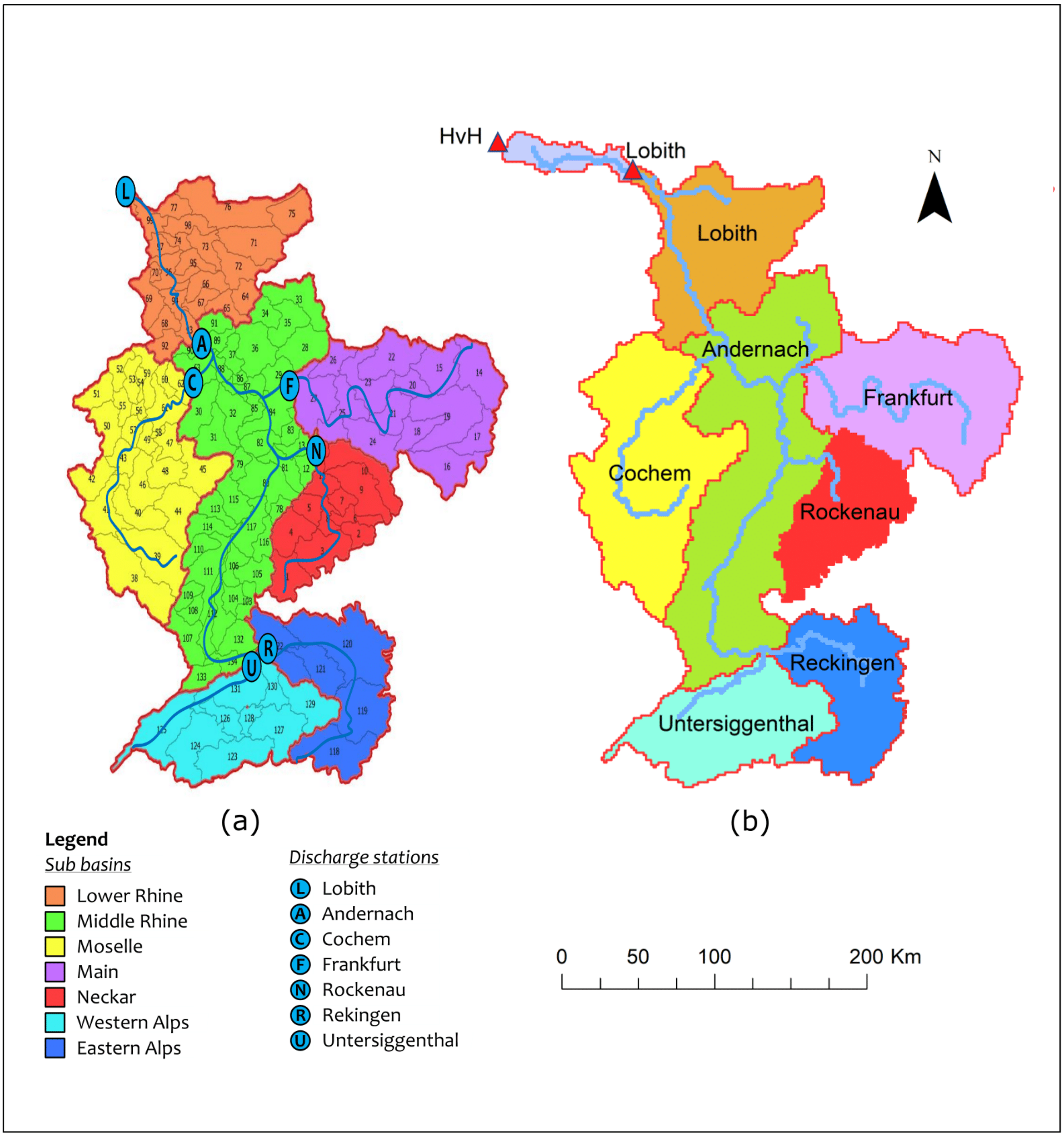

Figure S1: Sub-basin discretization for HBV model (left, Huiskes 2016) and SPHY(right). HBV model used 134 sub basins to calibrate the model whereas, SPHY model was calibrated at 7 sub basins. The major sub-basin in the model are delineated at the same location.

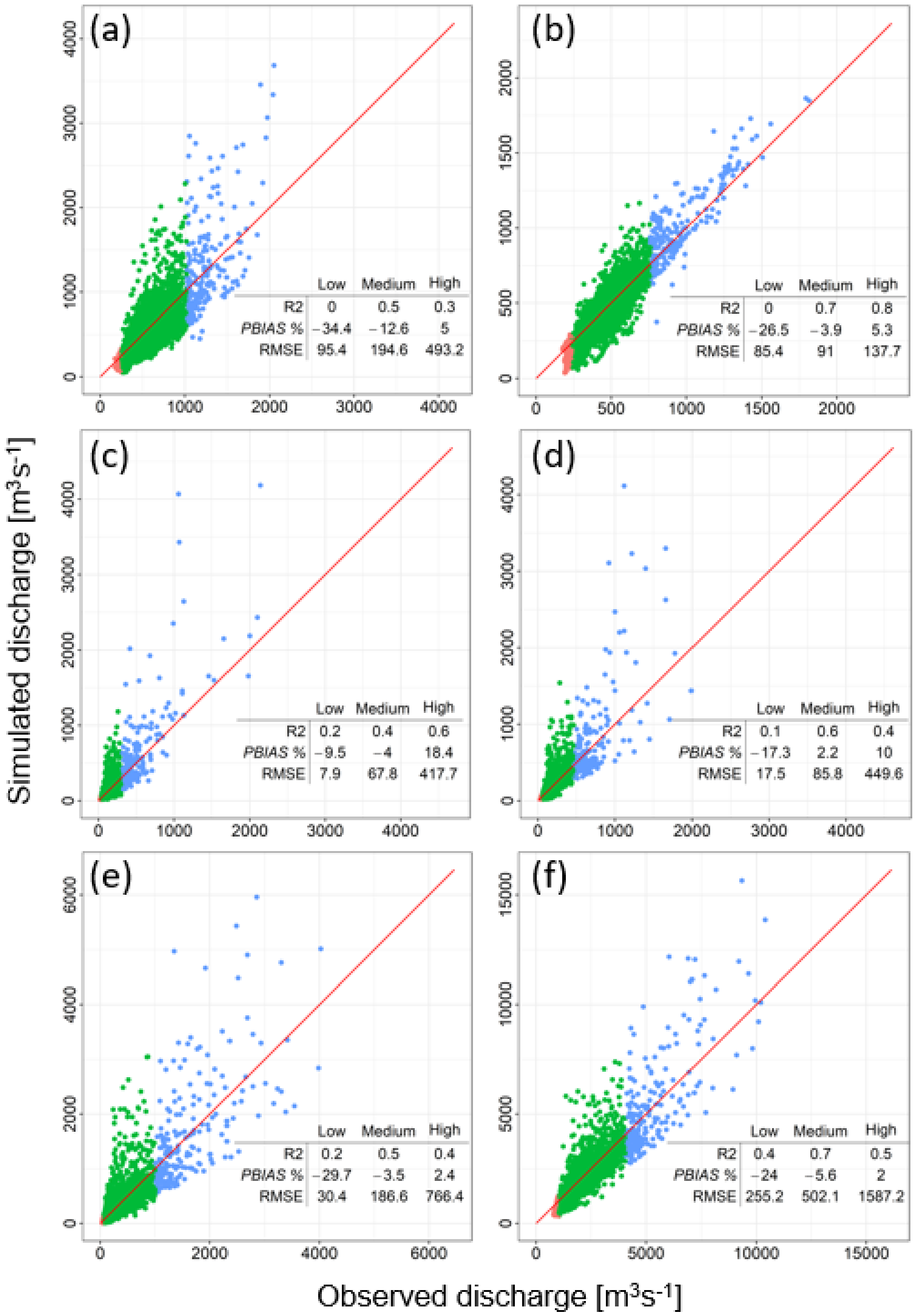

Figure S2: Observed versus modeled daily streamflow at (a) Untersiggenthal (b) Reckingen (c) Rockenau, (d) Frankfurt (e) Cochem and (f) Andernach for SPHY model. Colours indicate three ranges based on observed quantiles: "low" (<5%, red), "med" (5–95%, green) and "high" (>95%, blue). The solid red line represents the 1:1 slope. The tables on the figure represent the performance statistics coefficient of determination (R^2), percent bias (PBIAS, %) and root mean square error (RMSE, m^3s^{-1}) for three ranges of quantiles.

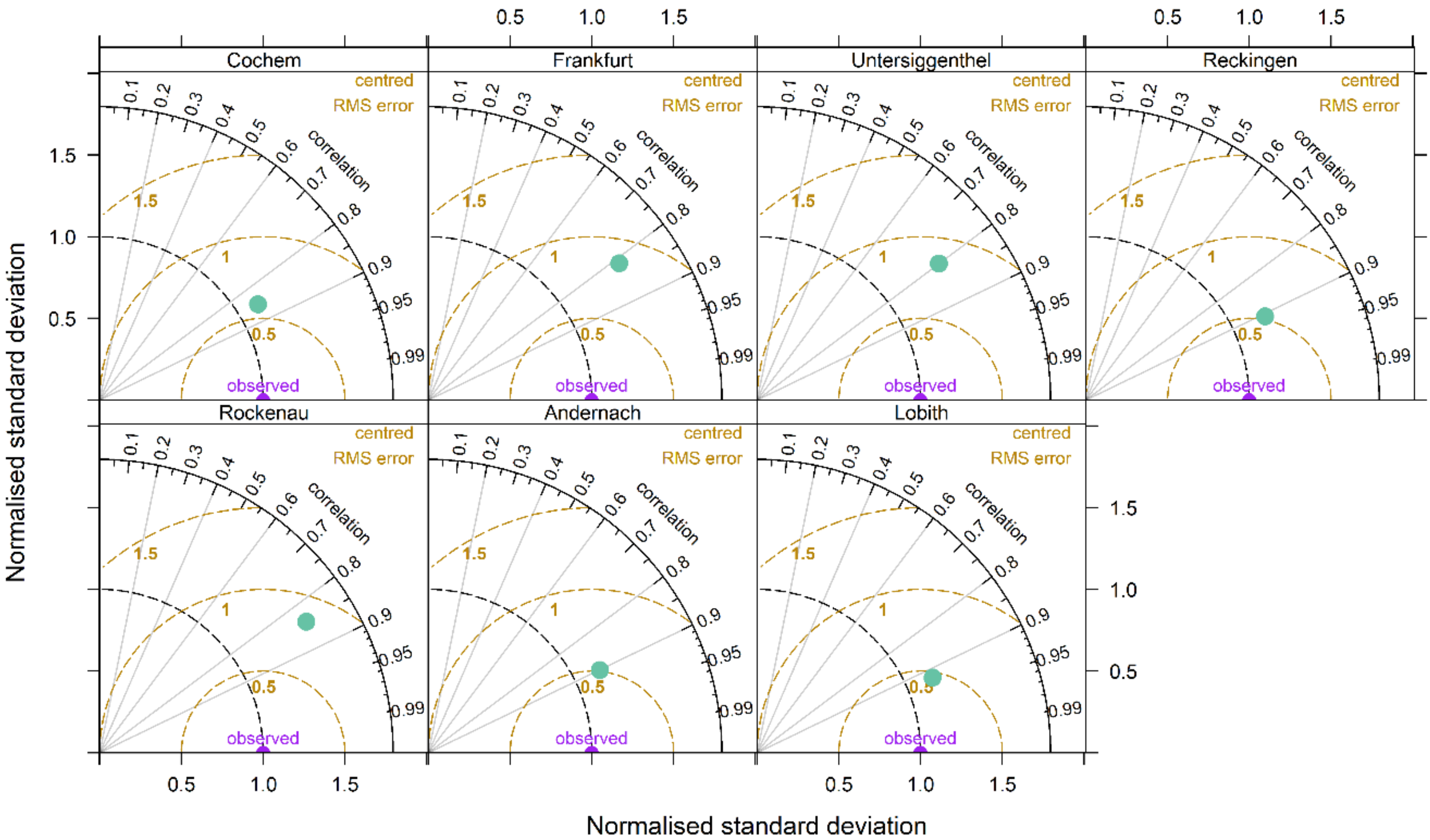

Figure S3: Taylor diagram at different locations for SPHY model. The axis represents the normalized standard deviation (sd). The purple triangle represents the observed and the green dot represents the simulated discharge respectively. These statistics compared here are correlation coefficient (R), sd and root-mean-square error. The plot is made using "TaylorDiagram" function from "open air" R package (Carslaw and Ropkins 2012).

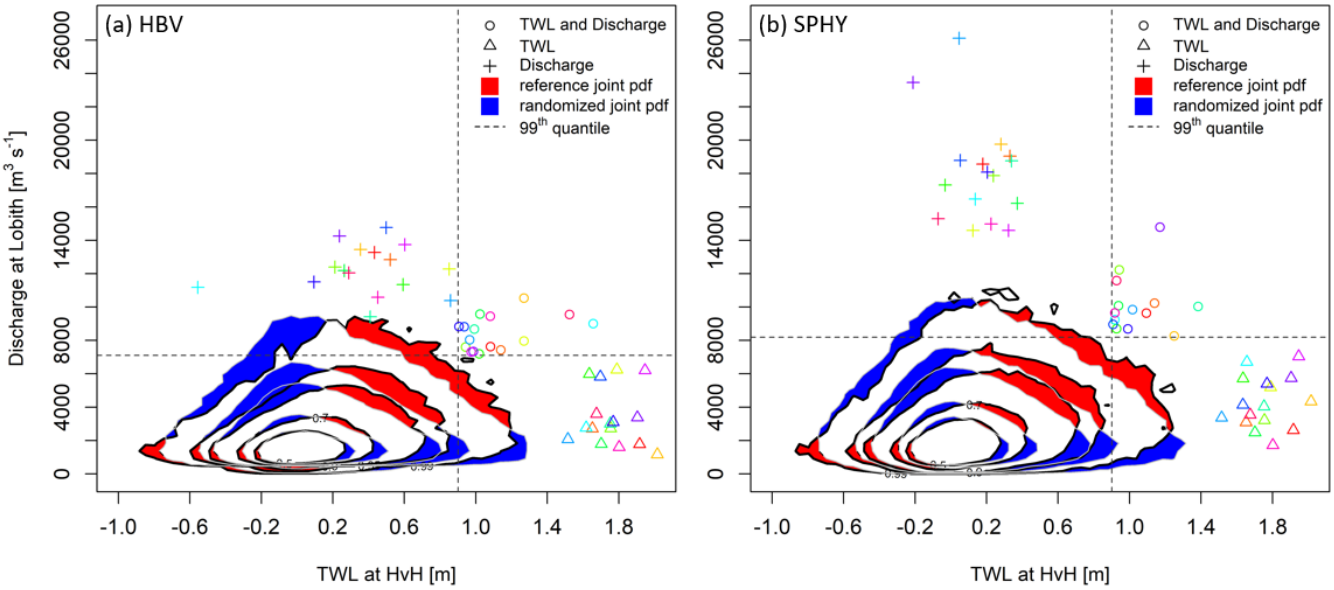

Figure S4: Joint probability density for total water level at HvH and discharge at Lobith (6-day lagged) for winter six months period (left-(a) HBV and right-(b) SPHY). The contours show different quantiles of the joint distribution (50th, 70th, 90th, 95th and 99th). Shading is used to contrast density estimates from direct model output and randomized (shuffled) data, where correlations have been artificially removed. Red/blue shading indicates regions where model data is more/less populated than the shuffled data. The thin dashed lines show the 99th quantiles of each variable in the respective dataset. The colored points correspond to the highest TWL (Δ), discharge (+) and compound events (o) for individual ensemble members.

Equation SE1: Equation for the NSE index calculation.

$$NSE = 1 - \frac{\sum_{t=1}^{T}(Q_m^t - Q_o^t)^2}{\sum_{t=1}^{T}(Q_o^t - \bar{Q}_o)^2}$$

Where;
$\bar{Q}_o$ is the mean of observed discharges.
Q_m is modeled discharges at timestep t.
Q_o is observed discharges at timestep t.

References

Carslaw D and Ropkins K 2012 openair - An R package for air quality data analysis *Environ. Model. Softw.* **27**–**28** 52–61

Huiskes I 2016 *Using Ensemble Streamflow Predictions for extreme discharge purposes in the river Rhine* (University of Twente) Online: https://www.utwente.nl/en/et/wem/education/msc-thesis/2016/huiskes.pdf