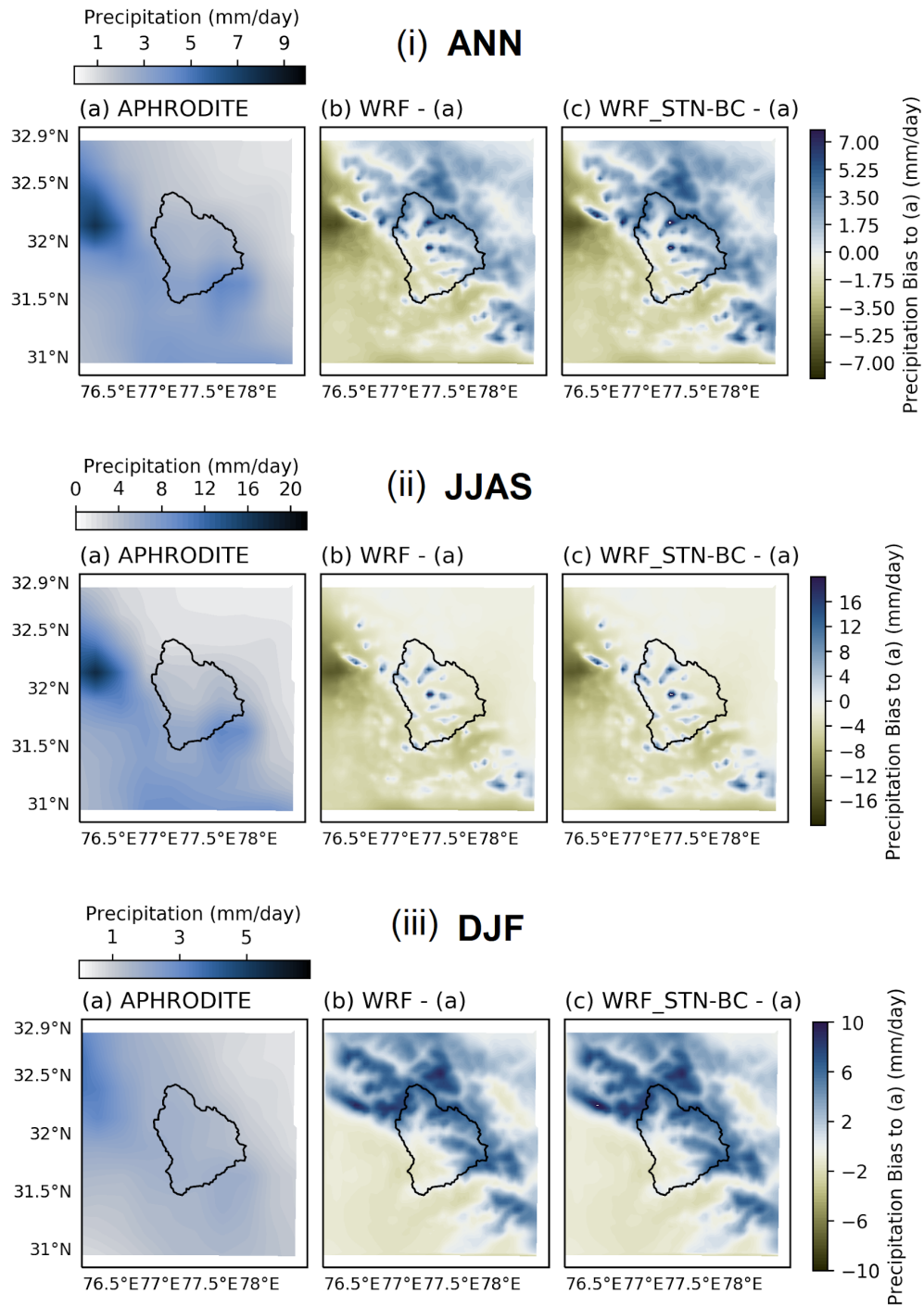
Figure S6- (a) shows the observed average annual precipitation for (i) ANN, (ii) JJAS, and (iii) DJF. (b) shows the average annual precipitation bias to (a) for MP8_KF experiment with WRF for (i) ANN, (ii) JJAS, and (iii) DJF. (c) shows the bias corrected precipitation of the WRF MP8_KF experiment using observed precipitation at seven locations for (i) ANN, (ii) JJAS, and (iii) DJF.