Diagnostic efficiency - a diagnostic approach for model evaluation

Robin Schwemmle[1], Dominic Demand[1], Markus Weiler[1]

Albert-Ludwigs-Universität Freiburg

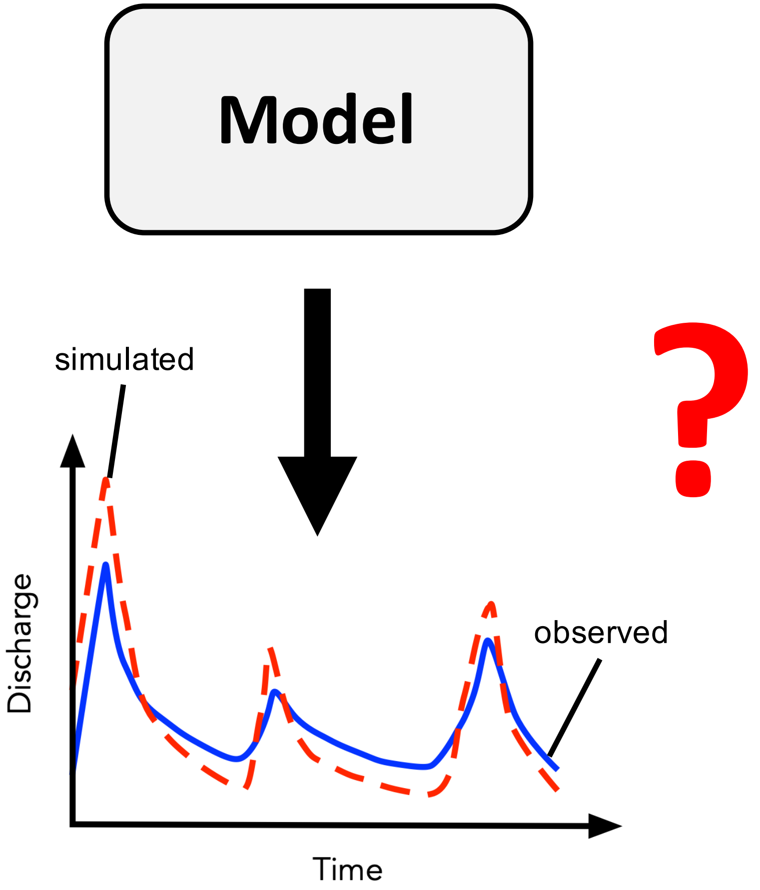
robin.schwemmle@hydrology.uni-freiburg.de

[1]Chair of Hydrology
University of Freiburg
Freiburg, Germany

UNI FREIBURG

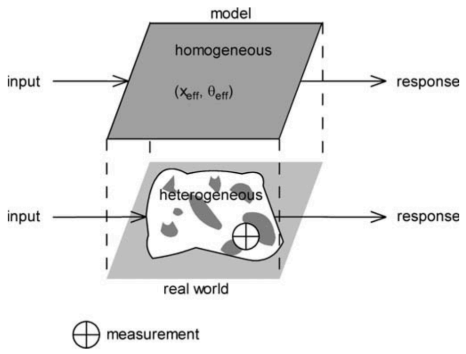
HYDROLOGY

Diagnostic model evaluation



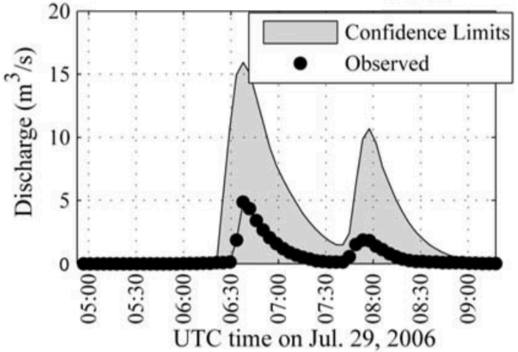
Origin of errors in hydrological simulations

Model parameters



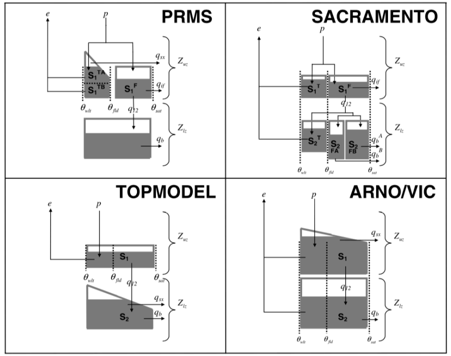
Wagener and Gupta (2005), *SERRA*

Input data



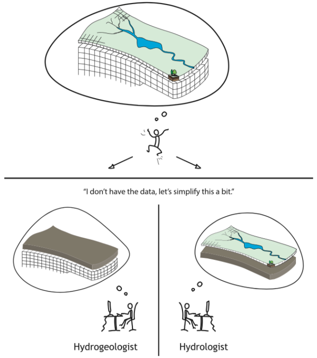
Yatheendras et al. (2008), *WRR*

Model structure



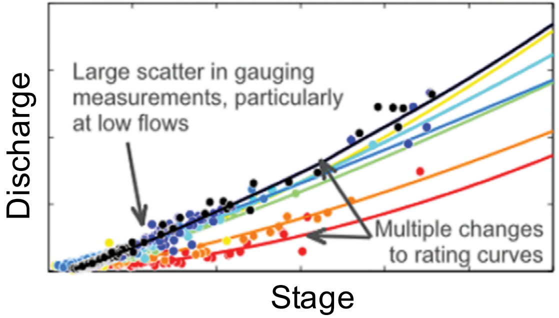
Clark et al. (2008), *WRR*

Initial and boundary conditions



Staudinger et al. (2015), *JoH*

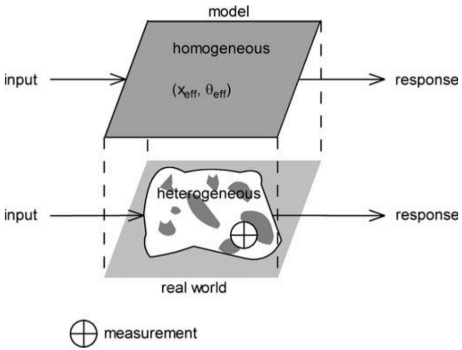
Observations



Coxon et al. (2015), *WRR*

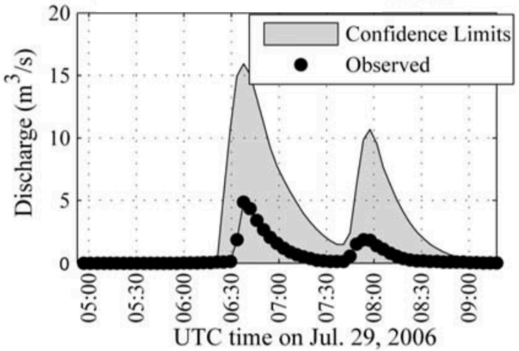
Origin of errors in hydrological simulations

Model parameters



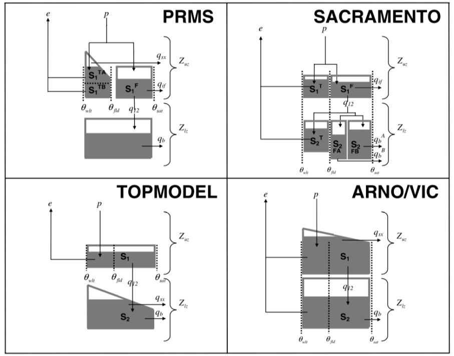
Wagener and Gupta (2005), *SERRA*

Input data



Yatheendras et al. (2008), *WRR*

Model structure



Clark et al. (2008), *WRR*

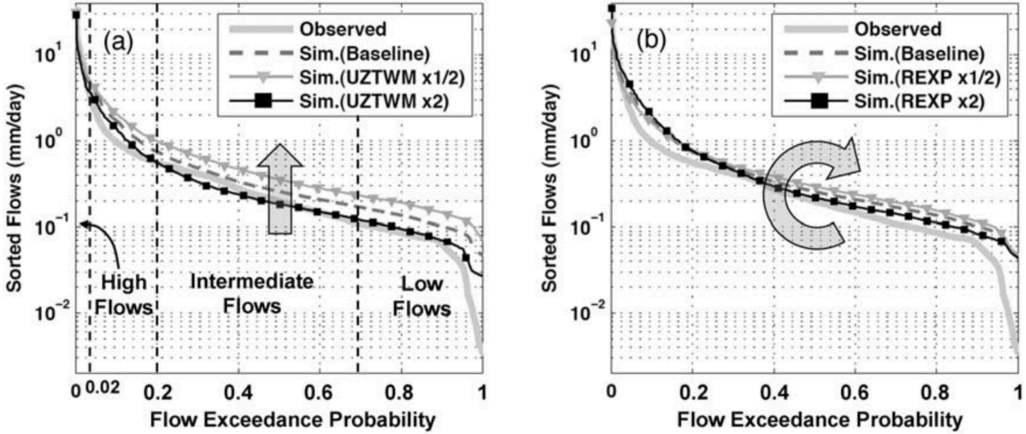
Research gap

- Traditional efficiency measures (e.g. Kling-Gupta Efficiency)

  → diagnostic meaning of a number between **-∞** and **1?**

- Diagnostic signatures (e.g. high flow bias)

  → comprehensive visualization**?**



Yilmaz et al. (2008), *WRR*

Research objectives

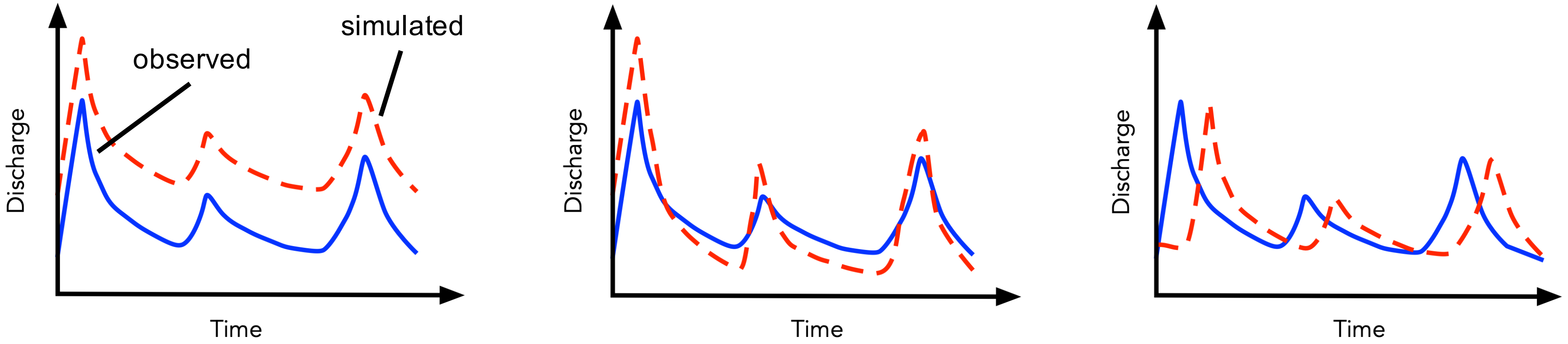
- **Defintion of a diagnostic efficiency measure** based on flow duration curve and correlation which hints to the **origin of errors**
- Visualising **contributions of different model errors**
- Provide an easily extendable **evaluation tool**

diag-eff on GitHub: https://github.com/schwemro/diag-eff

Documentation (including tutorials): https://diag-eff.readthedocs.io/en/latest/

Diagnostic Efficiency (*DE*)

$$DE = 1 - \sqrt{constant\ error + dynamic\ error + timing\ error}$$

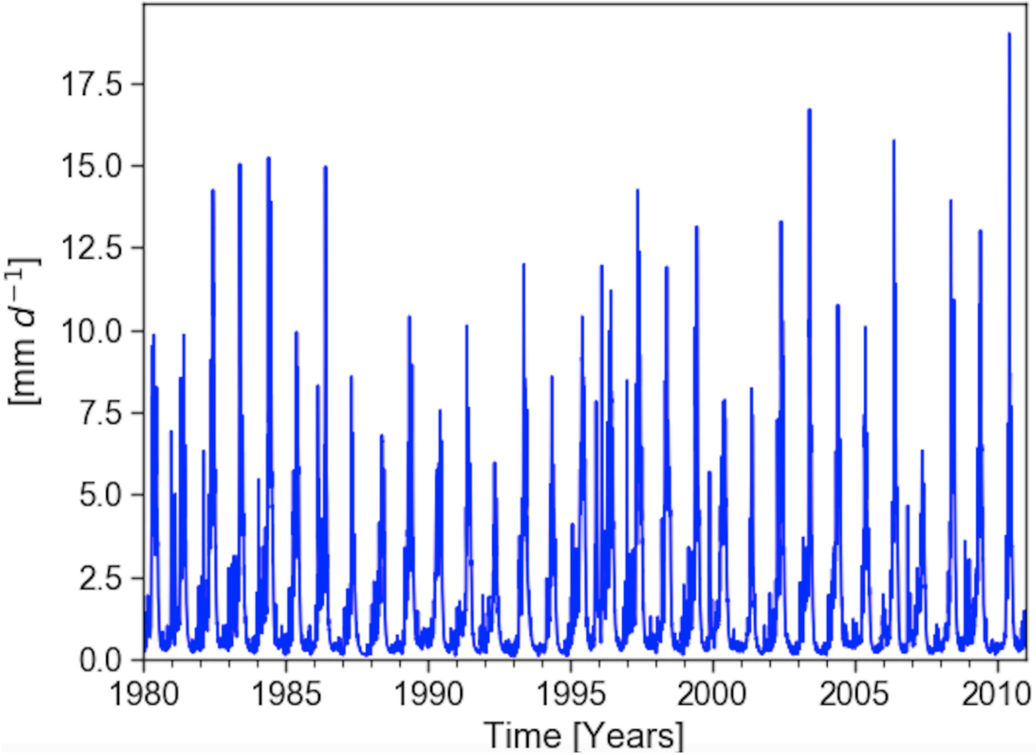


$$DE = 1 - \sqrt{\overline{B_{rel}}^2 + |B_{area}|^2 + (r-1)^2}$$

Schwemmle et al. (2020) in prep., (will be submitted to HESS shortly after EGU2020)

Mimicking errors

gauge_id: 13331500; gauge_name: Minam 160 River near Minam, OR, U.S.



CAMELS dataset; Addor et al. (2017)

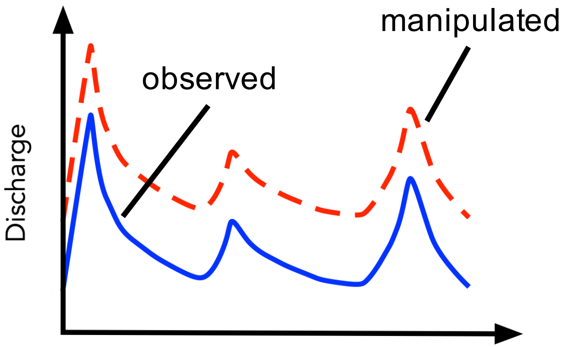
Mimicking errors by **systematic manipulation** of an **observed streamflow time series:**

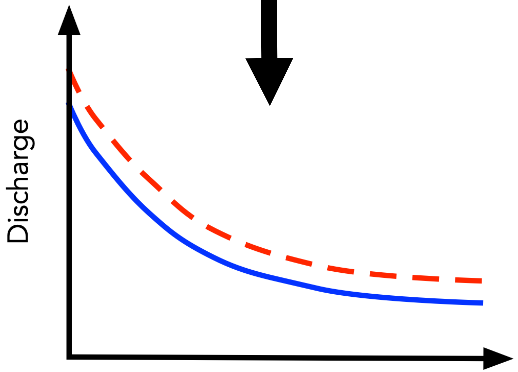
→ constant error (i)

→ dynamic error (ii)

→ timing error (iii)

Mimicking constant error (i)





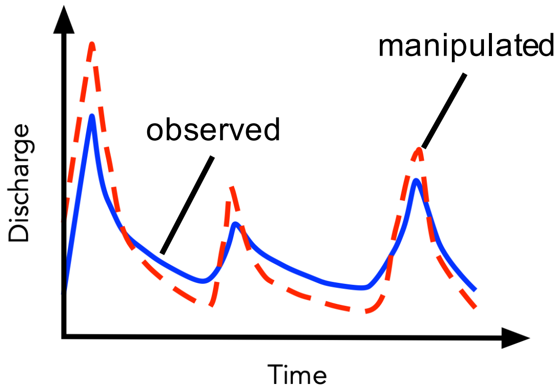
Multiplying the **observed time series** with a **constant >1** to **generate a positive offset**

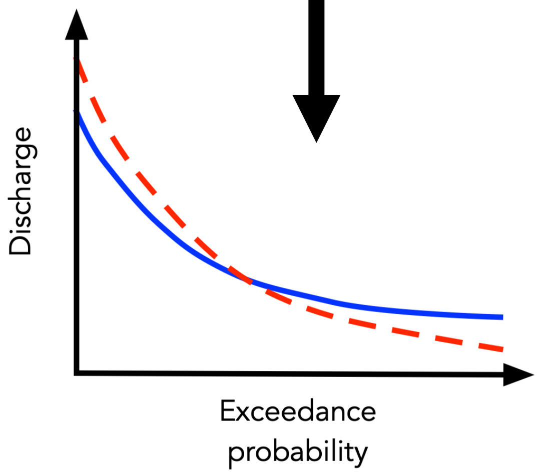
$$DE = 1 - \sqrt{\overline{B_{rel}}^2 + |B_{area}|^2 + (r-1)^2}$$

arithmetic mean of the relative bias

$$\overline{B_{rel}} = \frac{1}{N} \sum_{i=0}^{i=1} B_{rel}(\mathrm{i})$$

Mimicking dynamic error (ii)





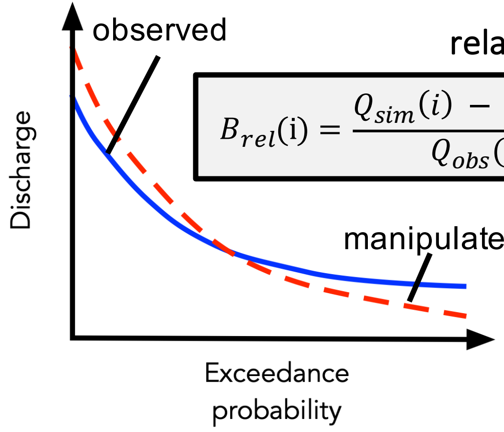
Multiplying the **observed time series** with an **linearly interpolated vector** (1.5, ..., 0.5) to **increase high flows** and **decrease low flows**

$$DE = 1 - \sqrt{\overline{B_{rel}}^2 + |B_{area}|^2 + (r-1)^2}$$

area of the residual bias

$$|B_{area}| = \int_0^1 |B_{res}(i)|\, di$$

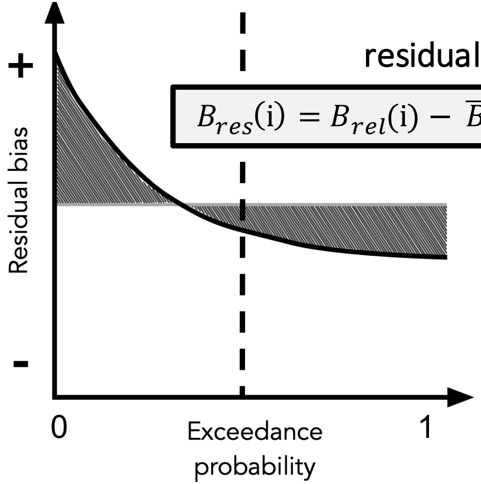
Calculation of the dynamic error term



relative bias **1**

$$B_{rel}(\mathrm{i}) = \frac{Q_{sim}(i) - Q_{obs}(i)}{Q_{obs}(i)}$$

$$DE = 1 - \sqrt{\overline{B_{rel}}^2 + |B_{area}|^2 + (r-1)^2}$$



residual bias **2**

$$B_{res}(\mathrm{i}) = B_{rel}(\mathrm{i}) - \overline{B_{rel}}$$

area of the residual bias **3**

$$|B_{area}| = \int_0^1 |B_{res}(i)|\ di$$

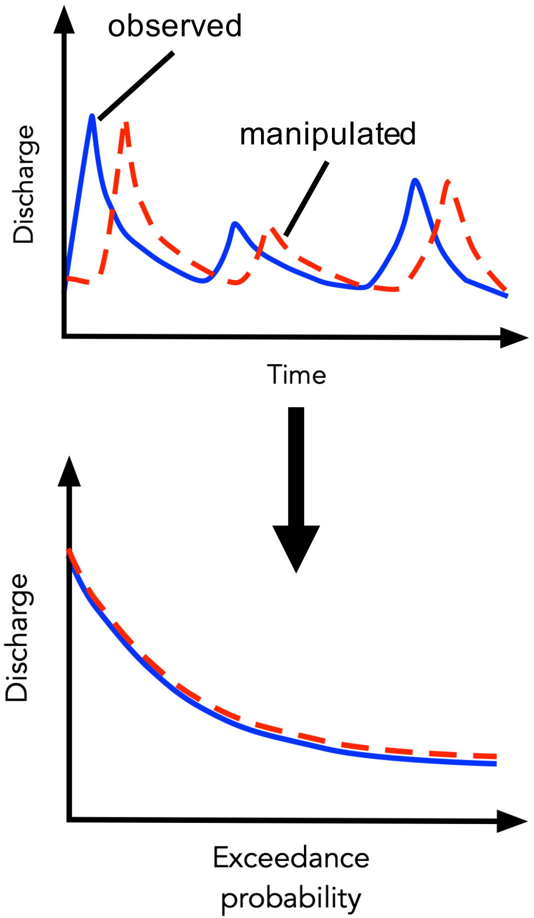
direction of dynamic error **4**

$$B_{dir} = \int_0^{0.5} B_{res}(i)\ di$$

slope of the residual bias **5**

$$B_{slope} = \begin{cases} |B_{area}| \cdot (-1), & B_{dir} > 0 \\ |B_{area}| \quad , & B_{dir} < 0 \\ 0 \quad , & B_{dir} = 0 \end{cases}$$

Mimicking timing error (iii)

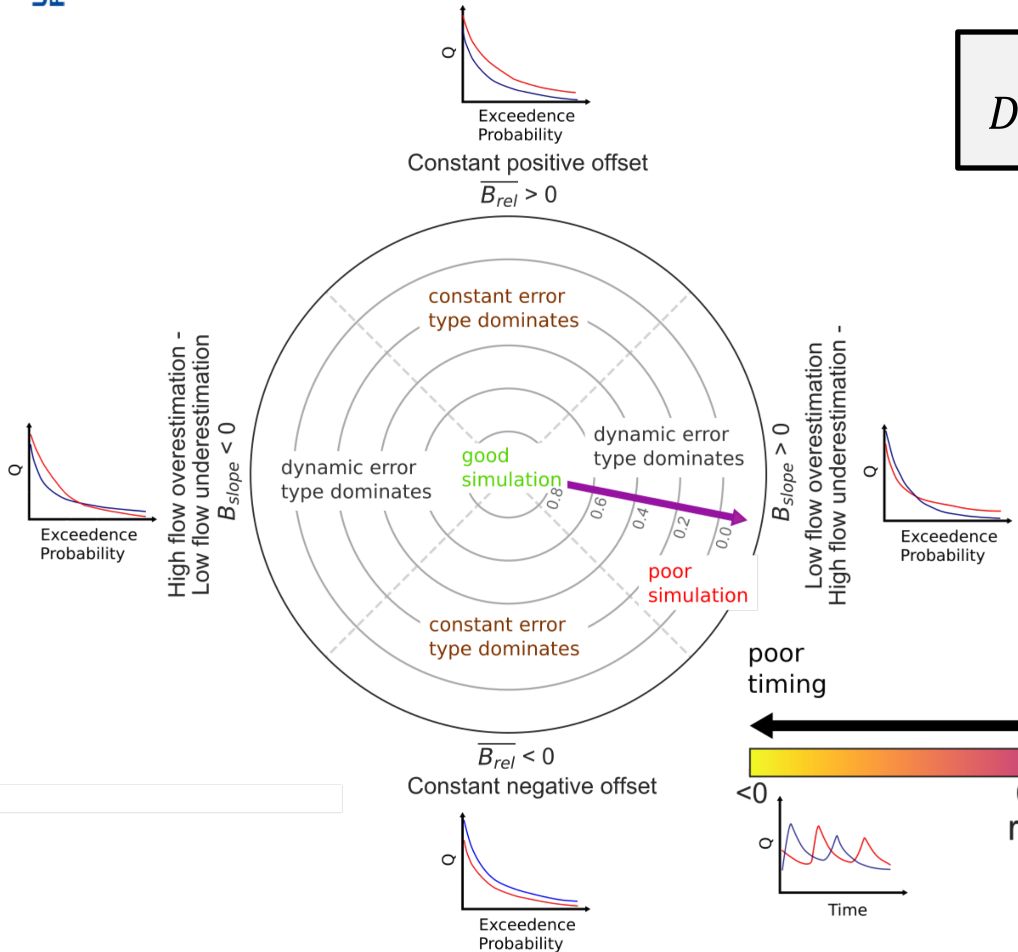


Randomizing the order of the observed time series

$$DE = 1 - \sqrt{\overline{B_{rel}}^2 + |B_{area}|^2 + (r-1)^2}$$

linear correlation between simulations and observations

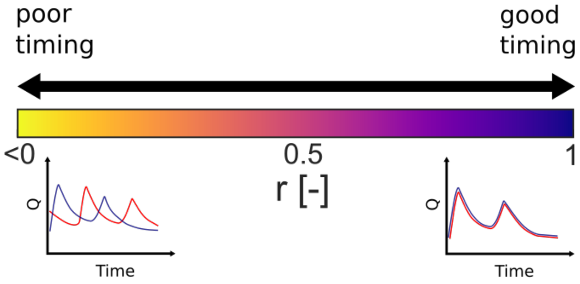
Diagnostic polar plot



$$DE = 1 - \sqrt{\overline{B_{rel}}^2 + |B_{area}|^2 + (r-1)^2}$$

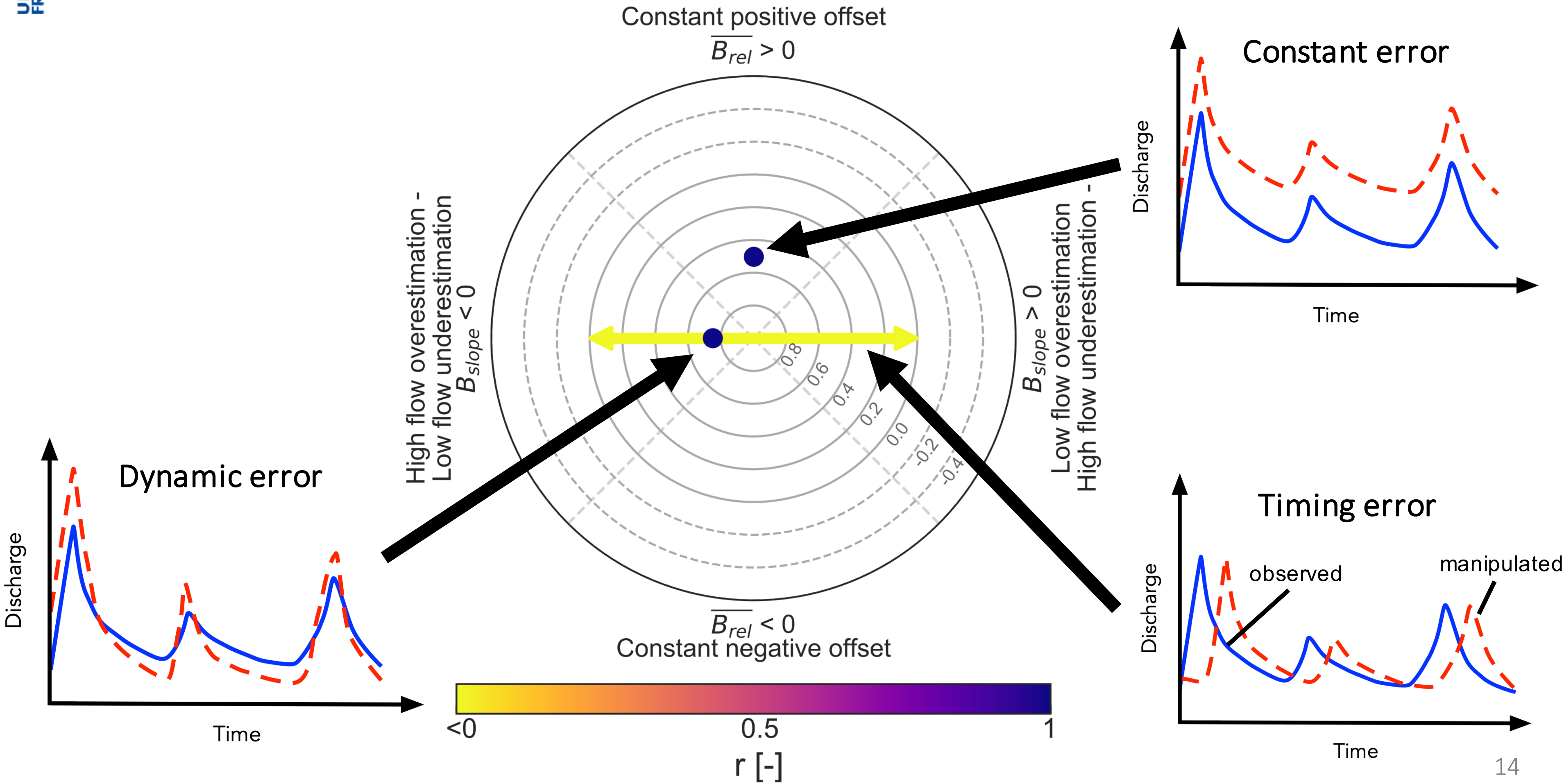
ratio between constant and dynamic error

$$\varphi = arctan2(\overline{B_{rel}}, B_{slope})$$



Schwemmle et al. (2020) in prep., (will be submitted to HESS shortly after EGU2020)

Proof of concept



Comparison to *KGE*

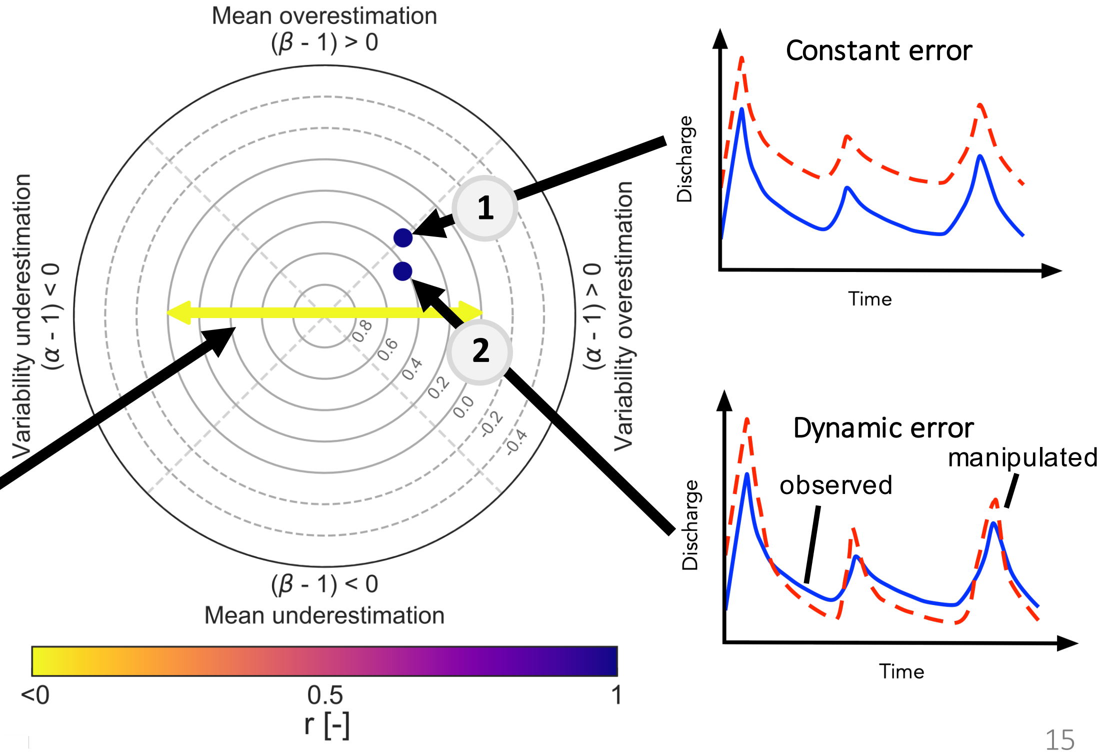
$$KGE = 1 - \sqrt{(\beta - 1)^2 + (\alpha - 1)^2 + (r - 1)^2}$$

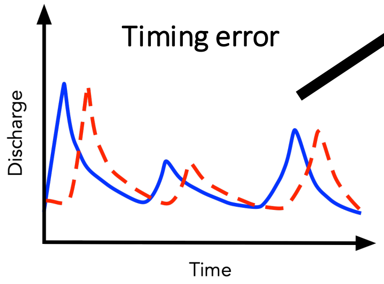
ratio between bias error and flow variability error

$$\varphi = arctan2(\beta - 1, \alpha - 1)$$

1 Polar plot for *KGE* falsely indicate an equal share of bias error and variability error

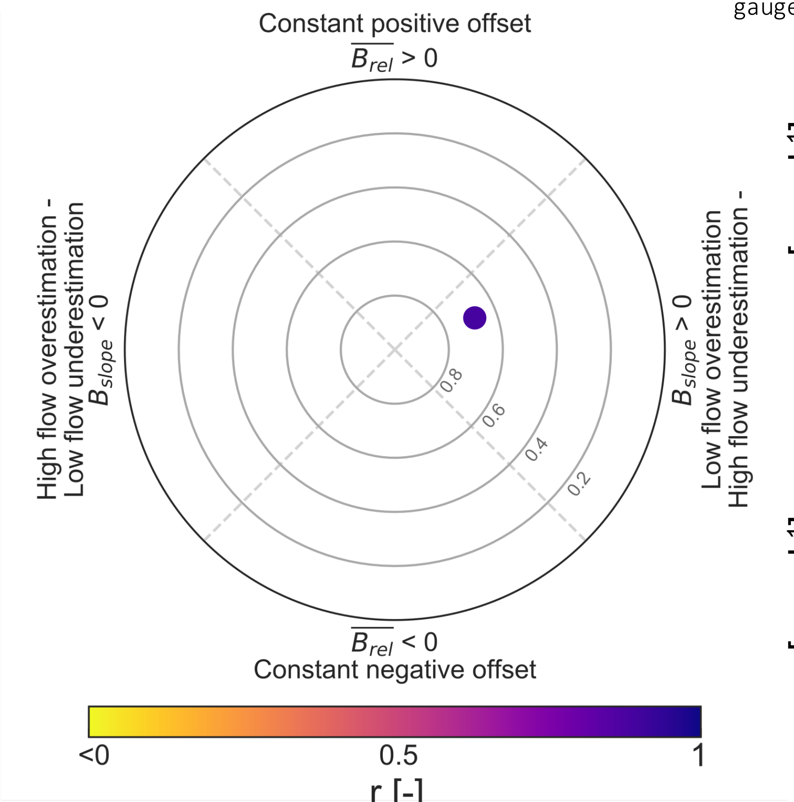
2 Contribution of bias error is falsely displayed

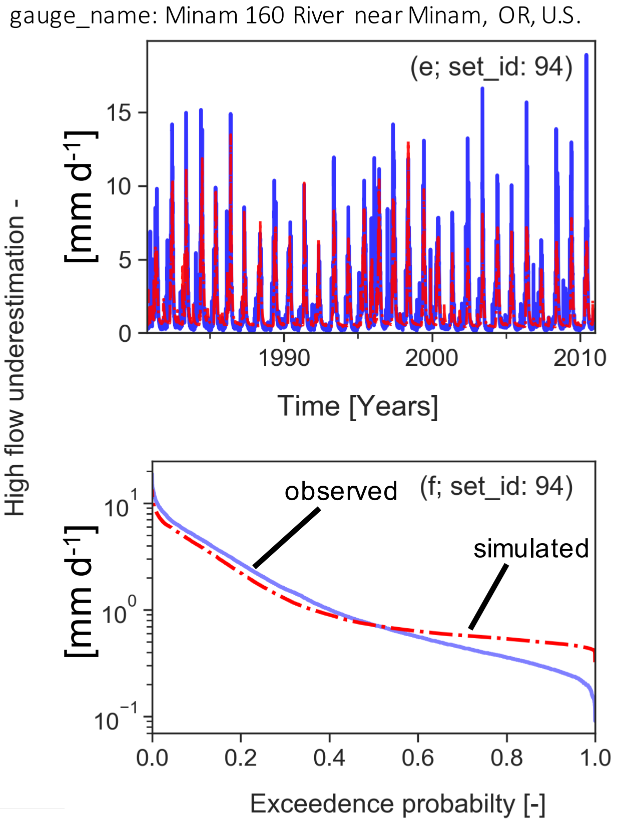




Real case example (Newman et al. (2015))

- positive dynamic error dominates
- slight positive constant error
- lowest share by timing

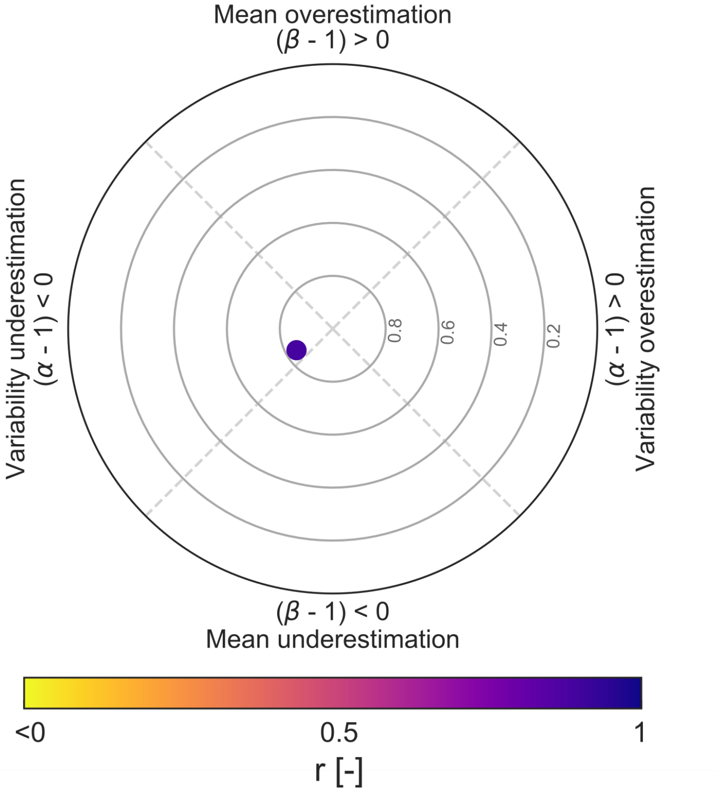


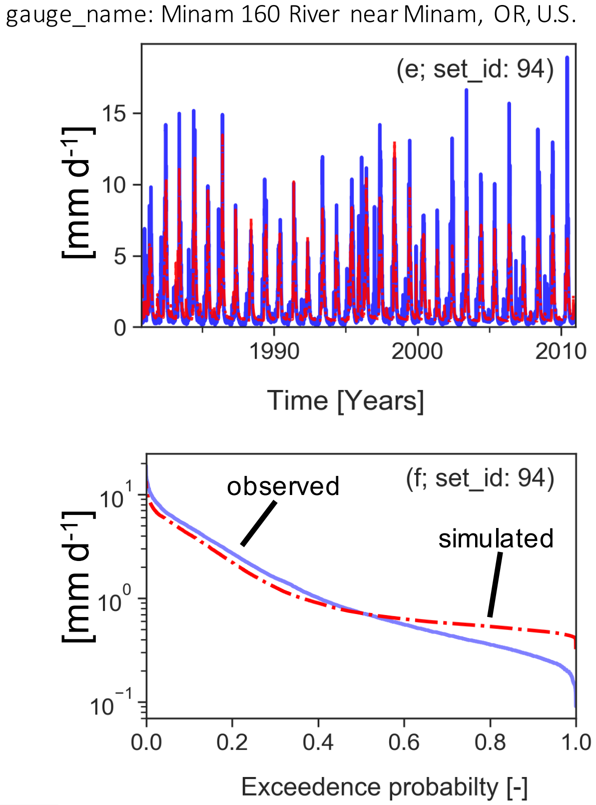


Schwemmle et al. (2020) in prep., (will be submitted to HESS shortly after EGU2020)

Comparison to *KGE*

Polar plot of *KGE* leads to different error contributions than diagnostic polar plots





Schwemmle et al. (2020) in prep., (will be submitted to HESS shortly after EGU2020)

Conclusions

- theoretical proof of concept and real world applicability
- diagnostic polar plots provide hints on the origin of errors
- blueprint for systematic development of other diagnostic efficiency measures

diag-eff on GitHub: https://github.com/schwemro/diag-eff

Documentation (including tutorials): https://diag-eff.readthedocs.io/en/latest/

robin.schwemmle@hydrology.uni-freiburg.de

References

Addor, N., Newman, A. J., Mizukami, N., and Clark, M. P.: The CAMELS data set: catchment attributes and meteorology for large-sample studies, in, version 2.0 ed., Boulder, CO: UCAR/NCAR, 2017.

Clark, M. P., Slater, A. G., Rupp, D. E., Woods, R. A., Vrugt, J. A., Gupta, H. V., Wagener, T., and Hay, L. E.: Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models, Water Resources Research, 44, 10.1029/2007wr006735, 2008.

Coxon, G., Freer, J., Westerberg, I. K., Wagener, T., Woods, R., and Smith, P. J.: A novel framework for discharge uncertainty quantification applied to 500 UK gauging stations, Water Resources Research, 51, 5531-5546, 10.1002/2014wr016532, 2015.

Newman, A. J., Clark, M. P., Sampson, K., Wood, A., Hay, L. E., Bock, A., Viger, R. J., Blodgett, D., Brekke, L., Arnold, J. R., Hopson, T., and Duan, Q.: Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: data set characteristics and assessment of regional variability in hydrologic model performance, Hydrol. Earth Syst. Sci., 19, 209-223, 2015.

Staudinger, M., Stoelzle, M., Cochand, F., Seibert, J., Weiler, M., and Hunkeler, D.: Your work is my boundary condition!: Challenges and approaches for a closer collaboration between hydrologists and hydrogeologists, Journal of Hydrology, 571, 235-243, 10.1016/j.jhydrol.2019.01.058, 2019.

Wagener, T., and Gupta, H. V.: Model identification for hydrological forecasting under uncertainty, Stochastic Environmental Research and Risk Assessment, 19, 378-387, 10.1007/s00477-005-0006-5, 2005.

Yatheendradas, S., Wagener, T., Gupta, H., Unkrich, C., Goodrich, D., Schaffner, M., and Stewart, A.: Understanding uncertainty in distributed flash flood forecasting for semiarid regions, Water Resources Research, 44, 10.1029/2007wr005940, 2008.

Yilmaz, K. K., Gupta, H. V., and Wagener, T.: A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model, Water Resources Research, 44, 10.1029/2007wr006716, 2008.