Authors' responses to comments

Regionalization of hydrological model parameters using gradient boosting machine

Zhihong Song, Jun Xia, Gangsheng Wang, Dunxian She, Chen Hu, Si Hong

We really appreciate the editor's and anonymous reviewers' valuable suggestions and comments on the manuscript. All suggestions are helpful to improve this manuscript. We have addressed the comments point-by-point and revised the manuscript accordingly.

Responses to Referee #1's comments

The manuscript is interesting and clearly written. I have a single comment; in particular, what are the practical reasons for regionalizing the parameters of the hydrological model using regression algorithms, given that available data cover uniformly the region of the study; therefore, regionalization is not needed in my opinion (parameters are already available).

Response: We sincerely appreciate the referee's positive comments on the manuscript. Our response to the reviewer's concern on regionalizing the model parameters is as follows and has been revised in the discussion part of the manuscript (Sec. 4.4, Page 29).

"Hydrologic models often rely on regionalization approaches to transfer information from small to large spatial scale (e.g., from gridcell to subbasin, watershed, and regional scale) (Beck et al., 2020; Mizukami et al., 2017), and from gauged to ungauged catchments (He et al., 2011; Hrachowitz et al., 2013; Pagliero et al., 2019; Parajka et al., 2013).

In this study, though parameters were calibrated and thus available at each gridcell,

the parameter values at around 450 gridcells were not reliable owning to poor model performance (i.e., KEG < 0) (Knoben et al., 2018; Koskinen et al., 2017; Sutanudjaja et al., 2018). Therefore, we only used the calibrated parameters with KGE ≥ 0 (i.e., representing better model performance) for regionalization of parameters. The model performance for 53% of these gridcells (with KGE < 0 prior to regionalization) were improved when we re-ran the model with regionalized parameters. Particularly, the KGE values in 37% of the gridcells (with KGE < 0 prior to regionalization) became positive, indicating a substantive improvement of the modeling performance.

Even though the parameters were well calibrated and available at each gridcell, one might think whether and which topographic and edaphic properties mediate these hydrological parameters. Our machine learning (i.e., GBM) based regionalization of parameters enables to estimate six key hydrological parameters using site-specific characteristics. Following the regionalization of parameters, our results of variable importance quantitatively indicate that the runoff generation parameters are majorly controlled by slope, saturated soil moisture content, and elevation. Moreover, the terrain attributes significantly regulate the runoff processes in relatively humid regions, while the saturated soil moisture content becomes a limiting factor in arid areas. The regionalization of parameters will improve our mechanistic understanding of the runoff generation processes and associated key hydrological parameters under different topographic and edaphic conditions."

Responses to Referee #2's comments

This discussion paper examines the use of gradient boosting machine learning to model the dependency of model parameters and estimate model parameters for four climate regions in China. The study also enhances the DTVGM model by incorporating the Penman-Monteith-Leuning (PML) equation. The discussion paper is well-presented, the study design is nicely conceived, and the results and discussion are presented fairly, with references to supporting figures, tables or cited literature, where appropriate.

I only have minor comments to enhance what is already a high quality manuscript.

Response: We sincerely appreciate the referee's positive comments on the manuscript. All the comments and suggestions have been replied to below and have been addressed in the revision.

(1) In several places, the model is referred to as the "China-wide hydrological model" (see L105). This is somewhat confusing. Is this a formal name for the hydrological model? If so, I would expect to see a citation after the phrase to reference the use of this name. If this is not the formal name, it may be more helpful to say "We ran a hydrological model developed by Beck et al (2020), for country of China in a spatially distributed…" In Section 2.1, the title could be changed to "Application of the hydrological model across China."

Response: Thanks for pointing this out! "China-wide hydrological model" is not a formal name for the model. We have rewritten the sentence to address this issue: "We ran a hydrological model, i.e., Distributed Time-Variant Gain Model with the Penman-Monteith-Leuning equation (DTVGM-PML) developed in this study, for country of China in a spatially distributed manner" [Line 106, Page 5]. We have also changed the title of Section 2.1 to "Application of the hydrological model across China". Other corresponding information has been revised [Line 103, Page 5; Line 130, Page 6; Line 549, Page 30].

(2) In my reading, there was some confusion about how you were able to compute evaluation metrics for all 15,640 grid cells (see Figure 5 for example). If you know the "truth" for runoff and ET at every grid cell, then why do you need a regionalization model?

Response: Our response to the reviewer's concern on regionalizing the model parameters is as follows and has been revised in the discussion part of the manuscript [Sec. 4.4, Page 29].

"Hydrologic models often rely on regionalization approaches to transfer information from small to large spatial scale (e.g., from gridcell to subbasin, watershed, and regional scale) (Beck et al., 2020; Mizukami et al., 2017), and from gauged to ungauged catchments (He et al., 2011; Hrachowitz et al., 2013; Pagliero et al., 2019; Parajka et al., 2013).

In this study, though parameters were calibrated and thus available at each gridcell, the parameter values at around 450 gridcells were not reliable owning to poor model performance (i.e., KEG < 0) (Knoben et al., 2018; Koskinen et al., 2017; Sutanudjaja et al., 2018). Therefore, we only used the calibrated parameters with KGE ≥ 0 (i.e., representing better model performance) for regionalization of parameters. The model performance for 53% of these gridcells (with KGE < 0 prior to regionalization) were improved when we re-ran the model with regionalized parameters. Particularly, the KGE values in 37% of the gridcells (with KGE < 0 prior to regionalization) became positive, indicating a substantive improvement of the modeling performance.

Even though the parameters were well calibrated and available at each gridcell, one might think whether and which topographic and edaphic properties mediate these hydrological parameters. Our machine learning (i.e., GBM) based regionalization of parameters enables to estimate six key hydrological parameters using site-specific characteristics. Following the regionalization of parameters, our results of variable importance quantitatively indicate that the runoff generation parameters are majorly

controlled by slope, saturated soil moisture content, and elevation. Moreover, the terrain attributes significantly regulate the runoff processes in relatively humid regions, while the saturated soil moisture content becomes a limiting factor in arid areas. The regionalization of parameters will improve our mechanistic understanding of the runoff generation processes and associated key hydrological parameters under different topographic and edaphic conditions."

(3) Also in Figure 5, it would be helpful to show the comparison of the model performance with and without the PML addition so that one can see in quantifiable terms how the addition of the equation improves the calibration and validation performance.

Response: Thanks for the valuable suggestion! We have shown the comparison of model performance with and without PML addition in Supplementary Fig. S2 (see Fig. R1). We have also added the following text to describe the results:

"A great deal of previous studies have highlighted the importance of incorporating the vegetation change information into hydrological models to achieve better performance in hydrological simulations (Donohue et al., 2007, 2010; Gerten, 2013; Ivanov et al., 2008; Lei et al., 2014; Thompson et al., 2011). Additionally, it has been demonstrated that coupling the PML equation into hydrological models can improve the hydrological simulations under vegetation greening conditions (Bai et al., 2018; Li et al., 2009; Zhang et al., 2009; Zhou et al., 2013)." [Line 135–140, Page 7]

"We also performed the comparison of model performance in hydrological simulation between the original DTVGM (without PML) and DTVGM-PML. As shown in Figure R1, the KGE and PBIAS values of runoff simulation (Fig. R1a, b) by DTVGM were lower than those from DTVGM-PML. The median KGE and PBIAS values of ET simulation (Fig. R1c, d) were comparable between the two models. In summary, DTVGM-PML can help to improve hydrological simulations relative to DTVGM. Additionally, the consideration of vegetation dynamics by the PML equation

in DTVGM-PML would improve the mechanistic understanding of hydrological response under vegetation greening, which is lacking in DTVGM." [Line 278–285, Page 14,15]
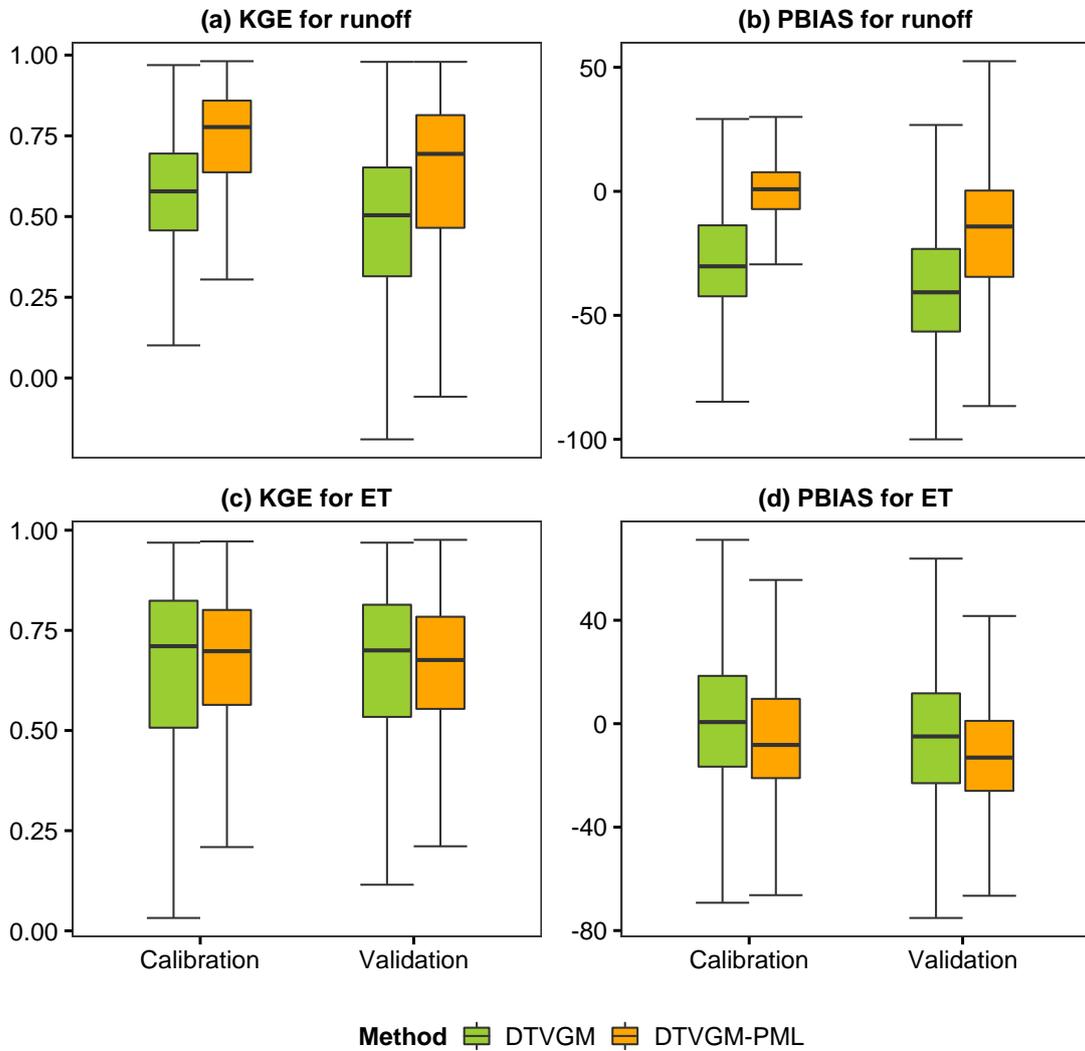

Figure R1. Comparison of model performance in runoff (a: KGE, and b: PBIAS) and ET (c: KGE, and d: PBIAS) simulation between DTVGM and DTVGM-PML in the calibration and validation periods. KGE denotes the Kling-Gupta efficiency. PBIAS denotes the percent bias".

(4) Section 2.4: You provide an excellent description of the evaluation criteria; however, in your use of the Taylor skill score, could you clarify how the skill is determined for

model parameters when you cannot know the true value of the parameters? L207-209 were somewhat confusing. This could be my lack of familiarity with the TSS, but it may be helpful to look over those lines to see if you could improve the explanation there.

Response: Thanks for your comments. We have provided more detailed information of the Taylor skill score (TSS) in the revised paper as follows:

"The Taylor skill score, as a comprehensive metric of correlation coefficient, standard deviation, and root mean square error, has been widely used in model evaluation (Mohan and Bhaskaran, 2019; Taylor, 2001)." [Line 214–216, Page 10]

$$TSS = \frac{4(1+r)^4}{\left(SDR+\frac{1}{SDR}\right)^2 (1+r_0)^4}, \tag{R1}$$

$$r = \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{\sigma_X \sigma_Y}, \tag{R2}$$

$$SDR = \frac{\sigma_X}{\sigma_Y}, \tag{R3}$$

$$\sigma_X = \sqrt{\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n}}, \ \sigma_Y = \sqrt{\frac{\sum_{i=1}^{n}(Y_i-\bar{Y})^2}{n}}, \tag{R4}$$

where X, Y are the calibrated parameter and regionalized parameter in DTVGM-PML (e.g., the runoff generation parameters, g_1 and g_2), respectively; the subscript i denotes the ith sample of the gridded parameter; $\bar{X}$, $\bar{Y}$ are the mean values of X and Y, respectively; σ_X, σ_Y are the spatial standard deviation of calibrated parameter and regionalized parameter, respectively; r represents the spatial correlation coefficient between X and Y; r_0 is the maximum correlation attainable and usually set to 0.999; SDR is the ratio of σ_X to σ_Y; and n is the total number of values for X (and Y)." [Line 227–237, Page 11]

"We also calculated the Taylor skill scores (TSS) with the grid-scale calibrated parameters as the reference parameters to evaluate the regionalization model performance for estimating each parameter at each grid. As shown in Figure S6 (Fig. R2), GBM obviously outperformed MLR with a higher TSS, suggesting that the GBM-regionalized parameters presented a higher spatial agreement with reference parameter

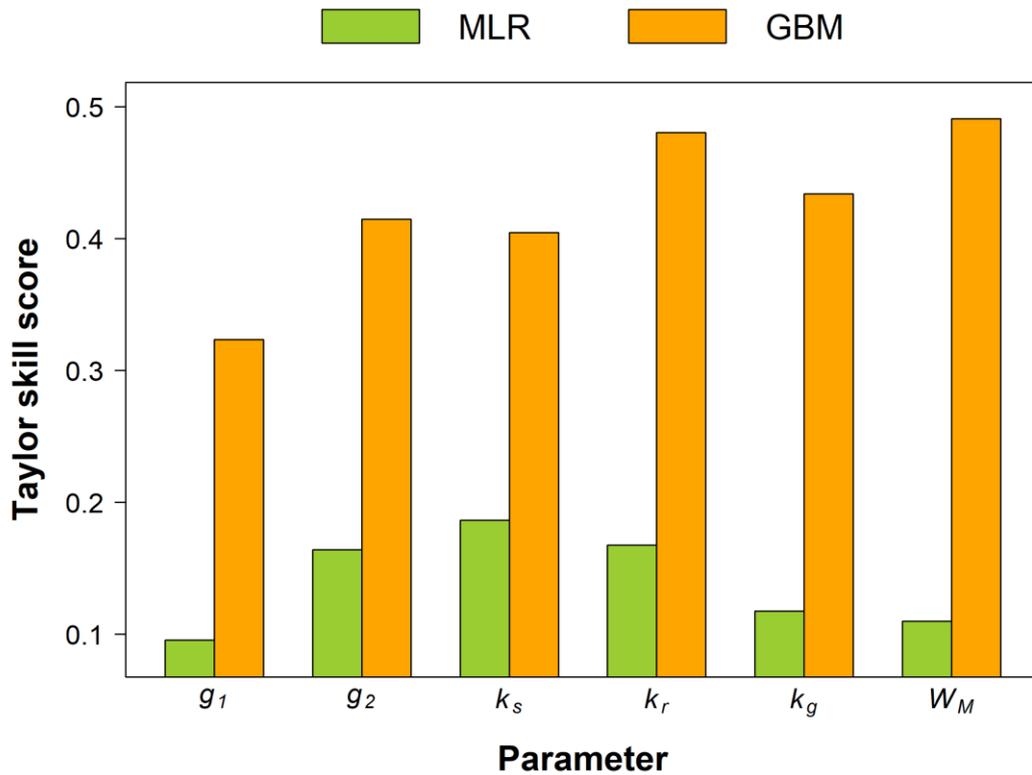values than the MLR-generated parameters." [Line 311–315, Page 16,17]

Figure R2. Taylor skill scores (TSS) of each parameter generated from the multiple linear regression (MLR) and the gradient boosting machine (GBM). The Taylor skill scores were computed using parameters from all grid cells across China.

(5) In the figure captions, some of the acronyms are spelled out, while others are not. It may be best to spell out all abbreviated words and their abbreviations in the captions so the reader does not have to refer back to the text.

Response: Thanks for the suggestion. We have modified corresponding Figure captions.

The figure caption of Fig. 4 "Spatial patterns of mean annual runoff (a) and ET (b) simulations by the DTVGM-PML during 1982–2012" has been modified as "Spatial patterns of mean annual runoff (a) and evapotranspiration (ET) (b) simulations by the DTVGM-PML during 1982–2012.".

The figure caption of Fig. 5 "Model performance of runoff (KGE: a, and PBIAS: b) and ET (KGE: c, and PBIAS: d) simulations in calibration and validation periods. The boxplot was generated using data from a total of 15640 grid cells over China" has been modified as "Model performance of runoff (KGE: a, and PBIAS: b) and evapotranspiration (ET) (KGE: c, and PBIAS: d) simulations in calibration and validation periods. The boxplot was generated using data from a total of 15640 grid cells over China. KGE denotes the Kling-Gupta efficiency. PBIAS denotes the percent bias".

The figure caption of Fig. 6 has been modified as "Performance evaluation of the multiple linear regression (MLR) and the gradient boosting machine (GBM) for six parameters, (a) g_1, (b) g_2, (c) k_s, (d) k_r, (e) k_g, (f) W_M, in humid region. MLR and GBM denote the multiple linear regression with stepwise selection and the gradient boosting machine model. The boxplot is generated from the 10 samples in k-fold cross validation. We use the non-parametric Kruskal-Wallis (KW) test to determine the significance of difference in the performance between MLR and GBM at a significance level of 0.05".

The figure caption of Fig. 7 has been modified as "Spatial patterns of model parameters: (a) g_1, (b) g_2, (c) k_s, (d) k_r, (e) k_g, (f) W_M, derived from (1) calibration (CLB), (2) multiple linear regression (MLR) and (3) gradient boosting machine (GBM)".

The figure caption of Fig. 8 has been modified as "Cumulative distribution functions (CDFs) of KGE for runoff simulation based on three parameter sets (black lines for calibration (CLB), blue lines for multiple linear regression (MLR), red lines for gradient boosting machine (GBM)) in the validation period over four climatic zones. KGE denotes the Kling-Gupta efficiency."

The figure caption of Fig. 10 has been modified as "Relative importance of explanatory variables from the gradient boosting machine (GBM) model for (a) g_1, (b) g_2, (c) k_s, (d) k_r, (e) k_g, (f) W_M in four climatic zones (1, humid; 2, semi-humid; 3, semi-arid; 4, arid). The mean relative importance over four regions for each parameter is shown in the right column (a5-g5). The mean relative importance over six parameters

for each climatic zone is shown in the bottom row (g1-g5). Explanatory variables include elev (Elevation), slp (Slope), snd (Sand content), slt (Silt content), cly (Clay content), fc (Field capacity), pw (Wilting point), thr (Residual moisture content), ths (Saturated moisture content), ksat (Saturated hydraulic conductivity)."

References

Bai, P., Liu, X., Zhang, Y., and Liu, C.: Incorporating vegetation dynamics noticeably improved performance of hydrological model under vegetation greening, Sci. Total Environ., 643, 610-622, 10.1016/j.scitotenv.2018.06.233, 2018

Beck, H.E., Pan, M., Lin, P., Seibert, J., Dijk, A.I.J.M., and Wood, E.F.: Global Fully Distributed Parameter Regionalization Based on Observed Streamflow From 4,229 Headwater Catchments, Journal of Geophysical Research: Atmospheres, 125(17), 10.1029/2019JD031485, 2020

Donohue, R.J., Roderick, M.L., and Mcvicar, T.R.: On the importance of including vegetation dynamics in Budyko's hydrological model, Hydrol. Earth Syst. Sci., 11(2), 983-995, 10.5194/hess-11-983-2007, 2007

Donohue, R.J., Roderick, M.L., and Mcvicar, T.R.: Can dynamic vegetation information improve the accuracy of Budyko's hydrological model? J. Hydrol., 390(1-2), 23-34, 10.1016/j.jhydrol.2010.06.025, 2010

Gerten, D.: A vital link: water and vegetation in the Anthropocene, Hydrol. Earth Syst. Sc., 17(10), 3841-3852, 10.5194/hess-17-3841-2013, 2013

He, Y., Bárdossy, A., and Zehe, E.: A review of regionalisation for continuous streamflow simulation, Hydrol. Earth Syst. Sc., 15(11), 3539-3553, 10.5194/hess-15-3539-2011, 2011

Hrachowitz, M., Savenije, H.H.G., Blöschl, G., Mcdonnell, J.J., Sivapalan, M., Pomeroy, J.W., Arheimer, B., Blume, T., Clark, M.P., Ehret, U., Fenicia, F., Freer, J.E., Gelfan, A., Gupta, H.V., Hughes, D.A., Hut, R.W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P.A., Uhlenbrook, S., Wagener, T., Winsemius, H.C., Woods, R.A., Zehe, E., and Cudennec, C.: A decade of Predictions in Ungauged Basins (PUB)—a review, Hydrological Sciences Journal, 58(6), 1198-1255, 10.1080/02626667.2013.803183, 2013

Ivanov, V.Y., Bras, R.L., and Vivoni, E.R.: Vegetation-hydrology dynamics in complex terrain of semiarid areas: 1. A mechanistic approach to modeling dynamic feedbacks, Water Resour. Res., 44(3),

10.1029/2006WR005588, 2008

Knoben, W.J.M., Woods, R.A., and Freer, J.E.: A Quantitative Hydrological Climate Classification Evaluated With Independent Streamflow Data, Water Resour. Res., 54(7), 5088-5109, https://doi.org/10.1029/2018WR022913, 2018

Koskinen, M., Tahvanainen, T., Sarkkola, S., Menberu, M.W., Laurén, A., Sallantaus, T., Marttila, H., Ronkanen, A., Parviainen, M., Tolvanen, A., Koivusalo, H., and Nieminen, M.: Restoration of nutrient-rich forestry-drained peatlands poses a risk for high exports of dissolved organic carbon, nitrogen, and phosphorus, Sci. Total Environ., 586, 858-869, https://doi.org/10.1016/j.scitotenv.2017.02.065, 2017

Lei, H., Yang, D., and Huang, M.: Impacts of climate change and vegetation dynamics on runoff in the mountainous region of the Haihe River basin in the past five decades, J. Hydrol., 511, 786-799, 10.1016/j.jhydrol.2014.02.029, 2014

Li, H., Zhang, Y., Chiew, F.H.S., and Xu, S.: Predicting runoff in ungauged catchments by using Xinanjiang model with MODIS leaf area index, J. Hydrol., 370(1-4), 155-162, 10.1016/j.jhydrol.2009.03.003, 2009

Mizukami, N., Clark, M.P., Newman, A.J., Wood, A.W., Gutmann, E.D., Nijssen, B., Rakovec, O., and Samaniego, L.: Towards seamless large-domain parameter estimation for hydrologic models, Water Resour. Res., 53(9), 8020-8040, 10.1002/2017WR020401, 2017

Mohan, S., and Bhaskaran, P.K.: Evaluation of CMIP5 climate model projections for surface wind speed over the Indian Ocean region, Clim. Dynam., 53(9-10), 5415-5435, 10.1007/s00382-019-04874-2, 2019

Pagliero, L., Bouraoui, F., Diels, J., Willems, P., and Mcintyre, N.: Investigating regionalization techniques for large-scale hydrological modelling, J. Hydrol., 570, 220-235, 10.1016/j.jhydrol.2018.12.071, 2019

Parajka, J., Viglione, A., Rogger, M., Salinas, J.L., Sivapalan, M., and Blöschl, G.: Comparative assessment of predictions in ungauged basins – Part 1: Runoff-hydrograph studies, Hydrol. Earth Syst. Sc., 17(5), 1783-1795, 10.5194/hess-17-1783-2013, 2013

Sutanudjaja, E.H., van Beek, R., Wanders, N., Wada, Y., Bosmans, J.H.C., Drost, N., van der Ent, R.J., de Graaf, I.E.M., Hoch, J.M., de Jong, K., Karssenberg, D., López López, P., Peßenteiner, S., Schmitz, O., Straatsma, M.W., Vannametee, E., Wisser, D., and Bierkens, M.F.P.: PCR-GLOBWB 2: a 5 arcmin global hydrological and water resources model, Geosci Model Dev, 11(6), 2429-2453, 10.5194/gmd-11-2429-2018, 2018

Taylor, K.E.: Summarizing multiple aspects of model performance in a single diagram, Journal of Geophysical Research: Atmospheres, 106(D7), 7183-7192, https://doi.org/10.1029/2000JD900719, 2001

Thompson, S.E., Harman, C.J., Konings, A.G., Sivapalan, M., Neal, A., and Troch, P.A.: Comparative hydrology across AmeriFlux sites: The variable roles of climate, vegetation, and groundwater, Water Resour. Res., 47(10), 10.1029/2010WR009797, 2011

Zhang, Y., Chiew, F.H.S., Zhang, L., and Li, H.: Use of Remotely Sensed Actual Evapotranspiration to Improve Rainfall‐Runoff Modeling in Southeast Australia, J. Hydrometeorol., 10(4), 969-980, 10.1175/2009JHM1061.1, 2009

Zhou, Y., Zhang, Y., Vaze, J., Lane, P., and Xu, S.: Improving runoff estimates using remote sensing vegetation data for bushfire impacted catchments, Agr. Forest Meteorol., 182-183, 332-341, 10.1016/j.agrformet.2013.04.018, 2013