

Response to Comments from Anonymous Reviewer #1

We thank reviewer #1 for the comments. The reviewer's comments are in *italic* below. The added/modified parts are highlighted in **blue**, both below and in revised manuscript. The revised marked-up manuscript (text and figures) is attached after the responses to reviewer #2's comments.

Reviewer #1: Interactive comment on "Assimilation of river discharge in a land surface model to improve estimates of the continental water cycles" by Fuxing Wang et al.

The paper describes a number of experiments assimilating GRDC runoff data into the ORCHIDEE land surface model across the Iberian peninsula. The assimilation adjusts the simulated runoff at sub-catchment scale where GRDC observations are available through a 'optimization parameter', effectively rescaling the simulated runoff towards the observations. The discharge bias is substantially reduced by adjusting this 'optimization parameter' and neighboring sub-catchments are corrected by extrapolating the parameter to these.

The paper is clear and it is well written. The study is likely to be very relevant for future studies possibly extending and improving on the presented concept.

I propose a minor revision.

Some General Remarks:

1. The validation is performed on the basis of using GRDC both as an 'observation' and an independent 'validation' dataset? This should be discussed very critically. I am not an expert in the field of continental runoff and possibly there is no other independent data source to have a better independent validation. In general this is however quite uncommon in assimilation studies, i.e. satellite observations might be assimilated to improve soil moisture and the results would be validated against independent in-situ measurements.

Answer: Yes, the GRDC is used for both river discharge assimilation and validation due to the fact that this is the only comprehensive dataset at global scale so far. To overcome this limitation, the assimilated river discharges are also validated over the catchments where the GRDC stations are discarded during assimilation (e.g., Alcala Del Rio station of the Guadalquivir River in Fig. 7). Although the validation datasets are from the same GRDC source, they are from other independent observation stations thus can be seen as an independent validation. This is a 'first order validation' since the stations are discarded from the assimilation, but we expect the first order is valid.

The explanations are added in **Lines 402-406**: ‘A common validation approach is to compare the assimilated river discharge with other independent data sources. However, the river discharge observations are limited, and the GRDC is the only comprehensive river discharge datasets at global scale so far. To overcome this limitation, the assimilated river discharges are also validated over the catchments where the GRDC stations are discarded during assimilation.’

And in **Lines 416-420**: ‘The discharges for certain sub-basins without assimilated observations (e.g., observation unavailable or GRDC stations discarded) are corrected by x as well. Although the validation datasets are from the same GRDC source, they are from other independent observation stations thus can be seen as an independent validation (‘first order validation’).

2. (1) *For probably this reason the authors compare the corrected evaporation against GLEAM. As stated, GLEAM uses a different precipitation, the entire comparison therefore is challenging. Did the authors consider using the same precipitation as input for their experiments? It should be quite simple.*

Answer: The experiments of using the GLEAM’s precipitation as input are not carried out due to three reasons. First, parts of the atmospheric forcing (e.g., air pressure, air humidity, wind speed, etc.) of GLEAM are not available, and they also impact evapotranspiration estimation. Second, it is not feasible to use GLEAM precipitation and maintaining a coherence with other forcing (e.g. radiation) taken from other sources. Third, the assimilation system is run with three different atmospheric forcing (WFDEI_GPCC, WFDEI_CRU, CRU_NCEP, see Fig. 12a-12d), and the evapotranspiration corrections are quite close (see Fig. 12a-12d). The explanations were added in **Lines 548-551**: ‘Due to the unavailability of parts of GLEAM’s atmospheric forcing (e.g., air pressure, air humidity, air speed, etc.) and difficulty of maintaining a coherence with other forcing, the assimilation system does not run with GLEAM’s precipitation input’.

2. (2) *Also, corrected evapotranspiration values could be compared to Fluxnet in-situ measurements. This should be either included or a strong case should be made why this was not done. The motivation of exactly / only using GLEAM should also be well presented. There are a number of alternative evapotranspiration products.*

Answer: The corrected evapotranspiration (E) is compared with several other evapotranspiration products (Fig. R1 below). The ‘FluxNex extrapolated’ dataset (1982-2008, at 0.5°) was obtained by applying machine learning approach to the upscaling of observations from the global network of eddy covariance towers FLUXNET (Jung et al., 2009). The precipitation of MPIBGC for model tree training comes from CRU. Vinukollu et al. (2011) generated global evaporation from multi-sensor remote sensing data using Penman-Monteith and Priestley-Taylor based approach (‘PUPM’ and ‘PUPT’ hereafter). This data is available at 0.5° from 1984 to 2007 with its precipitation forcing obtained from GPCP. Mueller et al. (2013) collected most of the existing evaporation products and produced

LandfluxEVAL from 1989 to 2005 at 1°. The ‘Diagnostic’ and ‘All’ categories are chosen to use as much observations as possible. These comparison results are consistent with the comparison with GLEAM.

The availability of both precipitation and E for GLEAM data allows to estimate $P-E$ which can be compared with the assimilate values, while the $P-E$ is not available for other datasets. Therefore, only the results of GLEAM are shown in the study.

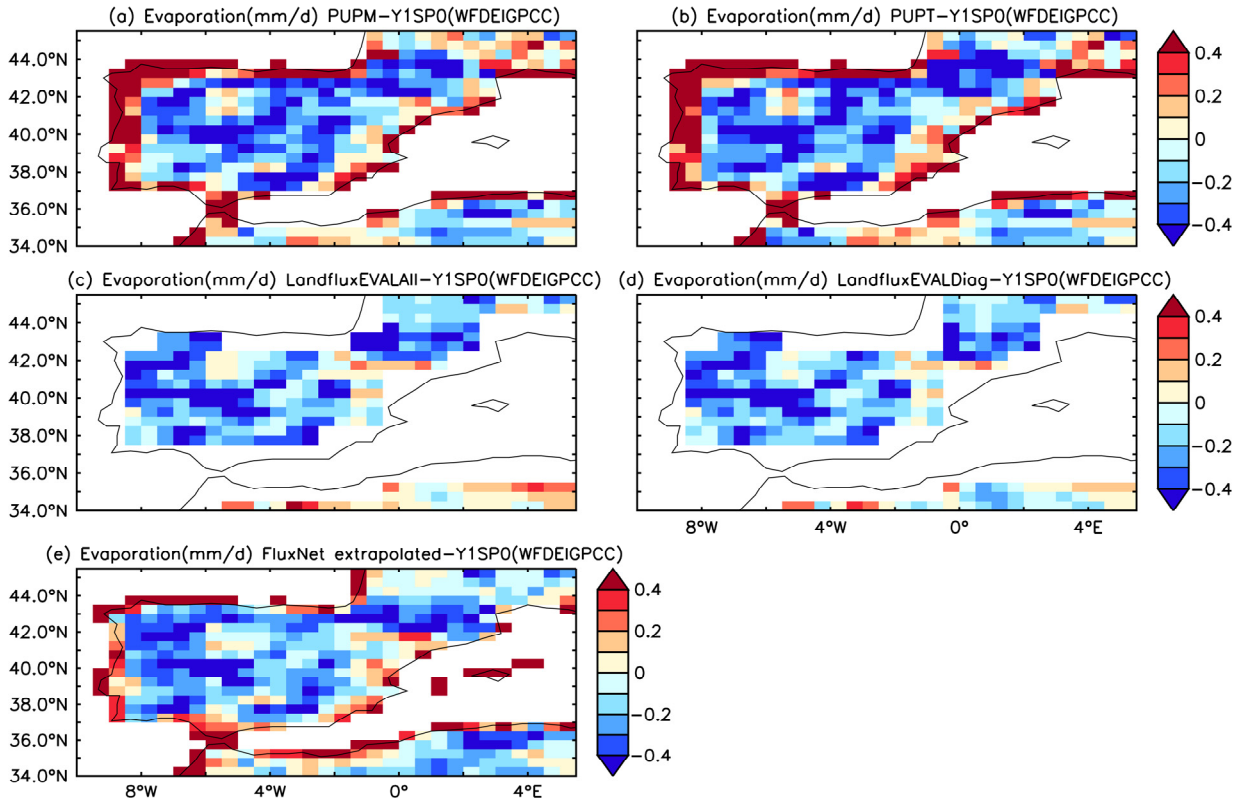


Figure R1. Comparison of assimilated E (mm/d) forced by WFDEI_GPCC with other products: PUPM (a), PUPT (b), LandfluxEVALAll (c), LandfluxEVALDiag (d), and ‘FluxNet extrapolated’ (e) by using multi-year averaged values from 1980 to 1989.

The explanations are added (**Lines 563-567**): ‘The results are quite consistent when comparing the corrected E with several other products which are obtained by using different methodology and forcing (e.g., Jung et al., 2009; Vinukollu et al., 2011; Mueller et al., 2013). Considering the availability of $P-E$ for GLEAM data which allows to compare it with the bias corrected value, only the results of GLEAM are shown’. Three new references were added (**Lines 701-703, 752-756, 838-840**):

Jung, M., Reichstein, M., and Bondeau, A.: Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model, *Biogeosciences*, 6, 2001–2013, doi:10.5194/bg-6-2001-2009, 2009.

Mueller, B., Hirschi, M., Jimenez, C., Ciais, P., Dirmeyer, P. A., Dolman, A. J., Fisher, J. B., Jung, M., Ludwig, F., Maignan, F., Miralles, D., McCabe, M. F., Reichstein, M., Sheffield, J., Wang, K. C., Wood, E. F., Zhang, Y., and Seneviratne, S. I.: Benchmark products for land evapotranspiration: LandFlux-EVAL multi-dataset synthesis, *Hydrol. Earth Syst. Sci.*, 17, 3707-3720, doi:10.5194/hess-17-3707-2013, 2013.

Vinulcollu, R.K., Wood, E.F., Ferguson, C.R., Fisher, J.B.: Global estimates of evapotranspiration for climate studies using multi-sensor remote sensing data: Evaluation of three process-based approaches, *Remote sensing of environment*, 115(3):801-23, 2011.

3. (1) *The correction factor x is applied to each sub-catchment for runoff. It was not quite clear to me how the evapotranspiration was then corrected, presumably at a grid cell level? This duality between correcting at catchment scale but the model essentially being a distributed one computing the water balance at each grid cell should be made clearer. The model runoff is corrected as it was a lumped conceptual land surface model but the relationship between this and the land surface heterogeneity is not clear to me.*

Answer: The evapotranspiration and the ‘runoff + drainage’ are corrected at the same grid cell level (dashed lines in Fig. 1a). Over each grid cell, the ‘runoff + drainage’ is corrected by x , while the corrected evapotranspiration is obtained from Eq. (5). Over each sub-catchment (with several model grids) where the GRDC is available, the correction factors over these model grid cells are the same. These explanations have been added at **Lines 158-159**: ‘Over each upstream area (dashed box in Fig. 1a), the optimal x of these model grid cells are the same. The ‘ $R + D$ ’ and E are corrected at the same grid cell level by x and Eq. (5), respectively.’

3. (2) *Also, is equifinality a serious issue? I suppose a number of optimized x can result in the same or very similar runoff downstream? Can this be mitigated by also looking at the correct seasonality of the generated runoff?*

Answer: Yes, the reviewer is right. A number of optimized x can result in very similar river discharge. Due to the high requirement of computing resources, the assimilation time step is yearly instead of monthly. Using monthly data with monthly x may not completely alleviate the equifinality issue. By using annual x values with monthly discharge, we assume that the seasonal variations of surface runoff and deep drainage are perfect.

Explanations were added in **Lines 623-627**: ‘To improve the calculation efficiency, this study uses annual mean correction factors without considering its seasonal variation thus the seasonal discharges do not improved. One issue of the x optimization approach could be the equifinality with a number of optimized x result in the similar river discharge at downstream.). Future developments can be made

towards generating ensemble optimal x to better assess the uncertainties associated to each parameter x .’

4. *The proposed method is supposed to be superior to more simple water-balance methods? Can this be somehow quantified?*

Answer: Yes, the proposed method is supposed to be superior to simple water-balance methods, because a LSM has better estimation of evapotranspiration (E) using physically based equations and takes advantage of spatial distribution of precipitation (P) and $P-E$. Furthermore, the LSM simulates river discharge at a higher frequency (daily) than the simple water-balance methods (monthly to annual). Explanations were added in **Lines 614-617**: ‘The proposed method is supposed to be superior to the simple water-balance methods, because a LSM estimates E at sub-diurnal scales with physically based equations and takes advantage of spatial distribution of the P and $P-E$ ’.

This is difficult to quantify because both the distributed LSM and the simple water-balance model can be adjusted with data assimilation, but we hope that the correction factor will be smaller for LSM than other water balance methods.

5. *Despite the in general high-level language there are a number of inaccuracies (for instance missing articles).*

Answer: The grammar was checked and the articles were added in several places (e.g., added ‘the’ before ‘Iberian Peninsula’, before ‘Mediterranean’, before ‘Jucar River’, before ‘Chelif’, before ‘ESA’, etc.).

Specific comments:

1. L155: . . . *for different parameters . . . , parameters includes also variables, such as soil moisture, runoff etc.?*

Answer: The sentence ‘for various parameters’ was revised to “for various variables” (**Line 166**).

2. L172: *Again, I’m getting confused with parameter and variable, I suppose parameter is x , but the actual runoff is a variable? Please take care with this throughout the text.*

Answer: Revised (**Lines 163**). The ‘parameter’ now refers to x , while others are ‘variable’.

3. L173: *The background error B is vital in DA, why was it chosen like this? More detail needed.*

Answer: Details were added at **Lines 185-187**: ‘The matrix **B** was determined based on expert knowledge of ORCHIDEE model (Kuppel et al., 2012; Santaren et al., 2014)’. The related references were provided (**Lines 716-718, 799-802**):

Kuppel, S., Peylin, P., Chevallier, F., Bacour, C., Maignan, F., and Richardson, A. D.: Constraining a global ecosystem model with multi-site eddy-covariance data, *Biogeosciences*, 9, 3757-3776, <https://doi.org/10.5194/bg-9-3757-2012>, 2012.

Santaren, D., Peylin, P., Bacour, C., Ciais, P., and Longdoz, B.: Ecosystem model optimization using in situ flux observations: benefit of Monte Carlo versus variational schemes and analyses of the year-to-year model performances, *Biogeosciences*, 11, 7137-7158, doi:10.5194/bg-11-7137-2014, 2014.

4. L218: *Is WFDEI not being updated? Please recheck.*

Answer: Yes, the WFDEI is updated with time (revised at **Line 248**).

5. L237: *Does each HTU have it's own location within a grid cell? Or is it more 'conceptual'. Might be helpful to clarify this in the model description. I'm assuming that they have a fixed location within each grid cell.*

Answer: The location of each HTU is not fixed within each grid cell. Each river basin is constructed by connecting a number of HTUs which are defined inside the ORCHIDEE grid boxes. Each HTU represents the section of the river basin within the grid box. Therefore, the location of each HTU within the grid cell is not fixed but depends on basin characters.

The explanations were added at **Lines 208-211**: ‘The HTU is constructed based on the Pfafstetter topological coding system and user defined size. Each HTU represents the section of the river basin within the grid box, and many HTUs forms a river basin (Nguyen-Quang et al., 2018). Therefore, the relative locations of HTUs in each grid cell are not fixed’.

More details of HTU are elaborated by Nguyen-Quang et al. (2018, **Lines 766-769**):

Nguyen-Quang T., Polcher, J., Ducharne, A., Arsouze. T., Zhou X., Schneider A., and Fita L.: ORCHIDEE-ROUTING: A new river routing scheme using a high resolution hydrological database. Submitted to *Geosci. Model Dev.*, 2018

6. L266: *What is meant by one optimization parameter? In my understanding the algorithm only perturbs x to find the optimum fit between the runoff simulations and observations? The river routing parameters are perturbed? Or does it depend on the number of upstream catchments with a separate x?. Not quite clear to me.*

Answer: It means to perturb x over one upstream catchment in each iteration, but not perturbing the river routing parameters. Different upstream catchments have different x ($x_1, x_2, x_3, \dots, x_N$). The sentence was revised to (**Lines 299-301**): ‘The ORCHIDAS with L-BFGS-B algorithm explores the full space of x by perturbing a separate x (x_i) over the i th upstream catchment ($i=1, 2, \dots, N_{opt}$; N_{opt} is the total number of optimized x depending on the number of observation stations) in each iteration’.

7. L274: . . . value ‘1’ and a ‘pre-estimated error’: ‘and’ should be ‘or’?

Answer: Revised (**Line 310**).

8. L282: the cost function is lower? The value of the cost function? Section Experiments design needs to be a bit clearer.

Answer: ‘The cost function is lower ...’ was revised to ‘The value of the cost function ...’ (**Line 318**).

The following modifications were done in section ‘Experiments design’ to make it clear.

At **Lines 301-304**: ‘To save computing time, the river routing parameterization (forced by corrected R and D) rather than the full ORCHIDEE is executed. The total execution time depends on the number of parameters to be optimized, the length of simulation years, and the number of iterations’.

At **Lines 307-308**: ‘Over the Iberian Peninsula, the range of x is defined between 0 and 20 which depends on the Q_{fg} and Q_{obs} .’

At **Section 2.4** and **Fig. 3**, the expressions of ‘ x_{prior} is set to a constant value 1’ and ‘ x_{prior} is set to pre-estimated-prior’ are names as x_{prior_1} and x_{prior_ref} , respectively.

At **Lines 321-324**: ‘The oscillation of J at the steps 3 and 5 could be due to the fact that the calculation of the gradient of J by finite difference is not optimal. It is also possible because the L-BFGS-B explores partly the physical range during the first few iteration to estimate the Hessian of the cost function for convergence’.

At **Lines 351-355**, the Eq. (10) for ‘Uncertainty’ and its explanations are moved to **Section 2.4**.

At **Lines 270-271**, explanations added: ‘For each GRDC station, the corresponding catchment surface in the model is estimated.’

9. L288: factor m corresponds to number of GRDC stations?

Answer: Yes. Explanations were added (**Line 327**): ‘(i.e., the number of GRDC station)’.

10. L304: *The river routing model runs at each grid cell? The distributed nature of the river routing model is not quite clear.*

Answer: No, the river routing model runs over all model grid cells together in each run. More explanations for Y1SP0 were given at **Lines 344-348**: ‘Take the Y1SP0 for example, in each iteration, the correction factor x is perturbed by m times. For each perturbation, the ORCHIDEE river routing model runs once with one x (e.g., x_i at the i th sub-catchment) being perturbed while the x of other sub-catchments are kept the same. Therefore, the total number of years required for m stations, n iterations and k years assimilation is $m \times n \times k$ ’.

11. L322: *higher than a factor of 1.5?*

Answer: Revised (**Line 373**).

12. L359: *“Summary” seems misnamed for the amount of text following*

Answer: The sections 3.2.1 and 3.2.2 (Summary) in the previous manuscript have been merged into section 3.2, thus the title ‘Summary’ is removed.

13. L369: *They most certainly do. . .*

Answer: The word ‘could’ was removed from the previous manuscript (**Line 430**).

14. L375: *→ can allow, remove ‘of’*

Answer: Removed (**Line 436**).

15. L377: *→ patterns, some inaccuracies in this area*

Answer: Revised (**Line 438**).

16. L383: *Is it also connected to topographic or other land surface features which might be not well presented by the forcing data or the model itself? Just wondering.*

Answer: Yes, it is. Revised in **Lines 442-444**: ‘Besides the atmospheric forcing, the uncertainties could also origin from boundary condition (e.g., topographic or other land surface features), model parameter, model structure or missing processes’.

17. L475: *GREAM → GLEAM*

Answer: Revised (**Line 561**).

18. L479: references, also maybe mention more global attempts to create gridded runoff data? (can be in the introduction).

Answer: Several references were cited at **Lines 570-572**: ‘e.g., Boukthir and Barnier, 2000; Mariotti et al., 2002; Struglia et al., 2004; Peucker-Ehrenbrink, 2009; Ludwig et al., 2009; Szczypta et al., 2012’.

A brief review of global gridded runoff data was added in ‘Introduction’ (**Lines 56-59**): ‘Although great efforts have been made for gridded river discharge data at global scale (e.g., RivDIS v1.1, Vorosmarty et al., 1998; Dai and Trenberth, 2002; Fekete et al., 2002), these data are usually at monthly or annual scales and have not been updated with time’. Two new references were added (**Lines 685-687, 841-843**):

Fekete, B. M., C. J. Vorosmarty, W. Grabs.: High-resolution fields of global runoff combining observed river discharge and simulated water balances, *Global Biogeochemical Cycles*, 16 (3): 15-1 to 15-10, 2002.

Vorosmarty, C. J., Fekete B. M., and Tucker B. A.: Global River Discharge, 1807-1991, V. 1.1 (RivDIS). ORNL DAAC, Oak Ridge, Tennessee, USA. <https://doi.org/10.3334/ORNLDAAC/199>, 1998.

19. L507: Throughout the paper most errors are attributed to the lack of human influences. For sure other factors also play a large role?

Answer: Yes. The role of other factors were mentioned in **Lines 602-603**: ‘The correction factor x can also cover errors in the model structure, model parameter, or boundary conditions (e.g., land surface characteristics imposed to the model)’.

20. Figure3 top: With the logarithmic scaling the lines mostly seem pretty horizontal. Is there a clearly visible gradient when using a different scale? Maybe add this as a window. Missing unit for J ?

Answer: A window for iterations 6-15 with normal y-axis and the unit for J (unit: 1) were added on **Fig. 3a**. Its caption was revised to: ‘The variation of cost function J (unit: 1; logarithmic y-axis) with iterations for x_{prior_1} ($x_{prior} = 1$, in blue) and for x_{prior_ref} ($x_{prior} = \text{pre-estimated-prior}$, in red). The iterations 6-15 are enlarged in the window (normal y-axis).’

Response to Comments from Anonymous Reviewer #2

We thank reviewer #2 for the comments. The reviewer's comments are in *italic* below. The added/modified parts are highlighted in blue, both below and in revised manuscript. The revised marked-up manuscript (text and figures) is attached after the responses to reviewer #2's comments.

Reviewer #2: Interactive comment on “Assimilation of river discharge in a land surface model to improve estimates of the continental water cycles” by Fuxing Wang et al.

The manuscript presents a calibration methodology to optimize a multiplicative factor on modeled surface runoff and deep drainage using river discharge observations. The study focus over Iberia using the ORCHIDEE land surface model, incorporating a river routing scheme and benefiting from the ORCHIDEE data assimilation system. This study is of general interest for the land surface and large-scale hydrological communities presenting a novel optimization/calibration methodology. The manuscript is well presented and organized, but there are a few points that require further attention before publication.

Comments:

1. “Data assimilation”: Data assimilation is normally associated with an “update” of the model state, e.g., via improved initial condition. In this study, merging modelled river discharge with observations is used to “obtain optimized discharge over the entire basin” (as mentioned in the abstract). Therefore I felt that the term “data assimilation” could be a bit misleading for the audience, since this manuscript shows a model optimization or calibration. I suggest that the authors make this point very clear to avoid confusion.

Answer: The data assimilation could be applied for different cases: (1) to correct initial condition (correcting state variable) which is mostly used for numerical weather prediction; (2) to correct the state variable during the data assimilation period (i.e., in this case both the trajectory of the model and the initial conditions are corrected); (3) to correct the parameter of a model by optimization. These different usages can be mixed. In the current study, the data assimilation refers to the 3rd case which is mainly used in ORCHIDEE data assimilation.

We find similar descriptions of data assimilation in several papers. For example, Reichle (2008) mentioned that ‘All data assimilation methods share the basic tenet of merging models and observations, yet the sophistication of the merging algorithm varies widely. Important differences also remain between the specific methods that are most suitable for a given application. Since atmospheric and oceanic dynamics are chaotic (that is, small errors in the initial condition can lead to large

differences at later times in the model integration), data assimilation in these areas is very much concerned with the estimation of initial conditions. By contrast, land surface dynamics are damped, and land surface assimilation is all about estimating errors in uncertain meteorological forcing (boundary) conditions and model parameterizations. Clearly, “one size does not fit all” in data assimilation’. Smith et al. (2013) explained that ‘It is most commonly used to produce initial conditions for state estimation: estimating model variables whilst keeping the model parameters fixed. However, it is also possible to use data assimilation to provide estimates of uncertain model parameters.’ Raoult et al. (2016) also wrote that ‘Optimisation techniques come under the umbrella of model–data fusion and range from simple ad hoc parameter tuning to rigorous data assimilation frameworks. These approaches have been used in a number of studies, covering various LSMs, to derive vectors of parameters that improve model–data fit significantly.’

For this reason, the expression of ‘data assimilation’ is kept in the paper, but a clarification is given at **Lines 91-97**: The data assimilation, a specific type of inverse problem, is generally applied for different cases: (1) to correct initial condition (correcting state variable) which is mostly used for numerical weather prediction; (2) to correct the state variable during the data assimilation period (i.e., in this case both the trajectory of the model and the initial conditions are corrected); and (3) to correct the parameter of a model by optimization. In the current study, the data assimilation refers to parameter optimization (the 3rd case).

References:

- Raoult, N. M., Jupp, T. E., Cox, P. M., and Luke, C. M.: Land-surface parameter optimisation using data assimilation techniques: the adJULES system V1.0, *Geosci. Model Dev.*, 9, 2833-2852, <https://doi.org/10.5194/gmd-9-2833-2016>, 2016.
- Reichle, R. H.: Data assimilation methods in the Earth sciences, *Adv. Water Resour.*, 31, 1411–1418, [doi:10.1016/j.advwatres.2008.01.001](https://doi.org/10.1016/j.advwatres.2008.01.001), 2008.
- Smith PJ, Thornhill GD, Dance SL, Lawless AS, Mason DC, Nichols NK. 2013. Data assimilation for state and parameter estimation: application to morphodynamic modelling. *Q. J. R. Meteorol. Soc.* 139: 314–327. [doi:10.1002/qj.1944](https://doi.org/10.1002/qj.1944)

2. River routing model: Since both references of the routing model are not published yet (Nguyen-Quang et al., 2017; Zhou et al., 2017) and this is a key component of this study it is important to have a bit more details on how the three linear reservoir are represented and which model parameters are used and were defined (e.g., water residence time). For example the aquifer level is referred later in the text due to spin-up, but it is not clear from the model description how the aquifers are represented in the model.

Answer: More descriptions for the three linear reservoir and the water residence time are added at **Lines 212-219**: ‘In each HTU, the water is routed through a cascade of three linear reservoirs characterized by their residence times: the groundwater, overland and stream reservoirs. The runoff and drainage are the inputs into the overland reservoir and groundwater reservoir, then they flowed into the stream reservoir of the downstream sub-grid basin. The residence times are determined by multiplying a constant reservoir factor (g) with a slope index (k). The g for stream, overland and groundwater reservoirs are 0.24, 3, and 25 day/km, respectively (Ngo-Duc et al., 2007). The slope index is a function of distance (d) and slope (S) between a pixel and its downstream pixel ($k=d/S^{1/2}$ defined by Ducharne et al., 2003).’

At **Lines 221-224**: ‘The river discharge is linear with R and D at annual scale over a small basin. In case of more than one observation stations are assimilated in a river basin (e.g., x_1 and x_2 in Fig. 1a), the river discharge at downstream is affected by the discharge of upstream thus it is not a linear system anymore. Therefore, the optimization is needed to deal with the x over the non-linear sub-basins’

At **Lines 135-136**: ‘The W and A terms refer to water storage and water stored in the aquifers, respectively.’

The description of aquifers are added at **Lines 196-198**: ‘In other words, the ORCHIDEE LSM assumes that the aquifer level is below the model bottom, and it neglects the upward water flow through capillary forces from its underlying aquifer.’

The two papers are available on line now at: <https://www.geosci-model-dev-discuss.net/gmd-2018-57/> and <https://www.hydrol-earth-syst-sci-discuss.net/hess-2018-88/> . The citation of the two papers has been updated at **Line 204** (Nguyen-Quang et al., 2018; Zhou et al., 2018). The reference lists are adjusted at **Lines 765-768** and **Lines 853-856**.

Nguyen-Quang, T., Polcher, J., Ducharne, A., Arsouze, T., Zhou, X., Schneider, A., and Fita, L.: ORCHIDEE-ROUTING: A new river routing scheme using a high resolution hydrological database, Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2018-57>, in review, 2018.

Zhou, X., Polcher, J., Yang, T., Hirabayashi, Y., and Nguyen-Quang, T.: Understanding the water cycle over the upper Tarim basin: retrospect the estimated discharge bias to atmospheric variables and model structure, Hydrol. Earth Syst. Sci. Discuss., <https://doi.org/10.5194/hess-2018-88>, in review, 2018.

3. How does the simple estimate of the correction factor used as prior (“ x_{prior} ”) compares with optimized values in figure 6? Are the changes significant for example in terms of improved correlation?

Answer: For easier understanding, the methods of ‘ $x_{prior} = 1$ ’ and ‘ $x_{prior} =$ pre-estimated-prior’ are named as x_{prior_1} and x_{prior_ref} , respectively (**Section 2.4 and Fig. 3**).

The x_{prior_ref} is compared with optimized correction factor in Fig. R1 below. The x_{prior_ref} captures the general distribution pattern of optimal x , but the correlation coefficient of using x_{prior_ref} is lower than that of using optimal x . In other words, the assimilated river discharge is improved through both choosing x_{prior_ref} and optimization. The role of optimization is to find an appropriate correction factor when there are several basins (with observations) overlaps at upstream.

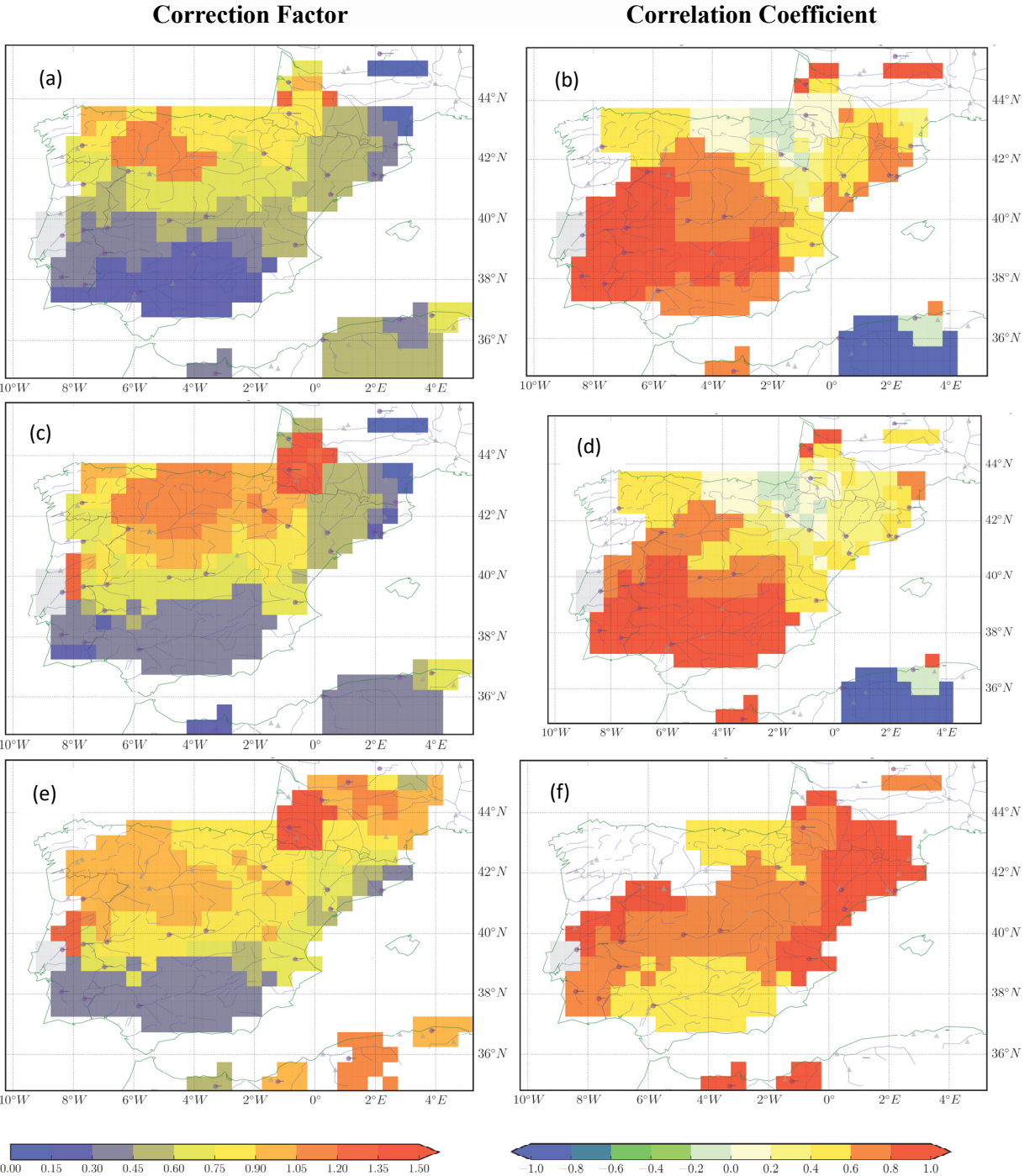


Figure R1. The x_{prior_ref} (left) and the correlation coefficient (right) of river discharge between observations and simulations from 1980 to 1989 for WFDEI_GPCC (1st row), WFDEI_CRU (2nd row) and CRU_NCEP (3rd row) forcing.

Explanations were added in **Lines 397-401**: ‘It should be mentioned that the x_{prior_ref} is able to capture the general distribution pattern of optimal x , but the performance of river discharge estimation is significantly improved through optimization. The role of optimization is to find an appropriate correction factor when there are several basins (with observations) overlaps at upstream’.

4. Role of forcing: To discard the role of precipitation forcing, the three datasets could be compared with a high resolution precipitation dataset (IB02, Belo-Pereira et al. 2011) also in terms of mean ratios: GPCC/IB02 CRU/IB02 NCEP/IB02 and compared with the “ x ” correction factor. I don’t see this as mandatory for the paper’s publication, but would make the results more robust.

Belo-Pereira M, Dutra E, Viterbo P. Evaluation of global precipitation data sets over the Iberian Peninsula. *Journal of Geophysical Research-Earth Surface*. 2011. 116: D20101. doi:10.1029/2010jd015481.

Answer: The precipitation of WFDEI_GPCC, WFDEI_CRU, and CRU_NCEP is compared with the IB02 precipitation data. The precipitation of the three forcing are higher than IB02 over most regions (Figs. R2a-R2c) but their spatial distributions are different with the proposed evaporation correction (Figs. 9e-9g). The ratios of WFDEI_GPCC/IB02, WFDEI_CRU/IB02 CRU_NCEP/IB02 are generally higher than 1 with few grid cells of ratios lower than 1 being distributed randomly (Figs. R2d-R2f). The pattern of the three ratios is not consistent with the optimized correction factor (Figs. 8a-8c), which indicates that the precipitation forcing error is not likely the dominant factor of the correction factor distribution.

These analysis are added in the revised manuscript (**Lines 452-459**): ‘This is also demonstrated by comparing the precipitations between the three forcing and IB02 dataset. Compared to IB02, all the three forcing overestimate rainfall in the Iberian Peninsula (Figs. S1a-S1c), but none of these error patterns resembles that of the proposed E correction (Figs. 9e-9g). Unlike the pattern of the correction factor (Figs. 8a-8c), the ratios of annual mean precipitation between the three forcing and IB02 are higher than 1 over most regions (Figs. S1d-S1f). Therefore, the precipitation forcing error is not likely the dominant factor in determining the correction factor distribution.’

The IB02 dataset is described at **Lines 254-257**: ‘The precipitation of the three forcing is compared with IB02 dataset which is a gridded daily rainfall dataset for Iberia Peninsula with 0.2° resolution covers 1950 to 2003 (Belo-Pereira et al., 2011). It is generated by using ordinary kriging from more than 2400 quality-controlled stations.’

The reference was added at **Lines 650-652**. The Figs. R2d-R2f below were added in the ‘Supplementary’ of the manuscript (Fig. S1).

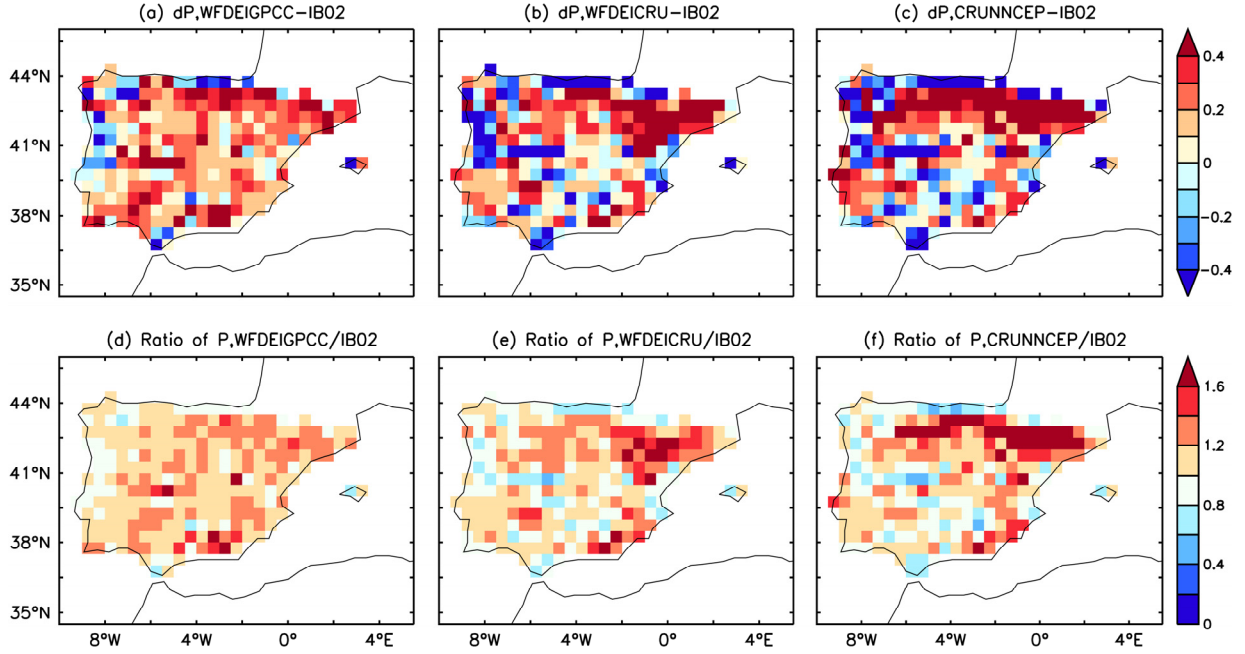


Figure R2. Comparison of precipitation (P , in mm/d) between IB02 and that used in the assimilation (a and d: WFDEI-GPCC; b and e: WFDEI-CRU; c and f: CRUNCEP) averaged from 1980 to 1989: 1st row for difference; 2nd row for ratio.

5. Impact on evaporation: Section 3.4 compares the first guess evaporation by the land-surface model with the changes in evaporation resulting for the correction as a post-processing. Would it be possible to re-run the LSM applying just a constant correction factor to evaporation ? I understand that this might be difficult to do while conserving energy, but even if energy is not conserved, it could show the impact of “improving” evaporation, that would then be reflected directly in R & D and should, in principle improve the discharge simulations.

Answer: We tested the possibility of improving river discharge by using a constant correction factor to evaporation. Theoretically, the modification of evaporation leads to a change in soil moisture thus surface runoff and deep drainage are changed. From Eq. (6), the correction factor for E (X_{Ecorr}) can be derived from x by Eq. (R1). The X_{Ecorr} is then applied to correct E (Eq. R2). Like the correction factor x , the X_{Ecorr} changes with year.

$$X_{Ecorr} \approx \frac{E + (1 - x) \cdot (R + D)}{E}, \quad (R1)$$

$$E_{corr} = X_{Ecorr} \cdot E \quad (R2)$$

The Eqs. R1 and R2 were implemented in ORCHIDEE LSM and the LSM was running over 1980-1984. The Fig. R3 shows the BIAS of river discharge after correcting evaporation in ORCHIDEE LSM. The absolute BIAS is reduced comparing with the reference run (forced by WFDEI_GPCC without correcting evaporation).

The BIAS becomes negative after correcting evaporation, which is probably because the evaporation correction factor X_{Ecorr} is greater than 1 over most cases, and it leads to a decrease in R+D with time evolution. Unlike the correction of runoff and drainage by using a constant factor (change with year) in current study, the correction of evaporation leads to a feedback on soil moisture which in turn affects the evaporation simulation. Therefore, both energy and water balance are not conserved in this case. Another solution of improving river discharge simulation by correcting evaporation could be to run the full ORCHIDEE LSM in the assimilation system with the same cost function as Eq. (7) in the manuscript. In this way, the intermediate variables are adjusted towards optimal river discharge with the modification of evaporation. Because the optimization by running the full ORCHIDEE model is very time consuming, this is not done in this paper but could be one of the future work.

The explanations were added in the revised manuscript (**Line 487-496**): ‘We also tested the possibility of improving the river discharge estimation by using a constant correction factor to evaporation (X_{Ecorr}). The X_{Ecorr} (different for each year) can be derived from Eq. (6).

$$X_{Ecorr} \approx \frac{E + (1 - x) \cdot (R + D)}{E}, \quad (11)$$

$$E_{corr} = X_{Ecorr} \cdot E \quad (12)$$

Although the Eqs. 11-12 are able to improve river discharge estimation by modifying soil moisture, the energy and water balance are not conserved. One solution could be to run the full ORCHIDEE LSM in the assimilation system with the same cost function as Eq. (7). In this way, the intermediate variables are adjusted towards optimal river discharge with the modification of evaporation. This approach executes the full ORCHIDEE model thus is very time consuming and is beyond the scope of the current study.’

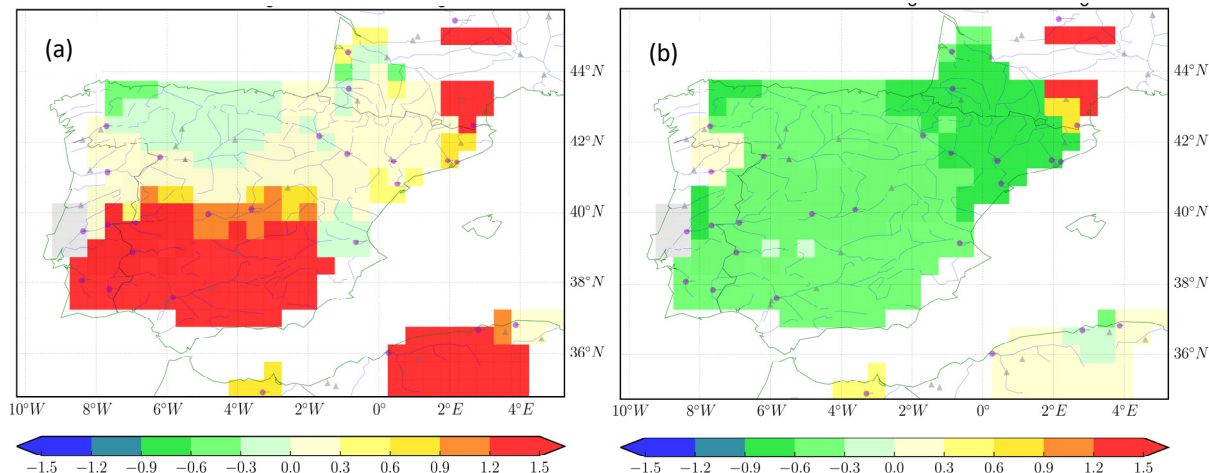


Figure R3. The BIAS of simulated river discharge before (left) and after (right) correcting evaporation in LSM by correction factor from 1980 to 1984.

6. Comparison with GLEAM: It would be beneficial to also present the comparison between the original Evaporation and GLEAM in addition to the results in Fig. 12 (could be an extra panel). Considering the results shown, I find it difficult to understand the sentence “ In 473: “This result further confirms that And some processes are probably missing in GREAM v3.1”. Please expand on this discussion to clarify the basis for this assumption.

Answer: The comparison between the original Evaporation and GLEAM is shown in Fig. S4 below (and **Figs. 12a-12d** in the revised manuscript).

The explanations have been added at **Lines 551-554**: ‘We find large difference between GLEAM and FG, which indicates that the evaporation is quite uncertain for different estimations. The geographical distribution and magnitude of difference in E between GLEAM and FG is highly consistent with that between GLEAM and bias corrected values by using different forcing (Figs. 12a-12c, and 12e-12g).’

The sentences have been revised at **Lines 558-561** to avoid confusion: ‘Because the bias corrected $P-E$ are corrected by GRDC observed river discharge, the $P-E$ (\approx river discharge) of GLEAM is very likely to be higher than GRDC observations over the Iberia. This result indicates that some processes are probably also missing in GLEAM v3.1.’

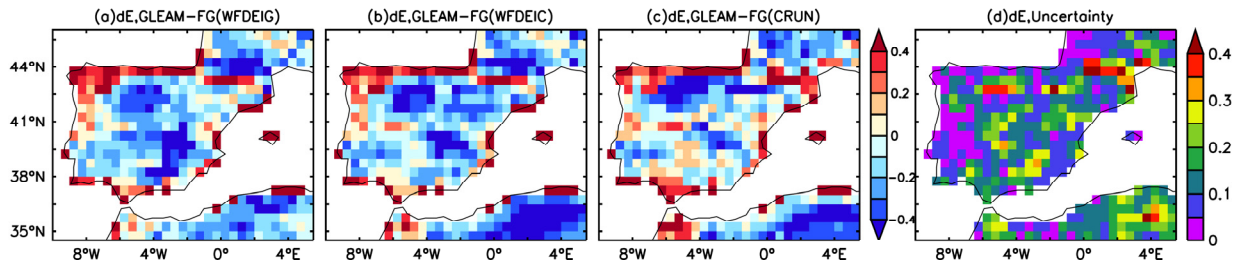


Figure R4. Comparison of evaporation (E , in mm/d) between GLEAM (v3.1) and FG values using different forcing (a: WFDEI-GPCC; b: WFDEI-CRU; c: CRUNCEP; d: uncertainty of using different forcing) averaged from 1980 to 1989.

Details:

1. Ln 21: “earth’s water cycle”

Answer: Revised (**Line 21**).

2. Ln 324: The relative bias shown in figure 5 highlight the biases in the South since the absolute values are low. The absolute biases might be higher in the northern areas.

Answer: Fig. R5 below plots the absolute bias over 1980-1989 by three different forcing. The high values of absolute bias are distributed in both northern and southern areas, and its spatial distribution is different for different forcing. To avoid confusion, the BIAS is named normalized bias (*Norm_BIAS*) in the revised manuscript.

The explanations have been added at **Lines 369-371**: ‘The spatial pattern of the absolute bias in river discharge varies with the atmospheric forcing (not shown). The normalized bias is then applied to measure the river discharge simulation.

The expression was also revised at **Line 374**: ‘The *Norm_BIAS* is small (within +/- 0.3) over north, west and southeast of the region (Figs. 5b, 5d and 5f)’.

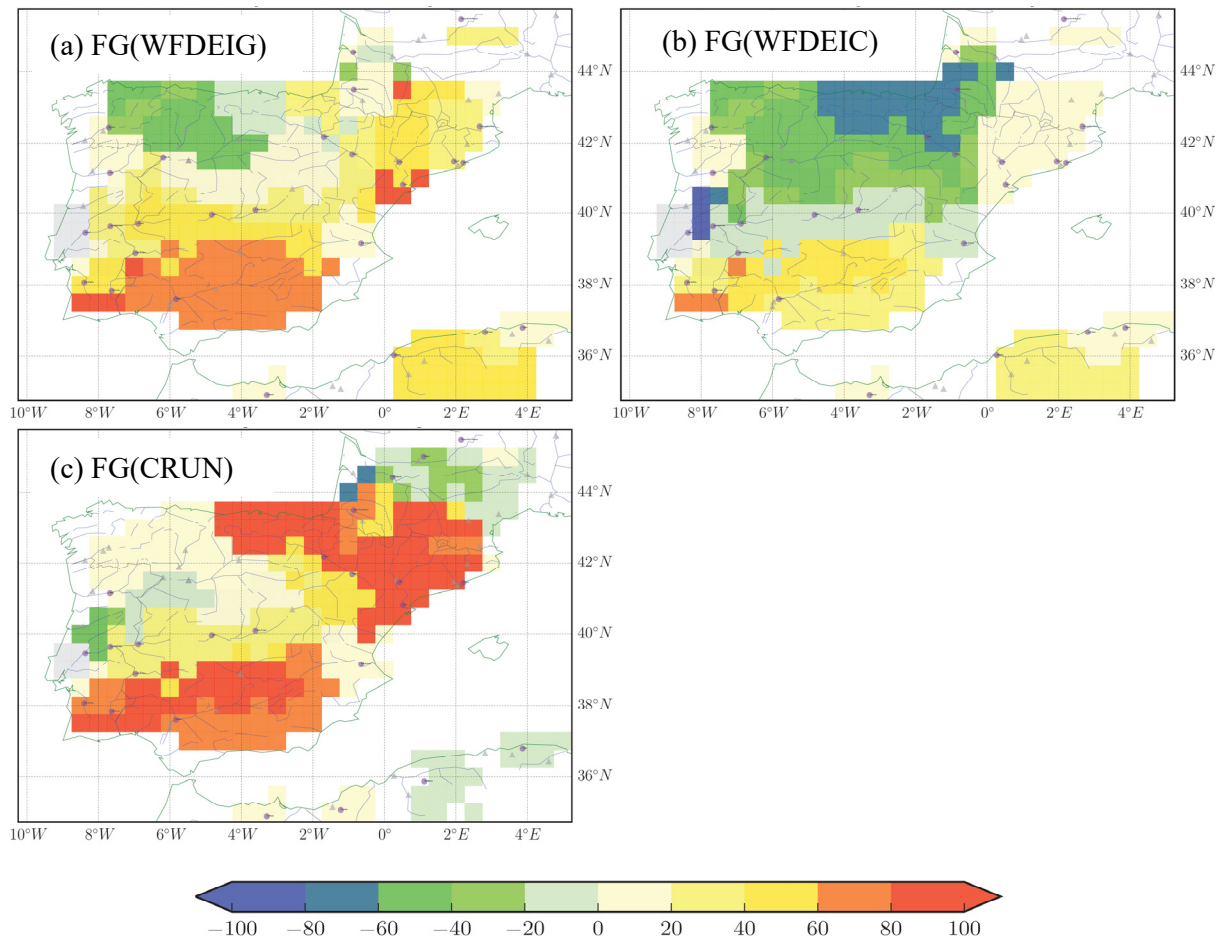


Figure R5. The absolute bias (m^3/s) of river discharge simulations from 1980 to 1989 using WFDEI_GPCP (a), WFDEI_CRU (b) and CRU_NCEP (c) forcing.

3. Ln 351: Should be: “Fig. 7 shows the annual mean” and not “annual cycle”

Answer: Revised to ‘annual mean’ (**Line 407**).

4. Ln 357: Looking that the stations distribution in Figure 2, the station Alcala Del Rio looks very close to Cantillana. If this is the case, the good results in Alcala Del Rio might be just a direct effect of the use of Cantillana observations, and it does not “validate the hypothesis that x is distributed homogeneously over the upstream basin”. Please provide the distance between the stations and difference in upstream area and mean Q_{obs} to show that Alcala Del Rio has other tributaries than just Cantillana to justify this sentence.

Answer: Based on GRDC observations, the distance and the difference in upstream area between Alcala Del Rio and Cantillana stations are 15.3 km and 2124 km^2 (46995 km^2 and 44871 km^2 , respectively). Between the two stations, there are several tributaries flow to Alcala Del Rio station, which leads to different annual mean river discharges at Cantillana (49.7 m^3/y) and Alcala Del Rio

stations (94.8 m³/y). This result illustrates that this approach is able to correct the river discharge over the entire basin. The above numbers and expressions were provided in the revised manuscript at **Lines 409-416**: ‘The observation of this station is not assimilated due to its large upstream area difference (15.53%>10%) between model (55635 km²) and GRDC (46995 km²). The overestimated discharge simulated by the model at this station is also corrected because it benefits from the correction factor estimated at the Cantillana station (-5.83°W, 37.59°N; 44871 km²) which locates at the 15.3 km upstream of Alcala Del Rio station of the Guadalquivir River (southwest of the Iberian Peninsula). Between the two stations, there are several tributaries flow to Alcala Del Rio station, which leads to different annual mean river discharges at Cantillana (49.7 m³/y) and Alcala Del Rio stations (94.8 m³/y). This result illustrates that this approach is able to correct the river discharge over the entire basin.’

5. Ln 429: It is not clear that the simulations “underestimate the inter-annual variability”. Could you provide the standard-deviation of the annual means of the observations and simulations?

Answer: The standard-deviations of the annual means for the FG(WFDEIG) and FG(WFDEIC) are 28.8 m³/s and 25.2 m³/s, respectively. They are lower than observation (33.8 m³/s). The values are provided at **Line 506-509**: “... while the FG(WFDEIG) and FG(WFDEIC) underestimate the inter-annual variability comparing with observations (Fig. 10a-10b). The standard-deviation of the annual means for observation, FG(WFDEIG), FG(WFDEIC) and FG(CRUN) are 33.8 m³/s, 28.8 m³/s, 25.2 m³/s and 34.3 m³/s, respectively.”

6. Ln 436 (results in Fig. 10): If we assume that the increase in discharge is due to an increase of groundwater abstraction should we expect decrease of the correction factor since this is a process which is not represented in the model? The opposite sign with an increase of the correction factor, with higher corrections in 1980 (around 0.2) and lower in 1989 (around 0.6) suggests that the correction factor is correcting for other processes and not human intervention? I think this is worth some discussion.

Answer: The following sentences and references were removed to avoid confusion: “The groundwater usage occupies about 90%, 16% and 44% in upper, middle and lower Guadiana river basin (Aldaya and Llamas, 2008). The groundwater abstraction increases (irrigation intensifies) during this period (Llamas and Garrido, 2007), which causes a reduction in soil water storage capacity and an increase in river discharge (Valverde et al., 2015)”.

Llamas, M. R. and Garrido, A.: Lessons from intensive groundwater use in Spain: Economic and social benefits and conflicts, In: Giordano M, Villholth KG (eds) The agricultural groundwater revolution: Opportunities and threats to development, Chapter 13. CABI International,

Oxfordshire, 266–295, 2007.

Valverde, P., Serralheiro, R., de Carvalho, M., Maia, R., Oliveira, B., and Ramos, V.: Climate change impacts on irrigated agriculture in the Guadiana river basin (Portugal), *Agric Water Manag*, 152:17–30, doi:10.1016/j.agwat.2014.12.012, 2015.

New discussions (including a reference) and a new reference were added at **Lines 511-515** and **Lines 835-837**: ‘Besides, there are many interconnected wetlands and structurally complex hydrogeological boundaries between the two upper-Guadiana aquifer in the upper Guadiana River basin (Van Loon and Van Lanen, 2013). These complex features are difficult to represent in model thus large bias exist in river discharge of ORCHIDEE. The correction factor corrects these model defects (Fig. 10c) and it demonstrates good skill in correcting the inter-annual variability of discharge and runoff coefficient (Fig. 10a-10b).’

Van Loon, A. F. and Van Lanen H. A. J.: Making the distinction between water scarcity and drought using an observation-modeling framework, *Water Resour Res*, 49, doi:10.1002/wrcr.20147, 2013.

1 **Assimilation of river discharge in a land surface model to**
2 **improve estimates of the continental water cycles**

3
4 Fuxing WANG¹, Jan POLCHER¹, Philippe PEYLIN², and Vladislav BASTRIKOV²

5
6 ¹Laboratoire de Météorologie Dynamique, IPSL, CNRS, Ecole Polytechnique, 91128, Palaiseau,
7 France

8 ²Laboratoire des sciences du climat et de l'environnement, IPSL, CEA, Orme des Merisiers,
9 91191, Gif sur Yvette, France

10
11
12 Manuscript revised on June 12, 2018

13 To be submitted to *Hydrology and Earth System Sciences (HESS)*

14
15 *Correspondence to:

16 Fuxing Wang

17 Email: fuxing.wang@lmd.jussieu.fr

18 Tel: 0033 (0)1 69 33 51 80

19

20 **Abstract:**

21 The river discharge plays an important role in earth's [water cycle](#), but it is difficult to
22 estimate due to un-gauged rivers, human activities, and measurement errors. One approach is based
23 on the observed flux and a simple annual water balance model (ignoring human processes) for
24 ungauged rivers, but it only provides annual mean values which is insufficient for oceanic
25 modellings. Another way is by forcing a land surface model (LSM) with atmospheric conditions.
26 It provides daily values but with uncertainties associated to models.

27 We use data assimilation techniques by merging the modelled river discharges by
28 ORCHIDEE (without human processes currently) LSM and the observations from Global Runoff
29 Data Center (GRDC) to obtain optimized discharges over the entire basin. The 'model systematic
30 errors' and 'human impacts' (e.g., dam operation, irrigation, etc.) are taken into account by an
31 optimization parameter x (with annual variation), which is applied to correct model intermediate
32 variables runoff and drainage over each sub-watershed. The method is illustrated over [the](#) Iberian
33 Peninsula with 27 GRDC stations over the period 1979-1989. ORCHIDEE represents a realistic
34 discharge over north of [the](#) Iberian Peninsula with small model systematic errors, while the model
35 overestimates discharges by 30%-150% over south and northeast region where the blue water
36 footprint is large. The [normalized](#) bias has been significantly reduced to less than 30% after
37 assimilation, and the assimilation result is not sensitive to assimilation strategies. This method also
38 corrects the discharge bias for the basins without observations assimilated by extrapolating the
39 correction from adjacent basins. The 'correction' increases the inter-annual variability of river
40 discharge because of the fluctuation of water usage. The E ($P-E$) of GLEAM (Global Land
41 Evaporation Amsterdam Model, v3.1a) is lower (higher) than the bias corrected value, which could
42 be due to the different P forcing and probably the missing processes in the GLEAM model.

43 Key words: river discharge; data assimilation; human processes; water cycle; land surface model;
44 the Mediterranean

45

46 **1. Introduction**

47 The river discharge is an essential component of the earth's water cycles, which can be
48 used as an indicator of the hydrological cycle intensification (Munier et al., 2012). It is important
49 not only for water resources management, climate studies, ecosystem health over land (Syed et al.,
50 2010; Sichangi et al, 2016), but also for providing freshwater inflow to ocean (Dai and Trenberth,
51 2002). The freshwater flux at the sea surface has significant influence on the climate system (e.g.,
52 ENSO, ocean dynamics) and on ocean salinity (Kang et al., 2017). The fresh water inputs for ocean
53 model usually requires high frequency data (e.g., daily or 10-daily, Scherbakov and Malakhova
54 2011). Besides, as the ocean model with high spatial resolution (e.g., < 10 km) demonstrates better
55 skills than coarse resolution model (Bricheno et al., 2014; Wang et al., 2017), there is also a
56 requirement of high resolution fresh water fluxes. [Although great efforts have been made for
57 gridded river discharge data at global scale \(e.g., RivDIS v1.1, Vorosmarty et al., 1998; Dai and
58 Trenberth, 2002; Fekete et al., 2002\), these data are usually at monthly or annual scales and have
59 not been updated with time.](#) Therefore, it is of great interest to estimate large scale river discharge
60 over the long-term at high temporal and spatial resolution and low uncertainty.

61 Estimating the river discharge input to ocean is a difficult endeavor for several reasons.
62 First, there are many un-gauged rivers that are difficult to evaluate. Second, most large rivers are
63 gauged by national agencies, and these data are difficult to access for public users. Besides, the
64 number of operational gauging stations is decreasing worldwide (Syed et al., 2010; Sichangi et al,
65 2016). Third, even though the observations are available, the observed river flow at the outlet is
66 not well known because it is difficult to get gauging stations close to the river mouth and many
67 observations are affected by human activities especially in semi-arid regions (Jordà et al., 2017).

68 One approach to estimate the freshwater inflow into ocean is based on the observed water
69 fluxes over data-rich regions and a simple annual water balance model, precipitation inputs minus
70 the evaporation, which ignoring human usage and other processes over ungauged basins (e.g.,
71 Szczypta et al. 2012; Peucker-Ehrenbrink, 2009; Mariotti et al., 2002; Struglia et al. 2004; Boukthir
72 and Barnier, 2000; Ludwig et al., 2009). This method is the basis of most water balance studies
73 and oceanic modelling activities but it has several limitations. First, there are uncertainties in
74 observations related to measurement method and post-processing method. These uncertainties are

75 difficult to quantify due to the incomplete information (Jordà et al., 2017). Second, only annual
76 mean values are available over un-gauged basins (about 40% for the Mediterranean; 42% over
77 globe excluding Greenland and Antarctica, Clark et al., 2015) by simple runoff models, which are
78 not sufficient for oceanic modellings.

79 Riverine input can also be obtained through forcing a state of the art land surface model
80 (LSM) or global hydrological model (GHM) with bias corrected atmospheric conditions (e.g., aus
81 der Beek et al., 2012; Bouraoui et al. 2010; Jin et al., 2010; Sevault et al., 2014). These numerical
82 models can estimate river discharge at higher frequency and over more un-gauged basins (Jordà et
83 al., 2017), but they are associated with modelling uncertainties. First, models are designed and
84 have proved the ability to capture the natural water cycles, but relatively less progress has been
85 made in parameterizing human processes (Pokhrel et al., 2017). The water flow of many
86 catchments has been strongly regulated by human through irrigation use, dam operation, etc. (e.g.,
87 the southern shores of the Mediterranean). Second, there are large discrepancies among models
88 resulting from the differences in model inputs, parameterizations, and atmospheric forcing data
89 (Ngo-Duc et al., 2007; Wang et al., 2016; Liu et al. 2017).

90 The objective of the present study is to illustrate a novel approach based on assimilation
91 techniques applied to LSM to estimate continental water cycles (riverine fresh water). [The data
92 assimilation, a specific type of inverse problem, is generally applied for different cases: \(1\) to
93 correct initial condition \(correcting state variable\) which is mostly used for numerical weather
94 prediction; \(2\) to correct the state variable during the data assimilation period \(i.e., in this case
95 both the trajectory of the model and the initial conditions are corrected\); and \(3\) to correct the
96 parameter of a model by optimization. In the current study, the data assimilation refers to the 3rd
97 case.](#) This assimilation approach merges the data from the model (ORCHIDEE LSM) and the
98 observed river discharge from the Global Runoff Data Centre (GRDC, 56068 Koblenz, Germany).
99 This will allow to compensate for model systematic errors or missing processes and provide
100 estimates of the riverine input into the sea at high temporal and spatial resolution. Although
101 previous works exist on assimilation of river discharge (e.g., Li et al., 2015; Bauer-Gottwein et al.,
102 2015; Pauwels et al., 2009), these studies mainly focus on the stream flow prediction over
103 individual catchments. They are difficult to extend to long-term scale and large catchment due to
104 the observations and computing time limitations.

105 This paper focuses on the methodology and its illustration in a Mediterranean region (the
106 Iberian Peninsula) which is considered one of the most vulnerable regions to climate change due
107 to its geographic and socio-economic characteristics (Vargas-Amelin and Pindado, 2014).
108 Although the amount of river discharge is relatively small (about one third to half of precipitation
109 amount; Tixeront, 1970; Shaltout and Omstedt 2015), it is an important source of fresh water
110 entering the Mediterranean Sea and it plays an important role in sustaining the marine productivity
111 (Bouraoui et al., 2010) and overturning circulation (Verri et al., 2017). The river discharges to the
112 Mediterranean Sea underwent important changes during recent decades. This variation is
113 particularly important for this region because of its scarce water resource with increasing water
114 demand for domestic, industrial, irrigation and tourism activities, as well as its drier and warmer
115 conditions under climate change (Romanou et al., 2010). Considering the high stress on the water
116 resources in the Mediterranean region, accurate estimation of the actual resources is important.

117 The methods (including the model, datasets and numerical experiment) are described in
118 Sect. 2. The results and discussions are given in Sect. 3. Conclusions are drawn in Sect. 4.

119 **2. Methods**

120 **2.1. The theoretical background**

121 The theoretical basis of the LSM assimilation for the study is the vertical and lateral water
122 balance. The precipitation (P) input of a basin is transferred into either evaporation, surface runoff
123 (R), deep drainage (D) (eventually the R and D reaching the channel and leaving in the form of
124 river discharge), or stored in the ground.

$$125 \quad \frac{dW}{dt} = P - (R + D) - E, \quad (1)$$

126 Over long period, the change of water storage $\frac{dW}{dt}$ is small ($\frac{dW}{dt} \approx 0$), thus

$$127 \quad P - E \approx R + D \quad (2)$$

128 The lateral water balance over a basin (e.g., the sub-catchment 2 in blue in Fig. 1a) is given
129 by:

130
$$\frac{dA_2}{dt} = \left[\int_{S_2} (R_2 + D_2) ds \right] - Q_2 + Q_1, \quad (3)$$

131 where S_2 is the area of sub-catchment 2; A_2 is the water stored in the aquifers of area S_2 ; Q_2 and Q_1
 132 are the river discharge at outlet of each sub-catchment, and they are calculated by the integral of
 133 runoff and drainage over the sub-catchment area S_1 and S_2 . We assume the A_2 variation at annual
 134 scale is small ($\frac{dA_2}{dt} \approx 0$) due to its slow variability, although it can be nonzero due to the human
 135 intervention (e.g., over Indo-Gangetic Basin, MacDonald et al., 2016). [The \$W\$ and \$A\$ terms refer to](#)
 136 [water storage and water stored in the aquifers, respectively](#). The Eqs. (1)-(3) describe the basic
 137 water cycle processes in the LSMs.

138 Despite that the LSMs have developed rapidly during the last few decades, few models
 139 take into account the human water usage processes. Due to this limitation, LSMs are usually
 140 accompanied with errors in reproducing discharge and evaporation in areas where these processes
 141 are dominant. Assuming the P forcing is known in LSM, the modelled water continuity imposes a
 142 balance of errors between E , R and D . However, the R and D are conceptual variables, and their
 143 errors are impossible to evaluate by observations directly. The field measurements of E over large
 144 area are also scarce due to land surface heterogeneity (Kalma et al., 2008). Fortunately, the
 145 observations of river discharge (Q_{obs}) are available. By fitting modelled discharge with Q_{obs} , we
 146 can correct model intermediate variables in Eqs. (1)-(3) (e.g., correct R and D by a correction factor
 147 x , Fig. 1a) in order to get bias corrected river discharge (Q_{corr}).

148
$$Q_{corr} = \int_{catchment} (x \cdot R + x \cdot D) dS, \quad (4)$$

149 Recalling the $\frac{dW}{dt}$ is small and P is known, we then transfer the x into vertical water balance
 150 and close the horizontal water balance by the corrected evaporation (E_{corr}):

151
$$E_{corr} \approx P - x \cdot (R + D), \quad (5)$$

152 The impacts of assimilation on E (ΔE) can be derived from the optimal x , R , and D :

153
$$\Delta E = E_{corr} - E \approx (1 - x) \cdot (R + D), \quad (6)$$

154 The key problem remains to determine the optimal x (described in Sect. 2.2.2). Each
 155 discharge observation station corresponds to an optimal correction factor x since the discharge is
 156 the only representative of the integral over the basin. The total number of x depends on the number
 157 of available stations. The optimal x over each observation station is applied to its entire upstream
 158 area. Over each upstream area (dashed box in Fig. 1a), the optimal x of these model grid cells are
 159 the same. The ‘ $R + D$ ’ and E are corrected at the same grid cell level by x and Eq. (5), respectively.

160 2.2. The models

161 2.2.1. Assimilation strategy and ORCHIDAS

162 The optimal x is obtained from the ORCHIDEE Data Assimilation System (ORCHIDAS,
 163 <https://orchidas.lsce.ipsl.fr/>). It was designed to optimize the variables related to water, energy and
 164 carbon cycles in ORCHIDEE (Organising Carbon and Hydrology in Dynamic Ecosystems;
 165 Krinner et al. 2005; De Rosnay et al., 2002) LSM by using various observations (e.g. in situ,
 166 satellite, etc.). The ORCHIDAS has been applied over different regions for various variables and
 167 demonstrated good performance (Santaren et al., 2007; Kuppel et al., 2012; MacBean et al., 2015).
 168 More details of ORCHIDAS are presented by Peylin et al. (2016).

169 In this work, the ORCHIDAS drives the ORCHIDEE routing scheme which is
 170 computationally less expensive than the full ORCHIDEE model (Fig. 1b). The data assimilation
 171 approach relies on the minimization of a misfit function $J(x)$ (aka cost function) by successive calls
 172 to “gradient-descent” minimization algorithm L-BFGS-B (Limited-memory Broyden-Fletcher-
 173 Goldfarb-Shanno algorithm with simple Box constraints, Byrd et al., 1995).

174 A new vector of parameter values x is estimated at each iteration. The $J(\mathbf{x})$ measures the
 175 mismatch between the vector of observed river discharges Q_{obs} and corresponding simulated
 176 values $Q_{sim}(x)$, as well as between the optimized correction factors x and its prior information x_{prior} :

$$177 \quad J(\mathbf{x}) = [\mathbf{Q}_{obs} - \mathbf{Q}_{sim}(\mathbf{x})]^t \mathbf{R}^{-1} [\mathbf{Q}_{obs} - \mathbf{Q}_{sim}(\mathbf{x})] + (\mathbf{x} - \mathbf{x}_{prior})^t \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_{prior}), \quad (7)$$

178 where \mathbf{R} and \mathbf{B} represent the prior error covariance matrices for observations and parameters,
 179 respectively. Diagonal elements of \mathbf{R} matrix represent the data uncertainties, which include both

180 the measurement errors (systematic and random) and model errors, we have defined it as the root
181 mean squared error (RMSE) between the prior model simulations and the observed river
182 discharges. Non-diagonal elements describe correlations between the data, which however are
183 difficult to presume correctly, and are usually neglected. The prior parameter uncertainties (matrix
184 **B**) have been set to 40% of the range of variation of correction factors obtained from the ratio Q_{obs}
185 and first guess value of river discharge simulation (Q_{fg}) obtained from x_{prior} . The matrix **B** was
186 determined based on the expert knowledge of ORCHIDEE model (Kuppel et al., 2012; Santaren
187 et al., 2014). Correlations between prior parameter values have not been considered. The gradient
188 of the $J(x)$ is calculated for all the parameters by finite difference approach at each iteration
189 (Kuppel et al., 2012).

190 **2.2.2. ORCHIDEE LSM with high-resolution river routing model**

191 The ORCHIDEE LSM is the land component of Institut Pierre Simon Laplace Climate
192 Model (IPSL-CM), which simulates energy, water and carbon cycles between the soil and
193 atmosphere. The unsaturated water flow is described at each land point by the one-dimensional
194 Richards equation with 2 m soil discretized to 11 levels. The surface runoff and deep drainage at
195 bottom layer are computed by Horton overland flow and free drainage (equals to hydraulic
196 conductivity), respectively. In other words, the ORCHIDEE LSM assumes that the aquifer level
197 is below the model bottom, and it neglects the upward water flow through capillary forces from its
198 underlying aquifer. The evaporation is partitioned into transpiration, bare soil evaporation,
199 interception loss and snow sublimation.

200 The ORCHIDEE is coupled with the ocean model through the river routing scheme
201 (Polcher, 2003; Ducharne et al. 2003; Guimberteau et al., 2012) which computes river discharge
202 by integrating the surface runoff and deep drainage over the basin. A high-resolution river routing
203 scheme was developed recently, which allows to better describe of catchments boundaries, flow
204 direction, and water residence time (Nguyen-Quang et al., 2018; Zhou et al., 2018). It is based on
205 the HydroSHED (Hydrological data and maps based on SHuttle Elevation Derivatives at
206 multiple Scales; <http://www.hydrosheds.org/>; Lehner et al., 2008) map with 1 km spatial resolution.
207 There are several hydrological transfer units (HTUs) in one ORCHIDEE grid-cell (e.g., 100 in the
208 current study). The HTU is constructed based on the Pfafstetter topological coding system and

209 user defined size. Each HTU represents the section of the river basin within the grid box, and many
210 HTUs forms a river basin (Nguyen-Quang et al., 2018). Therefore, the relative locations of HTUs
211 in each grid cell are not fixed.

212 In each HTU, the water is routed through a cascade of three linear reservoirs characterized
213 by their residence times: the groundwater, overland and stream reservoirs. The runoff and drainage
214 are the inputs into the overland reservoir and groundwater reservoir, then they flowed into the
215 stream reservoir of the downstream sub-grid basin. The residence times are determined by
216 multiplying a constant reservoir factor (g) with a slope index (k). The g for stream, overland and
217 groundwater reservoirs are 0.24, 3, and 25 day/km, respectively (Ngo-Duc et al., 2007). The slope
218 index is a function of distance (d) and slope (S) between a pixel and its downstream pixel ($k=d/S^{1/2}$
219 defined by Ducharme et al., 2003). The water can flow either to the next HTU within the same grid
220 cell or to the neighboring cell. The river discharge is diagnosed at the HTU level in the assimilation.
221 The river discharge is linear with R and D at annual scale over a small basin. In case of more than
222 one observation stations are assimilated in a river basin (e.g., x_1 and x_2 in Fig. 1a), the river
223 discharge at downstream is affected by the discharge of upstream thus it is not a linear system
224 anymore. Therefore, the optimization is needed to deal with the x over the non-linear sub-basins.

225 The time steps for the ORCHIDEE model and routing scheme are 30 minutes and 3 hours,
226 respectively. The spatial resolution of the model depends on the resolution of the atmospheric
227 forcing, and it is 0.5° for the current study (given in Sect. 2.3.2). The soil texture map is from
228 United States Department of Agriculture (USDA) with 12 soil textures (Reynolds et al. 2000). The
229 vegetation map is from the European Space Agency Climate Change Initiative (ESA CCI,
230 <https://www.esa-landcover-cci.org/>) reduced to the 13 plant functional types represented by the
231 model.

232 **2.3. The study domain and the datasets**

233 **2.3.1. Study domain**

234 The assimilation system is applied over the Iberian Peninsula. This region is dominated by
235 two climate types: the oceanic climate in the Atlantic coastal region and the Mediterranean
236 climate over most of Portugal and Spain. The annual precipitation is extremely unevenly

237 distributed with more than 1500 mm over northeastern Portugal, much of coastal Galicia and along
238 the southern borders of the Pyrenees but less than 300 mm over southeast Spain (Estrela et al.,
239 2012). Over Spain, agriculture occupies approximately 50% of the land area (e.g., year 2014,
240 <https://data.worldbank.org/indicator/AG.LND.AGRI.ZS>), and with around 1200 large dams
241 (European Working Group on Dams and Floods, 2010).

242 **2.3.2. The meteorology forcing**

243 In order to study the sensitivity of the optimization results to different forcing data, three
244 meteorology forcing are used: WFDEI_GPCC, WFDEI_CRU and CRU_NCEP. The
245 WFDEI_GPCC and WFDEI_CRU (3-hourly, 0.5°) are based on the WFDEI meteorological
246 forcing data which was produced using WATCH (WATER and global CHange) Forcing Data
247 (WFD) methodology applied to ERA-Interim data at 0.5° (Weedon et al., 2014; [http://www.eu-
248 watch.org/data_availability](http://www.eu-watch.org/data_availability)). The WFDEI is from 1979 and updates until now with eight
249 meteorological variables at 3-hourly time steps. The precipitation of WFDEI_GPCC and
250 WFDEI_CRU is corrected by GPCC (Global Precipitation Climatology Centre) and CRU
251 (Climatic Research Unit), respectively. The CRU_NCEP (6-hourly, 0.5°) combines the CRU
252 TS.3.1 (0.5°, monthly) climatology covering 1901-2012 and the NCEP (National Centers for
253 Environmental Prediction) reanalysis (2.5°, 6-hour) beginning in 1948
254 (<https://vesg.ipsl.upmc.fr/thredds/fileServer/store/p529viov/cruncep/readme.html>). The
255 precipitation of the three forcing is compared with the IB02 which is a gridded daily rainfall dataset
256 for the Iberia Peninsula with 0.2° resolution covers 1950 to 2003 (Belo-Pereira et al., 2011). It is
257 generated by using ordinary kriging from more than 2400 quality-controlled stations.

258 **2.3.3. The GRDC dataset**

259 The Global Runoff Database collects the monthly river discharge from most basin agencies
260 around the world (more than 9,300 stations) with an average record length of 43 years. Although
261 the quality of the observations is unknown (e.g., monitoring the river transect, velocity
262 measurements, etc.), the GRDC datasets are the most complete river discharge dataset available
263 today. It is hosted by the German Federal Institute of Hydrology

264 (Bundesanstalt für Gewässerkunde or BfG;
265 www.bafg.de/GRDC/EN/Home/homepage_node.html).

266 **2.3.4. Integration of GRDC in ORCHIDEE**

267 The location of some stations in the GRDC dataset might be incorrect for either the
268 longitude or latitude coordinate due to simple typos, logical errors in the original coordinates, or a
269 swapped order of the coordinate digits (Lehner, 2012). Due to this uncertainty, a quality control is
270 applied for GRDC when matching it with the corresponding HTUs in the river routing model. **For**
271 **each GRDC station, the corresponding catchment surface in the model is estimated.** The matching
272 process is stringent, and the GRDC qualification is restricted by two matching criteria: (1) the
273 difference in upstream area between GRDC and the model is less than a pre-defined percentage;
274 (2) the distance between GRDC and the model is less than a pre-defined distance. The higher the
275 two thresholds are, the more the matched GRDC stations can be positioned on the model's basin
276 representation. Meanwhile, the high threshold increases the uncertainties of the GRDC data due to
277 the errors in location and upstream area. By compromising between the two contradictory
278 requirements (the number of GRDC stations and the precise of the data), we choose the threshold
279 for upstream area difference and distance to be 10% and 25 km, respectively. Under this constraint,
280 27 GRDC stations are qualified among all 65 stations over the Iberian Peninsula domain (10°W-
281 5.5°E, 34°N-45.5°N; Fig. 2). It should be noted one GRDC station can match with several model
282 HTUs that locate in different model grids. In this case, the HTU with the lowest upstream area
283 difference is chosen. Therefore, the GRDC station is not necessarily in the same model grid as the
284 model HTU.

285 **2.3.5. The evaporation products**

286 The bias corrected evaporation deduced from the assimilation is compared with the
287 GLEAM (Global Land Evaporation Amsterdam Model; Martens et al., 2017;
288 <https://www.gleam.eu/>) product. GLEAM provides daily evaporation from 1984 to 2011 at 0.25°.
289 The evaporation is estimated by a minimalistic Priestley-Taylor potential evaporation model with
290 the majority of inputs estimated from remote sensing. It uses the microwave-derived soil moisture,
291 land surface temperature and vegetation density, and the detailed estimation of rainfall interception

292 loss. The rainfall interception loss is estimated separately using the Gash analytical model which
293 considers the canopy storage capacity, coverage, and the ratio of mean evaporation rate from wet
294 canopy. There are several versions of GLEAM data available, and we choose the latest version
295 v3.1a. The precipitation forcing of GLEAM v3.1a is from the Multi-Source Weighted-Ensemble
296 Precipitation (v1.2).

297 **2.4. Experiments design**

298 An ORCHIDEE simulation is performed to obtain the Q_{fg} and the corresponding R and D .
299 The ORCHIDAS with L-BFGS-B algorithm explores the full space of x by perturbing a separate
300 $x (x_i)$ over the i th upstream catchment ($i=1, 2, \dots, N_{opt}$; N_{opt} is the total number of optimized x
301 depending on the number of observation stations) in each iteration. To save computing time, the
302 river routing parameterization (forced by corrected R and D) rather than the full ORCHIDEE is
303 executed. The total execution time depends on the number of parameters to be optimized, the
304 length of simulation years, and the number of iterations. Multi-level parallelisms of the
305 assimilation are implemented to achieve the high computational efficiency. In each iteration, the
306 assimilation can run with N_{opt} ‘river routing’ simulations, with each ‘river routing’ model
307 parallelized with $N_{routing}$ CPUs ($N_{opt}=27$, $N_{routing}=16$ over the study domain). Over the Iberian
308 Peninsula, the range of x is defined between 0 and 20 which is determined by Q_{fg} and Q_{obs} .

309 In order to check the impacts of prior information x_{prior} on the optimization convergence
310 time, the x_{prior} is set to a constant value ‘1’ (x_{prior_1}) or a ‘pre-estimated-prior’ (x_{prior_ref} , defined as
311 the ratio of Q_{obs}/Q_{fg}), separately. The optimal x values are assigned over the whole study domain.
312 The x of the sub-catchment without GRDC station available is set to 1 (no correction). The
313 climatology values (e.g., over 1979-2014) are applied to fill the observation missing values over
314 certain period. In case of more than one GRDC stations locate in the same model grid, the averaged
315 correction factor is used.

316 The optimization results are not sensitive to the choice of x_{prior} , but the convergence time
317 indeed depends on x_{prior} . Fig. 3a shows that the x_{prior_ref} method requires less iteration to converge
318 than x_{prior_1} (7 and 15-20 iterations, respectively). The value of the cost function of x_{prior_ref} method
319 is lower than that of x_{prior_1} for all iteration steps. The normalized bias ($Norm_BIAS$) of discharge

320 after 7 iterations is less than 0.3 for the x_{prior_ref} method, while it is larger than 0.6 over most south
 321 regions for x_{prior_J} (Figs. 3b and 3c). The oscillation of J at the steps 3 and 5 could be due to the
 322 fact that the calculation of the gradient of J by finite difference is not optimal. It is also possible
 323 because the L-BFGS-B explores partly the physical range during the first few iteration to estimate
 324 the Hessian of the cost function for convergence.

$$325 \quad Norm_BIAS = \frac{Q_{sim} - Q_{obs}}{Q_{obs}}, \quad (8)$$

326 We choose x_{prior} set by x_{prior_ref} for n years ($n=10$, 1980-1989) experiment with iteration
 327 number k being 15 and number of correction factor m (i.e., the number of GRDC station) being 27.
 328 The x values vary with different years. Due to the slow variation in aquifer levels, a spin-up is
 329 necessary before optimization to get equilibrium of aquifer levels in LSM. The spin-up creates the
 330 aquifer initial states ($A^0_{corr}, A^1_{corr}, A^2_{corr}, \dots, A^{10}_{corr}$) at the start of the assimilation cycles over each
 331 ORCHIDEE model grid (Fig. 4), making it adapt to the bias corrected aquifer states.

$$332 \quad \frac{dA^i_{corr}}{dt} = \left[\int_S x(R_2 + D_2) \right] - Q_{corr,2} + Q_{corr,1}, 0 \leq i \leq 10 \quad (9)$$

333 To test different assumptions of errors in initial conditions, we implemented different
 334 optimization methods with each method results in a group ($m \times n$) of optimal x (Fig. 4). In method
 335 1, the optimization is carried out year by year with one-year spin-up for each iteration ('Y1SP1'
 336 here after). The x of the optimization year is applied during simulation. The method 2 is similar
 337 with Y1SP1 except that it uses optimized aquifer levels from the previous year ('Y1SP0' here
 338 after). This method assumes the final state variables (aquifer levels) of the optimal solution at the
 339 current optimization year is the best initial condition for the following assimilation year. In method
 340 3, the optimization is done over 10 years continuously with 1-year spin-up at the beginning of each
 341 10-year simulation ('Y10C' here after). The Y10C optimizes 270 x over 10 years together, while
 342 the Y1SP1 and Y1SP0 optimize the 10 years separately with 27 x each year. The 'river routing'
 343 model running years required by the three methods are 8100 ($=m \times 2 \times n \times k$), 4050 ($=m \times n \times k$) and
 344 44550 [$=m \times n \times (n+1) \times k$], respectively. Take the Y1SP0 for example, in each iteration, the
 345 correction factor x is perturbed by m times. For each perturbation, the ORCHIDEE river routing
 346 model runs once with one x (e.g., x_i at the i th sub-catchment) being perturbed while the x of other

347 sub-catchments are kept the same. Therefore, the total number of years required for m stations, n
348 iterations and k years assimilation is $m \times n \times k$. For all experiments, the optimization is carried out at
349 daily scale, and the diagnostics are performed for annual averages where we assume the water
350 storage variation is neglectable.

351 In order to further identify the impacts of atmospheric forcing on optimizations (e.g.,
352 optimal correction factor x), we measure the ‘Uncertainty’ of the variable (‘ var ’ in equation; ‘ var ’
353 refers to x , corrected evaporation, etc.) by Eq. (10). The higher the ‘Uncertainty’ is, the larger the
354 uncertainty is. The 0 value means that all the three ‘ var ’ values are equal.

$$355 \quad \text{Uncertainty}(var) = \frac{|var_1 - var_2| + |var_2 - var_3| + |var_1 - var_3|}{3} \quad (10)$$

356 3. Results and discussions

357 3.1. Evaluation of river discharge without assimilation

358 Fig. 5 displays the first guess simulation forced with different atmospheric forcing:
359 WFDEI_GPCC (Figs. 5a-5b), WFDEI_CRU (Figs. 5c-5d), and CRU_NCEP (Figs. 5e-5f). The
360 *Norm_BIAS* and correlation coefficient (computed by the annual mean values) are used to measure
361 the qualities of the simulated discharge. The diagnostics at each GRDC station are spread to the
362 entire upstream basin which contributes to the errors in discharge at downstream. The correlation
363 coefficient between FG (forced by WFDEI_GPCC and WFDEI_CRU) and observation is greater
364 than 0.6 over most regions, but it is less than 0.2 over certain regions (e.g., middle and southeast
365 of the Iberian Peninsula Figs. 5a and 5c). The correlation coefficient obtained by using
366 CRU_NCEP forcing is less than 0.2 for most regions (middle and west of the Iberian Peninsula),
367 which is worse than the simulation from WFDEI_GPCC and WFDEI_CRU. Wang et al. (2016)
368 also show the relatively poor performance of CRU_NCEP in simulating global land surface
369 hydrology and heat fluxes by using the Community Land Model (CLM4.5). The spatial pattern of
370 the absolute bias in river discharge varies with the atmospheric forcing (not shown). The
371 normalized bias is then applied to measure the river discharge simulation. The *Norm_BIAS* in
372 discharge shows consistent spatial distribution for simulations of three forcing. The *Norm_BIAS*
373 (positive) is higher than a factor of 1.5 over south and northeast of the Iberian Peninsula, which

374 means the overestimation of river discharge. The *Norm_BIAS* is small (within +/- 0.3) over north,
375 west and southeast of the region (Figs. 5b, 5d and 5f).

376 **3.2. Comparison of the three optimization strategies forced by WFDEI_GPCC**

377 We apply the three assimilate approaches (Y1SP1, Y1SP0, Y10C) to ORCHIDEE
378 simulations to correct the bias in discharge simulation by WFDEI_GPCC forcing. Fig. 6 (left)
379 displays the geographical distribution of the average correction factor x obtained after the
380 assimilation. The x values range between 0 and 1.5 over the study domain. The perfect discharge
381 simulation corresponds to x equal 1. The x value lower than 1 means the discharge in FG
382 (WFDEI_GPCC) is overestimated and thus a decrease of R and D is required, and vice versa for x
383 being higher than 1. The further the x away from 1, the larger the corrections of runoff and drainage
384 are. The three methods display similar spatial distribution pattern with x being less than 0.5 over
385 south and east of the Iberian Peninsula and x being higher than 1 over north of the Iberian Peninsula.
386 This spatial distribution of x is highly consistent with the pattern of *Norm_BIAS* in FG (discharge
387 overestimated in south and northeast, underestimated in north).

388 Fig. 6 (central column) shows the correlation coefficient between corrected discharge and
389 GRDC observations. After assimilation, the correlation of the optimized discharge and
390 observations is larger than 0.8 over most regions. The correlation coefficient for assimilated
391 discharge and observation is less than 0.6 (but higher than 0.4) over some regions and seems very
392 dependent on the forcing. This is probably because there is a contradiction of x between the
393 upstream and downstream stations and thus the method has difficulties finding a compromise (e.g.,
394 over the Ebro basin). In general, the regions with low correlation coefficient are forcing dependent,
395 while the regions with high correlation coefficient are very consistent among different forcing. Fig.
396 6 (right) gives the *Norm_BIAS* in discharge between assimilations and observations. After
397 assimilation, this positive bias in river discharge has been significantly reduced (within ± 0.3). It
398 should be mentioned that the x_{prior_ref} is able to capture the general distribution pattern of optimal
399 x , but the performance of river discharge estimation is significantly improved through optimization.
400 The role of optimization is to find an appropriate correction factor when there are several basins
401 (with observations) overlaps at upstream

402 A common validation approach is to compare the assimilated river discharge with other
403 independent data sources. However, the river discharge observations are limited, and the GRDC
404 is the only comprehensive river discharge datasets at global scale so far. To overcome this
405 limitation, the assimilated river discharges are also validated over the catchments where the GRDC
406 stations are discarded during assimilation. Fig. 7 shows the annual mean of river discharge over
407 the Alcala Del Rio station (-5.98°W, 37.52°N) on the Guadalquivir river (locates at southwest of
408 Spain) before and after correction. The observation of this station is not assimilated due to its large
409 upstream area difference (15.53% > 10%) between model (55635 km²) and GRDC (46995 km²).
410 The overestimated discharge simulated by the model at this station is also corrected because it
411 benefits from the correction factor estimated at the Cantillana station (-5.83°W, 37.59°N; 44871
412 km²) which locates at the 15.3 km upstream of Alcala Del Rio station of the Guadalquivir River
413 (southwest of the Iberian Peninsula). Between the two stations, there are several tributaries flow
414 to Alcala Del Rio station, which leads to different annual mean river discharges at Cantillana (49.7
415 m³/y) and Alcala Del Rio stations (94.8 m³/y). This result illustrates that this approach is able to
416 correct the river discharge over the entire basin. The discharges for certain sub-basins without
417 assimilated observations (e.g., observation unavailable or GRDC stations discarded) are corrected
418 by x as well. Although the validation datasets are from the same GRDC source, they are from other
419 independent observation stations thus can be seen as an independent validation ('first order
420 validation').

421 In summary, all the three methods (Y1SP1, Y1SP0, and Y10C) are able to improve the
422 river discharge simulation by ORCHIDEE LSM. The correlation coefficient and *Norm_BIAS* in
423 discharge obtained from the three methods are generally consistent. The correlation coefficient of
424 Y10C method in northeast is lower than that of Y1SP0 and Y1SP1, which is probably resulted
425 from its poor quality of atmospheric forcing. The Y1SP0 consumes less computing time than
426 Y1SP1 and Y10C, and it does not worsen the optimization results. By compromising between the
427 accuracy of results and the computing time, we choose Y1SP0 method for the further assimilation.

428 The above assimilations are performed with the same forcing (WFDEI-GPCC) by
429 assuming the errors in discharge are caused by model defect (e.g., model parameterization, model
430 structure, etc.). The uncertainties of simulated discharge also result from the atmospheric forcing.
431 The role of atmospheric forcing in assimilation is discussed in following section.

432 3.3. The sensitivity of the optimizations to atmospheric forcing

433 In order to understand the response of the optimizations to different atmospheric forcing
434 with different precipitation sources, the ORCHIDAS was also run with WFDEI_CRU and
435 CRU_NCEP forcing using Y1SP0 optimization strategy. Using two other different forcing for the
436 assimilation can allow us to understand how important the forcing uncertainty affects the
437 correction factor. The multi-year mean correction factor x obtained from WFDEI_CRU (Fig. 8a)
438 CRU_GPCC (Fig. 8b), and WFDEI_GPCC (Fig. 8c) displays quite consistent spatial patterns. The
439 coverage of low correction factor (blue in Figs. 8a-8b, corresponds to large correction) obtained
440 from CRU-NCEP is larger than that obtained from WFDEI_CRU and WFDEI_GPCC. This is
441 because the positive bias in discharge of FG simulation forced by CRU-NCEP is larger than that
442 by WFDEI_CRU and WFDEI_GPCC. Besides the atmospheric forcing, the uncertainties could
443 also origin from boundary condition (e.g., topographic or other land surface features), model
444 parameter, model structure or missing processes. For all forcing, the x is less than 0.3 (but greater
445 than 0) over south, which implies that the error in discharge is probably resulted from the missing
446 model processes (human activity). Over north, the x are close to 1 (discharge well simulated) for
447 all the three forcing, which indicates the correction comes from model ‘random’ error (nature
448 discharge) rather than the system error (e.g., missing processes).

449 The uncertainty of x by three forcing is small for most regions (Fig. 8d). The high
450 uncertainty of x over the Adoure (southwestern France) and the Chelif (in Algeria) river basins
451 corresponds to the large uncertainty in the different atmospheric forcing. This result demonstrates
452 the obtained correction factor x is robust in spite of using different atmospheric forcing. This is
453 also demonstrated by comparing the precipitations between the three forcing and the IB02 dataset.
454 Compared to the IB02, all the three forcing overestimate rainfall in the Iberian Peninsula (Figs.
455 S1a-S1c), but none of these error patterns resembles that of the proposed E correction (Figs. 9e-
456 9g). Unlike the pattern of the correction factor (Figs. 8a-8c), the ratios of annual mean precipitation
457 between the three forcing and the IB02 are higher than 1 over most regions (Figs. S1d-S1f).
458 Therefore, the precipitation forcing error is not likely the dominant factor in determining the
459 correction factor distribution.

460 In summary, the assimilation approach is able to correct errors in lateral water balance
461 despite using different forcing. Recalling that the corrected $R+D$ (through x) and the precipitation
462 are known, we then transfer the optimal correction factor x to the vertical water balance equation
463 (Eq. 5) to derive the bias corrected evaporation. This will enable us to understand the impacts of
464 assimilation on evaporation.

465 **3.4. Evaporation estimations through the optimal correction factor**

466 The evaporation of FG simulation by different forcing show quite consistent spatial
467 distribution (Figs. 9a-9c) and small uncertainty (<0.2 mm/d, Fig. 9d) with the value being higher
468 over north than south. The change of evaporation (dE) induced by the correction is consistent for
469 three forcing (Figs. 9e-9g) with low uncertainties (Fig. 9h). It should be mentioned that the
470 evaporation for the regions without GRDC stations are not corrected (i.e., correction factor x equals
471 1) such as southern France, western Portugal, and northwest, south and southeast of Spain (blank
472 regions in Fig. 8). The dE is positive (around 0.2 to 0.4 mm/d) over south and northeast where the
473 evaporation is underestimated in FG. Cazcarro et al. (2015) show large blue water footprint
474 (volume of surface and groundwater consumed for production an item) of human activity over
475 south (Jaén, Sevilla, and Malaga provinces), northeast (Palencia, Burgos, La Rioja, Navarra and
476 Valladolid provinces), north (Tarragona province) and middle (Toledo province) of Spain (Map.
477 1 of that paper). The large dE over south and northeast obtained in current study is consistent with
478 the blue water footprint of Cazcarro et al. (2015). Figs 9i-9k plot the change of the ratio of water
479 demand (dE) and water supply ($R+D$). This ratio measures the degree of water shortage. The
480 greater the ratio, the higher level of water shortage. The ratio is larger over south and northeast of
481 Spain, which is consistent with the results from other studies that measures the water deficits
482 (Rodríguez-Díaz et al., 2007) and water exploitation index (Pedro-Monzonís et al., 2015) in Spain.
483 Since we assume that the missing human processes is the main error in ORCHIDEE, the dE and
484 $dE/(R+D)$ indicate the changes induced by human processes. The spatial patterns of dE and
485 $dE/(R+D)$ are quite consistent with human water exploitation, thus the model missing processes
486 (e.g., human water usage) is considered as the dominant contribution to x .

487 We also tested the possibility of improving the river discharge estimation by using a annual
488 constant correction factor to evaporation (X_{Ecorr}), which can be derived from Eq. (6).

489
$$X_{Ecorr} \approx \frac{E + (1 - x) \cdot (R + D)}{E}, \quad (11)$$

490
$$E_{corr} = X_{Ecorr} \cdot E \quad (12)$$

491 Although the Eqs. 11-12 are able to improve river discharge estimation by modifying soil
 492 moisture, the energy and water balance are not conserved. One solution could be to run the full
 493 ORCHIDEE LSM in the assimilation system with the same cost function as Eq. (7). In this way,
 494 the intermediate variables are adjusted towards optimal river discharge with the modification of
 495 evaporation. This approach executes the full ORCHIDEE model thus is very time consuming and
 496 is beyond the scope of the current study.

497 **3.5. The inter-annual variation of correction factor and water cycle**

498 **3.5.1. The inter-annual cycles**

499 All the results so far are obtained by averaging multi-year mean values which provides us
 500 the bias correction information at spatial scale. To understand the inter-annual cycles of the
 501 correction and its possible contribution, we analyze the assimilation results over two stations at
 502 south of Spain where the discharge correction is large during the period of 1980 - 1989 (Fig. 8).

503 The Puente De Palmas station locates on the Guadiana River (southwest of the Iberian
 504 Peninsula) with an upstream area of 48515 km². The three FG simulations (with different forcing)
 505 significantly overestimate the river discharge and the runoff coefficient (ratio of discharge and
 506 precipitation), while the FG(WFDEIG) and FG(WFDEIC) underestimate the inter-annual
 507 variability comparing with observations (Fig. 10a-10b). The standard-deviation of the annual
 508 means for observation, FG(WFDEIG), FG(WFDEIC) and FG(CRUN) are 33.8 m³/s, 28.8 m³/s,
 509 25.2 m³/s and 34.3 m³/s, respectively. One reason could be the variation of water usage by
 510 irrigated agriculture which occupies 90% of the blue water usage (surface water and groundwater)
 511 in this semiarid basin (Aldaya and Llamas, 2008) or model errors. Besides, there are many
 512 interconnected wetlands and structurally complex hydrogeological boundaries between the two
 513 upper-Guadiana aquifer in the upper Guadiana River basin (Van Loon and Van Lanen, 2013).
 514 These complex features are difficult to represent in model thus large bias exist in river discharge

515 of ORCHIDEE. The correction factor corrects these model defects (Fig. 10c) and it demonstrates
516 good skill in correcting the inter-annual variability of discharge and runoff coefficient (Fig. 10a-
517 10b).

518 The Masia De Pompo station (17876 km²) is on the Jucar River (southeast of Spain). The
519 observations over the year 1983, 1988-1989 are obtained from the climatology values due to
520 the unavailability of GRDC data during this period. During 1980-1989, the inter-annual
521 variation of observed discharge (and runoff coefficient) and FG simulation is quite inconsistent
522 (Figs. 10d-10e). This is probably caused by the surface water usage which occupies about 55%
523 over this basin (Kahil et al., 2016). Most of them are used for agriculture (>80%) and urban
524 (>10%). Although the improvements in assimilated discharge are small, the correction factor is
525 able to capture the inter-annual variability in observations (Figs. 10d and 10f).

526 In summary, the inter-annual variation river discharge of FG simulation and
527 observations does not agree each other over the Guadiana River basin and the Jucar River basin
528 during 1980-1989. The human water usage (e.g., groundwater or surface water extraction)
529 process, which is neglected in current ORCHIDEE model, is likely to play an important role in
530 river discharge variation. The optimized correction factor (varies each year) improves the inter-
531 annual variability of the modelled river discharge.

532 3.5.2. The geographical distribution

533 To further understand the inter-annual variability of corrections over the entire Iberian
534 Peninsula region, Fig. 11 plots the spatial distribution of inter-annual variability of correction
535 factor x and river discharge which is quantified by coefficient of variation as used by Déry et al.
536 (2011) and Siam and Eltahir Elfatih (2017). In FG (WFDEI_GPCC) simulation, the inter-annual
537 variation of discharge is lower than 0.4 over most regions, which indicates an underestimation of
538 inter-annual variability of river discharge in FG. The inter-annual variability of discharge is
539 increased after assimilation over south and northeast. This change could be attributed to the
540 fluctuation of correction factor (human water usage) over these regions. This result agrees with the
541 results (Map. 6) of Cazcarro et al. (2015) with more large dams in south and northeast (nature
542 discharge greatly affected by human) than northwest of Spain (nature discharge less affected by

543 human). The inter-annual variability of correction factor x and discharge for Y1SP0 (CRUN) is
544 different from others, which mainly results from the different atmospheric forcing.

545 **3.6. Comparison of bias corrected evaporation with GLEAM data**

546 In order to evaluate the bias corrected evaporation, Figs. 12a-12h compare the GLEAM
547 product (v3.1a) with FG and with bias corrected E by assimilation using WFDEI_GPCC,
548 WFDEI_CRU and CRU_NCEP forcing. Due to the unavailability of parts of GLEAM's
549 atmospheric forcing (e.g., air pressure, air humidity, air speed, etc.) and difficulty of maintaining
550 a coherence with other forcing, the assimilation system does not run with GLEAM's precipitation
551 input. We find large difference between GLEAM and FG, which indicates that the evaporation is
552 quite uncertain for different estimations. The geographical distribution and magnitude of
553 difference in E between GLEAM and FG is highly consistent with that between GLEAM and bias
554 corrected values by using different forcing (Figs. 12a-12c, and 12e-12g). The systematic negative
555 difference is higher than the uncertainties of bias corrected E with different forcing (Figs. 12d and
556 12h). Parts of the differences are explained by the lower P of GLEAM than ORCHIDEE forcing
557 (Figs. 12i-12l). Generally, the $P-E$ (in mm/d) of GLEAM is higher than bias corrected value
558 associated with small uncertainties (Figs. 12m-12t). Because the bias corrected $P-E$ are corrected
559 by GRDC observed river discharge, the $P-E$ (\approx river discharge) of GLEAM is very likely to be
560 higher than GRDC observations over the Iberia. This result indicates that some processes are
561 probably also missing in GLEAM v3.1. We also compared our bias corrected E with GLEAM v1
562 data (Miralles et al., 2011), and we find the $P-E$ between GLEAM v1 and bias corrected values
563 are quite consistent for different forcing. The results are quite consistent when comparing the
564 corrected E with several other products which are obtained by using different methodology and
565 forcing (e.g., Jung et al., 2009; Vinukollu et al., 2011; Mueller et al., 2013). Considering the
566 availability of $P-E$ for GLEAM data which allows to compare it with the bias corrected value, only
567 the results of GLEAM are shown.

568 **4. Conclusions**

569 There has been several studies working on estimation of fresh water input from continent
570 to ocean (e.g., the Mediterranean Sea) based on observation or modelling approach (e.g., Boukthir
571 and Barnier, 2000; Mariotti et al., 2002; Struglia et al., 2004; Peucker-Ehrenbrink, 2009; Ludwig

572 [et al., 2009; Szczypta et al., 2012](#)). However, these estimations are limited either by the coarse
573 temporal resolution for observation approach or by the non-comprehensive representation of
574 physical processes (e.g., human activities) for modelling approach. As a result, the fresh water
575 estimations are accompanied with large uncertainties among varies studies. This proposed
576 methodology aims to improve the estimation of continental water cycles by merging the merits of
577 observations and modelling approach through data assimilation.

578 The basis of the method is the vertical and lateral water balance equations. The method
579 assumes that the precipitation minus evaporation from the model simulation is an appropriate first
580 guess so that all the errors in river discharge end up with runoff and drainage. Under this
581 assumption, the river discharges simulation at river outlet are expected to be improved by
582 correcting the runoff and drainage (inputs for river routing model).

583 The idea is achieved by embedding a river routing scheme of ORCHIDEE LSM and GRDC
584 river discharge observations into a data assimilation system (ORCHIDAS). The system can run
585 with multi-level parallel computing mode (both the routing model and [the](#) optimization are
586 parallelized). The river discharge is optimized through applying a correction factor x to model
587 runoff and drainage which translates errors in estimated $P-E$.

588 The method has been explained through its application over [the](#) Iberian Peninsula with 27
589 GRDC stations during 1979-1989 with x values being different each year. Main conclusions are:
590 First, the optimization results are not sensitive to x prior information x_{prior} , and assimilation
591 strategies, but the setting of x_{prior} by a ‘pre-estimated-prior’ (defined as Q_{obs}/Q_{fg}) indeed converges
592 faster than other x_{prior} values. The method Y1SP0 (the model spin-up uses the optimal aquifer
593 levels of previous optimization year) demonstrates high computing efficiency and comparable
594 discharge accuracy comparing with the other two methods (Y1SP0, Y10C), thus the Y1SP0 is
595 recommended (e.g., over [the](#) full Mediterranean catchment). Second, the largest correction of
596 discharge is found over south and northeast of [the](#) Iberian Peninsula. These regions are
597 characterized by large blue water footprint with large groundwater and surface water usage by
598 human activity. It implies that most of the corrections by x represents the missing human processes
599 (at least in the south of study domain). This is consistent with the fact that ORCHIDEE model
600 neglects the human processes (e.g., dam operation, irrigation, etc.). The discharge correction over

601 north of the Iberian Peninsula is relatively small, where is mainly due to model systematic error.
602 The correction factor x can also cover errors in the model structure, model parameter, or boundary
603 conditions (e.g., land surface characteristics imposed to the model). Third, the assimilated
604 discharges reveal lower bias (from $>100\%$ to $<30\%$) and higher inter-annual variability (due to the
605 fluctuation of water usage) than uncorrected ones. Fourth, the bias corrected evaporation are
606 compared with the GLEAM v3.1a product. The E of GLEAM is lower than the optimized E , while
607 the $P-E$ of GLEAM is higher than the optimized values. This different $P-E$ could be caused by the
608 different P forcing and the missing processes in the GLEAM model.

609 The method takes into account both gauged rivers (usually large rivers) and un-gauged
610 rivers, and it provides discharge estimates at daily scale from 1980 to 2014 with the time range
611 depend on atmospheric forcing. By using the correction factor of adjacent catchment, this method
612 also improves the river discharge simulation for the catchment without assimilating observations.
613 Besides, this method fills the gap of the data missing period (e.g., war, instruments, etc.) by
614 climatology values, thus the data are complete over the whole period. The proposed method is
615 supposed to be superior to the simple water-balance methods, because a LSM estimates E at sub-
616 diurnal scales with physically based equations and takes advantage of spatial distribution of the P
617 and $P-E$.

618 The result implies the necessity of parameterizing the human water uptake process in the
619 ORCHIDEE LSM. Besides, the poor quality of the river discharge observations (e.g., 68% stations
620 are discarded over the Iberian Peninsula) calls for a high quality data. The optimized correction
621 factors x are model and atmospheric forcing dependent. It is encouraged to apply this assimilation
622 method to other models, which will allow us to identify the sources of errors (e.g., model missing
623 process or forcing data). To improve the calculation efficiency, this study uses annual mean
624 correction factors without considering its seasonal variation thus the seasonal discharges do not
625 improved. One issue of the x optimization could be the equifinality with a number of optimized x
626 result in the similar river discharge at downstream. Future developments can be made towards
627 generating ensemble optimal x to better assess the uncertainties associated to each parameter x .
628 This assimilation method can be applied for water cycles studies, data inter-comparison, and
629 riverine fresh water estimation over other basins (e.g., the full catchment of the Mediterranean sea).

630 **Acknowledgments**

631 The authors gratefully acknowledge financial support provided by the STSE WACMOS-
632 MED (Water Cycle Multi-mission Observation Strategy for the Mediterranean) project under ESA
633 (Grant No. 4000114770/15/I-SBo) and the Earth2Observe (Global Earth Observation for
634 Integrated Water Resource Assessment) project of the FP7 (Grant No. 603608). The ClimServ
635 computational facilities at IPSL were used to perform all the simulations. [The authors also thank](#)
636 [the valuable and constructive comments from Emanuel Dutra \(Lisbon University\) and another](#)
637 [anonymous reviewer.](#)

638

639 **References:**

- 640 Aldaya, M. M. and Llamas, M. R.: Water footprint analysis for the Guadiana river basin, Value
641 of Water Research Report Series, No. 35, UNESCO–IHE Delft, The Netherland, 2008.
- 642 aus der Beek, T., Menzel, L., Rietbroek, R., Fenoglio-Marc, L., Grayek, S., Becker, M.,
643 Kusche, J., Stanev, E.V.: Modeling the water resources of the Black and
644 Mediterranean Sea river basins and their impact on regional mass changes, *J.*
645 *Geodyn.* 59–60, 157–167, <http://dx.doi.org/10.1016/j.jog.2011.11.011>, 2012.
- 646 Bauer-Gottwein, P., Jensen, I. H., Guzinski, R., Bredtoft, G. K. T., Hansen, S., and Michailovsky,
647 C. I.: Operational river discharge forecasting in poorly gauged basins: the Kavango River
648 basin case study, *Hydrol. Earth Syst. Sci.*, 19, 1469-1485, [https://doi.org/10.5194/hess-19-](https://doi.org/10.5194/hess-19-1469-2015)
649 [1469-2015](https://doi.org/10.5194/hess-19-1469-2015), 2015.
- 650 [Belo-Pereira, M., Dutra, E., and Viterbo, P.: Evaluation of global precipitation data sets over the](#)
651 [Iberian Peninsula, *J. Geophys. Res.*, 116, D20101, <https://doi.org/10.1029/2010jd015481>,](#)
652 [2011.](#)
- 653 Bricheno, L. M., Wolf, J. M., and Brown, J. M: Impacts of high resolution model downscaling in
654 coastal regions, *Cont. Shelf Res.*, 87, 1-16, 2014.
- 655 Boukthir, M. and Barnier, B.: Seasonal and inter-annual variations in the surface freshwater flux
656 in the Mediterranean Sea from the ECMWF re-analysis project, *Journal of Marine Systems*
657 24, 343–354, 2000
- 658 Bouraoui, F., Grizzetti, B. and Aloe, A.: Estimation of water fluxes into the Mediterranean Sea, *J.*
659 *Geophys. Res.*, 115, D21116, doi:10.1029/2009JD013451, 2010.
- 660 Byrd, R. H., Lu, P., Nocedal, J., and Zhu, C.: A limited memory algorithm for bound constrained
661 optimization, *SIAM J. Sci. Stat. Comput.*, 16, 1190–1208, 1995.
- 662 Cazarro, I., Duarte, R., Martín-Retortillo, M., Pinilla, V., and Serrano, A.: How sustainable is the
663 increase in the water footprint of the Spanish agricultural sector? A provincial analysis
664 between 1955 and 2005–2010, *Sustainability*, 7 (5), 5094-5119, doi:10.3390/su7055094,
665 2015.
- 666 Clark, E. A., Sheffield, J., van Vliet, M., Nijssen, B., and Lettenmaier, D. P.: Continental runoff
667 into the oceans (1950–2008), *J. Hydrometeor.*, 16, 1502–1520, doi:
668 <https://doi.org/10.1175/JHM-D-14-0183.1>, 2015

669 Dai, A. G. and Trenberth, K. E.: Estimates of freshwater discharge from continents: Latitudinal
670 and seasonal variations, *J. Hydrometeor*, 3, 660–687, 2002.

671 De Rosnay, P., Polcher, J., Bruen, M., and Laval, K.: Impact of a physically based soil water flow
672 and soil-plant interaction representation for modeling large-scale land surface processes, *J*
673 *Geophys Res*, 107: D11, doi:10.1029/2001JD000634, 2002.

674 Déry, S. J., Mlynowski, T. J., Hernández-Henríquez, M. A., and Straneo F.: Interannual variability
675 and interdecadal trends in Hudson Bay streamflow, *J. Marine Syst.*, 88, 341–351, 2011.

676 Ducharne, A., Golaz, C., Leblois, E., Laval, K., Polcher, J., Ledoux, E. and de Marsily, G.:
677 Development of a high resolution runoff routing model, calibration and application to assess
678 runoff from the LMD GCM, *J. Hydrol.*, 280, 207-228, 2003.

679 Estrela, T., Pérez-Martin, M.A., and Vargas, E.: Impacts of climate change on water resources in
680 Spain, *Hydrolog. Sci. J.*, 57(6), 1154-1167, doi: 10.1080/02626667.2012.702213, 2012.

681 European Working Group on Dams and Floods: Report on ‘Dams and floods in Europe, role of
682 dams in floods mitigation’, Pages 1-99, 2010.
683 [http://cnpqb.apambiente.pt/IcoldClub/jan2012/EWG%20FLOODS%20FINAL%20REPOR](http://cnpqb.apambiente.pt/IcoldClub/jan2012/EWG%20FLOODS%20FINAL%20REPORT.pdf)
684 [T.pdf](http://cnpqb.apambiente.pt/IcoldClub/jan2012/EWG%20FLOODS%20FINAL%20REPORT.pdf). Accessed 28 October 2017.

685 [Fekete, B. M., C. J. Vorosmarty, W. Grabs.: High-resolution fields of global runoff combining](#)
686 [observed river discharge and simulated water balances, *Global Biogeochemical Cycles*, 16](#)
687 [\(3\): 15-1 to 15-10, 2002.](#)

688 Guimberteau, M., Drapeau, G., Ronchail, J., Sultan, B., Polcher, J., Martinez, J., Prigent, C., Guyot,
689 J., Cochonneau, G., Espinoza, J., Filizola, N., Fraizy, P., Lavado, W., De Oliveira, E.,
690 Pombosa, R., Noriega, L. and Vauchel, P.: Discharge simulation in the sub-basins of the
691 Amazon using ORCHIDEE forced by new datasets, *Hydrol. Earth Syst. Sc.*, 16, 911-935,
692 2012.

693 Jin, F., Kitoh, A., and Alpert, P.: Water cycle changes over the Mediterranean: a comparison
694 study of a super-high-resolution global model with CMIP3, *Philos Trans R Soc A*, 368:
695 5137–5149, 2010.

696 Jordà, G., Von Schuckmann, K., Josey, S. A., Caniaux, G., Garcia-Lafuente, J., Sammartino, S.,
697 Özsoy, E., Polcher, J., Notarstefano, G., Poulain, P.-M., Adloff, F., Salat, J., Naranjo, C.,
698 Schroeder, K., Chiggiato, J., Sannino, G., and Macías, D.: The Mediterranean Sea Heat and

699 Mass Budgets: Estimates, Uncertainties and Perspectives, *Prog. Oceanogr.*,
700 doi:10.1016/j.pocean.2017.07.001, 2017.

701 Jung, M., Reichstein, M., and Bondeau, A.: Towards global empirical upscaling of FLUXNET
702 eddy covariance observations: validation of a model tree ensemble approach using a biosphere
703 model, *Biogeosciences*, 6, 2001–2013, doi:10.5194/bg-6-2001-2009, 2009.

704 Kahil, M., Albiac, J., and Dinar, A.: Improving the performance of water policies: Evidence
705 from drought in Spain, *Water*, 8, 34, 2016.

706 Kalma, J., McVicar, T., and McCabe, M.: Estimating land surface evaporation: a review of
707 methods using remotely sensed surface temperature data. *Surv. Geophys.*, 29 (4), 421-469,
708 doi: 10.1007/s10712-008-9037-z, 2008.

709 Kang, X., Zhang, R. and Wang G.: Effects of different freshwater flux representations in an ocean
710 general circulation model of the tropical Pacific, *Sci. Bull.*, 62: 345–351, 2017.

711 Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P.,
712 Sitch, S., and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled
713 atmosphere-biosphere system, *Global Biogeochem. Cycles*, 19: GB1015,
714 doi:10.1029/2003GB002199, 2005.

715 Kuppel, S., Peylin, P., Chevallier, F., Bacour, C., Maignan, F., and Richardson, A. D.:
716 Constraining a global ecosystem model with multi-site eddy-covariance data, *Biogeosciences*,
717 9, 3757-3776, <https://doi.org/10.5194/bg-9-3757-2012>, 2012.

718 Lehner, B., Verdin, K., and Jarvis, A.: New global hydrography derived from spaceborne elevation
719 data, *Eos, Transactions, AGU*, 89(10): 93-94, doi:10.1029/2008EO100001, 2008.

720 Lehner, B.: Derivation of watershed boundaries for GRDC gauging stations based on the
721 HydroSHEDS drainage network, GRDC Report Series, 41, Global Runoff Data Centre, 2012.
722 http://www.bafg.de/GRDC/EN/02_srvcs/24_rprtsrs/report_41.pdf?__blob=publicationFile.
723 Accessed: 29 September 2017.

724 Li, Y., Ryu, D., Western, A. W., and Wang, Q. J.: Assimilation of stream discharge for flood
725 forecasting: Updating a semidistributed model with an integrated data assimilation
726 scheme, *Water Resour. Res.*, 51, 3238–3258, doi:10.1002/2014WR016667, 2015.

727 Liu, X., Tang, Q., Cui, H., Mu, M., Gerten, D., Gosling, S., Masaki, Y., Satoh, Y., and
728 Wada, Y.: Multimodel uncertainty changes in simulated river flows induced by

729 human impact parameterizations, *Environ Res Lett.*, 12, 025009, doi: 10.1088/1748-
730 9326/aa5a3a, 2017.

731 Ludwig, W., Dumont, E., Meybeck, M. and Heussner, S.: River discharges of water and nutrients
732 to the Mediterranean and Black Sea: Major drivers for ecosystem changes during past and
733 future decades? *Prog. Oceanogr.*, 80, 199–217, doi:10.1016/j.pocean.2009.02.001, 2009.

734 [MacBean, N., Maignan, F., Peylin, P., Bacour, C., Bréon, F.-M., and Ciais, P.: Using satellite data
735 to improve the leaf phenology of a global terrestrial biosphere model, *Biogeosciences*, 12,
736 7185-7208, doi: 10.5194/bg-12-7185-2015, 2015.](#)

737 MacDonald, A. M., Bonsor, H. C., Ahmed, K. M., Burgess, W. G., Basharat, M., Calow, R. C.,
738 Dixit, A., Foster, S. S. D., Gopal, K., Lapworth, D. J., Lark, R. M., Moench, M., Mukherjee,
739 A., Rao, M. S., Shamsudduha, M., Smith, L., Taylor, R. G., Tucker, J., van Steenberg, F.
740 and Yadav, S. K.: Groundwater quality and depletion in the Indo-Gangetic Basin mapped
741 from in situ observations, *Nat. Geosci.*, 9, 762-766, 10.1038/ngeo2791, 2016.

742 Mariotti, A., Struglia, M. V., Zeng, N., and Lau, K-M.: The hydrological cycle in the
743 Mediterranean region and implications for the water budget of the Mediterranean Sea, *J.*
744 *Climate*, 15, 1674–1690, 2002.

745 Martens, B., Miralles, D.G., Lievens, H., van der Schalie, R., de Jeu, R.A.M., Fernández-Prieto,
746 D., Beck, H.E., Dorigo, W.A., and Verhoest, N.E.C.: GLEAM v3: satellite-based land
747 evaporation and root-zone soil moisture, *Geosci. Model Dev.*, 10, 1903–1925, doi:
748 10.5194/gmd-10-1903-2017, 2017.

749 Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., and
750 Dolman, A. J.: Global land-surface evaporation estimated from satellite-based observations,
751 *Hydrol. Earth Syst. Sci.*, 15, 453–469, doi:10.5194/hess-15-453-2011, 2011.

752 [Mueller, B., Hirschi, M., Jimenez, C., Ciais, P., Dirmeyer, P. A., Dolman, A. J., Fisher, J. B., Jung,
753 M., Ludwig, F., Maignan, F., Miralles, D., McCabe, M. F., Reichstein, M., Sheffield, J., Wang,
754 K. C., Wood, E. F., Zhang, Y., and Seneviratne, S. I.: Benchmark products for land
755 evapotranspiration: LandFlux-EVAL multi-dataset synthesis, *Hydrol. Earth Syst. Sci.*, 17,
756 3707-3720, doi:10.5194/hess-17-3707-2013, 2013.](#)

757 Munier, S., Palanisamy, H., Maisongrande, P., Cazenave, A., and Wood, E. F.: Global runoff
758 anomalies over 1993–2009 estimated from coupled Land–Ocean–Atmosphere water budgets

759 and its relation with climate variability, *Hydrol. Earth Syst. Sci.*, 16, 3647-3658,
760 <https://doi.org/10.5194/hess-16-3647-2012>, 2012.

761 Ngo-Duc, T., Laval, K., Ramillien, G., Polcher, J., and Cazenave, A.: Validation of the land water
762 storage simulated by Organising Carbon and Hydrology in Dynamic Ecosystems
763 (ORCHIDEE) with Gravity Recovery and Climate Experiment (GRACE) data, *Water Resour.*
764 *Res.*, 43, W04427, doi:10.1029/2006WR004941, 2007.

765 Nguyen-Quang, T., Polcher, J., Ducharne, A., Arsouze, T., Zhou, X., Schneider, A., and Fita, L.:
766 ORCHIDEE-ROUTING: A new river routing scheme using a high resolution hydrological
767 database, *Geosci. Model Dev. Discuss.*, <https://doi.org/10.5194/gmd-2018-57>, in review,
768 2018.

769 Pauwels, V. R. N., and De Lannoy, G. J. M.: Ensemble-based assimilation of discharge into
770 rainfall-runoff models: A comparison of approaches to mapping observational information to
771 state space, *Water Resour. Res.*, 45, W08428, doi:10.1029/2008WR007590, 2009.

772 Pedro-Monzonís M., Solera, A., Ferrer, J., Estrela, T., Paredes-Arquiola, J. A.: review of water
773 scarcity and drought indexes in water resources planning and management, *J*
774 *Hydrol*, 527:482–493. doi:10.1016/j.jhydrol.2015.05.003, 2015.

775 Peucker-Ehrenbrink, B.: Land2Sea database of river drainage basin sizes, annual water discharges,
776 and suspended sediment fluxes, *Geochem. Geophys. Geosyst.*, 10, Q06014,
777 doi:10.1029/2008GC002356, 2009.

778 Peylin, P., Bacour, C., MacBean, N., Leonard, S., Rayner, P., Kuppel, S., Koffi, E., Kane, A.,
779 Maignan, F., Chevallier, F., Ciais, P., and Prunet, P.: A new stepwise carbon cycle data
780 assimilation system using multiple data streams to constrain the simulated land surface carbon
781 cycle, *Geosci. Model Dev.*, 9, 3321-3346, <https://doi.org/10.5194/gmd-9-3321-2016>, 2016.

782 Pokhrel, Y. N., Felfelani, F., Shin, S., Yamada, T. J., and Satoh, Y.: Modeling large-scale human
783 alteration of land surface hydrology and climate, *Geoscience Letters*, 4(1): 1-13,
784 doi:10.1186/s40562-017-0076-5, 2017.

785 Polcher, J.: Les processus de surface a l'échelle globale et leurs interactions avec l'atmosphère,
786 Habilitation à diriger des recherches, Université Paris VI, Paris, France, 2003

787 Reynolds, C. A., Jackson, T. J., and Rawls, W. J.: Estimating soil water holding capacities by
788 linking the Food and Agriculture Organization Soil map of the world with global pedon

789 databases and continuous pedotransfer functions, *Water Resour. Res.*, 36(12): 3653–3662,
790 2000.

791 Romanou, A., Tselioudis, G., Zerefos, C. S., Clayson, C.-A., Curry, J. A., and Andersson, A.:
792 Evaporation–precipitation variability over the Mediterranean and the Black Seas from satellite
793 and reanalysis estimates, *J. Climate*, 23, 5268–5287, doi:10.1175/2010JCLI3525.1, 2010.

794 Rodríguez-Díaz, J. A., Knox, J. W., and Weatherhead, E. K.: Competing demands for irrigation
795 water: golf and agriculture in Spain, *Irrig. Drain.*, 56:541–549, 2007.

796 Santaren, D., Peylin, P., Viovy, N., and Ciais, P.: Optimizing a process-based ecosystem model
797 with eddy-covariance flux measurements: A pine forest in southern France, *Global*
798 *Biogeochem. Cy.*, 21, GB2013, doi: 10.1029/2006GB002834, 2007.

799 [Santaren, D., Peylin, P., Bacour, C., Ciais, P., and Longdoz, B.: Ecosystem model optimization](#)
800 [using in situ flux observations: benefit of Monte Carlo versus variational schemes and](#)
801 [analyses of the year-to-year model performances, *Biogeosciences*, 11, 7137-7158,](#)
802 [doi:10.5194/bg-11-7137-2014, 2014.](#)

803 Scherbakov, A. V. and Malakhova, V. V.: The Influence of Time Step Size on the Results of
804 Numerical Modeling of Global Ocean Climate, *Numerical Analysis and Applications*, 4(2),
805 175–187, 2011.

806 Szczypta, C., Decharme, B., Carrer, D., Calvet, J.-C., Lafont, S., Somot, S., Faroux, S., and Martin,
807 E.: Impact of precipitation and land biophysical variables on the simulated discharge of
808 European and Mediterranean rivers, *Hydrol. Earth Syst. Sci.*, 16, 3351-3370,
809 <https://doi.org/10.5194/hess-16-3351-2012>, 2012.

810 Sevault, F., Somot, S., Alias, A., Dubois, C., Lebeau-pin-Brossier, C., Nabat, P., Adloff, F., Déqué,
811 M., and Decharme, B.: A fully coupled Mediterranean regional climate system model: Design
812 and evaluation of the ocean component for the 1980-2012 period, *Tellus*, 66A, 23967,
813 doi:10.3402/tellusa.v66.23967, 2014.

814 Shaltout, M. and Omstedt, A.: Modelling the water and heat balances of the Mediterranean Sea
815 using a two-basin model and available meteorological, hydrological, and ocean data.
816 *Oceanologia*, 57:116–131, 2015.

817 Siam, M. S., and Eltahir Elfatih, A. B.: Climate change enhances interannual variability of the Nile
818 river flow, *Nat. Clim. Change*, doi: 10.1038/nclimate3273, 2017.

819 Sichangi, W.A, Wang, L., Yang, K., Chen, D., Wang, Z., Li, X., Zhou, J., Liu, W., and Kuria. D.:
820 Estimating continental river basin discharges using multiple remote sensing data sets, *Remote*
821 *Sens. Environ.*, 179: 36-53, <https://doi.org/10.1016/j.rse.2016.03.019>, 2016.

822 Struglia, M.V., Mariotti, A., and Filogrosso, A.: River discharge into the Mediterranean Sea:
823 climatology and aspects of the observed variability, *J. Clim.*, 17, 4740-4751,
824 doi: 10.1175/JCLI-3225.1, 2004.

825 Syed, T. H., Famiglietti, J. S., Chambers, D. P., Willis, J. K., and Hilburn, K.: Satellite-based
826 global-ocean mass balance estimates of interannual variability and emerging trends in
827 continental freshwater discharge, *Proc. Natl. Acad. Sci., USA*, 42, 17916–17921,
828 doi:10.1073/pnas.1003292107, 2010.

829 Tixeront, J.: Le bilan hydrologique de la Mer Noire et de la Mer Méditerranée, *Cahiers*
830 *Océanographiques*, 22(3), 227–237, 1970.

831 Vargas-Amelin, E. and Pindado, P.: The challenge of climate change in Spain: water resources,
832 agriculture and land, *J. Hydrol.*, 518, 243-249, doi:10.1016/j.jhydrol.2013.11.035, 2014.

833 Verri, G., Pinardi, N., Oddo, P., Ciliberti, S. A., and Coppini, G.: River runoff influences on the
834 Central Mediterranean Overturning Circulation, *Clim. Dynam.*, in press, 2017.

835 [Van Loon, A. F. and Van Lanen H. A. J.: Making the distinction between water scarcity and](#)
836 [drought using an observation-modeling framework, *Water Resour Res*, 49,](#)
837 [doi:10.1002/wrcr.20147, 2013.](#)

838 [Vinulcollu, R.K., Wood, E.F., Ferguson, C.R., Fisher, J.B.: Global estimates of evapotranspiration](#)
839 [for climate studies using multi-sensor remote sensing data: Evaluation of three process-based](#)
840 [approaches, *Remote sensing of environment*, 115\(3\):801-23, 2011.](#)

841 [Vorosmarty, C. J., Fekete B. M., and Tucker B. A.: Global River Discharge, 1807-1991, V. 1.1](#)
842 [\(RivDIS\). ORNL DAAC, Oak Ridge, Tennessee, USA.](#)
843 <https://doi.org/10.3334/ORNLDAAAC/199>, 1998.

844 Wang, A., Zeng, X., and Guo, D.: Estimates of global surface hydrology and heat fluxes from the
845 Community Land Model (CLM4.5) with four atmospheric forcing datasets, *J. Hydrometeorol.*,
846 17, 2493–2510, 2016.

847 Wang, Q., Wekerle, C., Danilov, S., Wang, X., and Jung, T.: A 4.5 km resolution Arctic Ocean
848 simulation with the global multi-resolution model FESOM1.4, *Geosci. Model Dev. Discuss.*,
849 <https://doi.org/10.5194/gmd-2017-136>, in review, 2017.

850 Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., and Viterbo P.: The WFDEI
851 meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim
852 reanalysis data, *Water Resour. Res.*, 50(9), 7505–7514, doi:10.1002/2014WR015638, 2014.
853 Zhou, X., Polcher, J., Yang, T., Hirabayashi, Y., and Nguyen-Quang, T.: Understanding the water
854 cycle over the upper Tarim basin: retrospect the estimated discharge bias to atmospheric
855 variables and model structure, *Hydrol. Earth Syst. Sci. Discuss.*, [https://doi.org/10.5194/hess-](https://doi.org/10.5194/hess-2018-88)
856 2018-88, in review, 2018.

857 **Figure captions:**

858 **Figure 1.** (a) The illustration of correcting river discharge (Q) simulation (simulation in blue solid
859 dot, observation in red star) by applying correction factors (x) to runoff and drainage over different
860 basins. The basin 1 and basin 2 are represented in yellow and blue, respectively. (b) The model
861 framework of the river discharge assimilation. The blue and red parts are run for ‘First Guess’ and
862 for assimilation, respectively.

863 **Figure 2.** The river network (blue lines) and the GRDC stations (solid dots represent the 27
864 qualified stations and the gray triangles represent unqualified stations) over the study domain.

865 **Figure 3.** (a) The variation of cost function J (unit: 1; logarithmic y-axis) with iterations for x_{prior_1}
866 ($x_{prior} = 1$, in blue) and for x_{prior_ref} ($x_{prior} = \text{pre-estimated-prior}$, in red). The iterations 6-15 are
867 enlarged in the window (normal y-axis). The $Norm_BIAS$ of optimized river discharge after 7
868 iterations for x_{prior_1} (b) and for x_{prior_ref} (c).

869 **Figure 4.** The set-up of assimilation experiments for n years ($n=10$, 1980-1989) and k iterations
870 ($k=10$) with m ($m=27$) correction factors (x) each year (x is different over years). (a) The i th year
871 (Y_i) optimization is initialized by the end of Y_{i-1} optimization; (b) the initial condition of Y_i
872 optimization is got by running Y_{i-1} optimization fed with the same x as Y_i ; (c) optimizing n years
873 together with one year spin-up at the beginning of n -year. The Y1SP0 and Y1SP1 divide the n -
874 year optimization into n 1-year optimization periods. The blue and red colors mean optimization
875 and spin-up simulations, respectively.

876 **Figure 5.** The river discharge simulations from 1980 to 1989 using WFDEI_GPCC (1st row),
877 WFDEI_CRU (2nd row) and CRU_NCEP (3rd row) forcing. Left: the correlation coefficient of
878 river discharge between observations and simulations; Right: the $Norm_BIAS$ of simulated river
879 discharge.

880 **Figure 6.** The optimization results from 1980 to 1989 using the three methods (1st row: Y1SP1;
881 2nd row: Y1SP0; 3rd row: Y10C) forced by WFDEI_GPCC. Left: the optimized correction factor
882 x ; Middle: the correlation coefficient of river discharge between observations and optimizations;
883 Right: the $Norm_BIAS$ of optimized river discharge.

884 **Figure 7.** The annual cycles of river discharge for ‘First Guess’ (FG) forced by WFDEI-GPCC
885 (black), Y1SP1 (blue), Y1SP0 (green), Y10C (yellow) and GRDC observations (red) over the

886 Alcala Del Rio station (-5.98°W, 37.52°N) on the Guadalquivir river. The dotted lines mean the
887 trend.

888 **Figure 8.** The correction factor x obtained from Y1SP0 forced by (a) WFDEI_CRU, (b)
889 CRU_NCEP, (c) WFDEI_GPCC, and (d) the ‘Uncertainty’ (defined by Eq. 10) of x by different
890 forcing. All values are averaged over 1980-1989.

891 **Figure 9.** The evaporation (E , in mm/d) before assimilation (1st line), change of evaporation (dE ,
892 in mm/d) after and before assimilation (2nd line), and the ratio of dE and runoff + drainage (3rd line)
893 for forcing WFDEI-GPCC (1st column), WFDEI-CRU (2nd column), CRU-NCEP (3rd column),
894 and the ‘Uncertainty’ (defined by Eq. 10) in different forcing (4th column) averaged from 1980 to
895 1989.

896 **Figure 10.** The optimization results by different atmospheric forcing (WFDEI-GPCC in black,
897 WFDEI-CRU in green, and CRU-NCEP in blue) over the Puente De Palmas station on Guadiana
898 River (a-d, -6.97°W, 38.88°N; 48515 km²) and over the Masia De Pompo station on the Jucar river
899 (e-h, -0.65°W, 39.15°N; 17876 km²): (a, d) annual river discharges; (b, e) runoff coefficient; (e, f)
900 optimized correction factor x for the simulated/assimilated river discharge (First Guess in dark
901 color, Y1SP0 in light color) with respect to GRDC observations (in red) from 1980 to 1989.

902 **Figure 11.** The inter-annual variation of correction factor x ($\frac{\sigma(x)}{\bar{x}}$; a, d, g), simulated river discharge
903 without assimilation ($\frac{\sigma(Q_{sim})}{Q_{sim}}$; b, e, h) and optimized river discharge ($\frac{\sigma(Q_{opt})}{Q_{opt}}$; c, f, i) for
904 Y1SP0_WFDEIGPCC (1st row), Y1SP0_WFDEICRU (2nd row) and Y1SP0_CRUNCEP (3rd row)
905 averaged over 1980-1989.

906 **Figure 12.** Comparison of evaporation (E , in mm/d, 1st line) between GLEAM (v3.1) and FG (First
907 Guess), as well as E (2nd line), precipitation (P , in mm/d, 3rd line), $P-E$ (in mm/d, 4th line) and $P-E$
908 (relative value between 0-1, 5th line) between GLEAM (v3.1) and assimilated values using
909 different forcing (1st column: WFDEI-GPCC; 2nd column: WFDEI-CRU; 3rd column: CRU-NCEP;
910 4th column: ‘Uncertainty’ (defined by Eq. 10) of using different forcing) averaged from 1980 to
911 1989.

Table 1. The assimilation and simulation experiments

Name	Atmospheric Forcing	Method
FG(WFDEIG)	WFDEI_GPCC	No assimilation
FG(WFDEIC)	WFDEI_CRU	No assimilation
FG(CRUN)	CRU_NCEP	No assimilation
Y1SP0(WFDEIG)	WFDEI_GPCC	Y1SP0 assimilation
Y1SP1(WFDEIG)	WFDEI_GPCC	Y1SP1 assimilation
Y10C(WFDEIG)	WFDEI_GPCC	Y10C assimilation
Y1SP0(WFDEIC)	WFDEI_CRU	Y1SP0 assimilation
Y1SP0(CRUN)	CRU_NCEP	Y1SP0 assimilation

Note: All runs are from 1980 to 1989 with 0.5° spatial resolution; FG stands for ‘[First Guess](#)’.

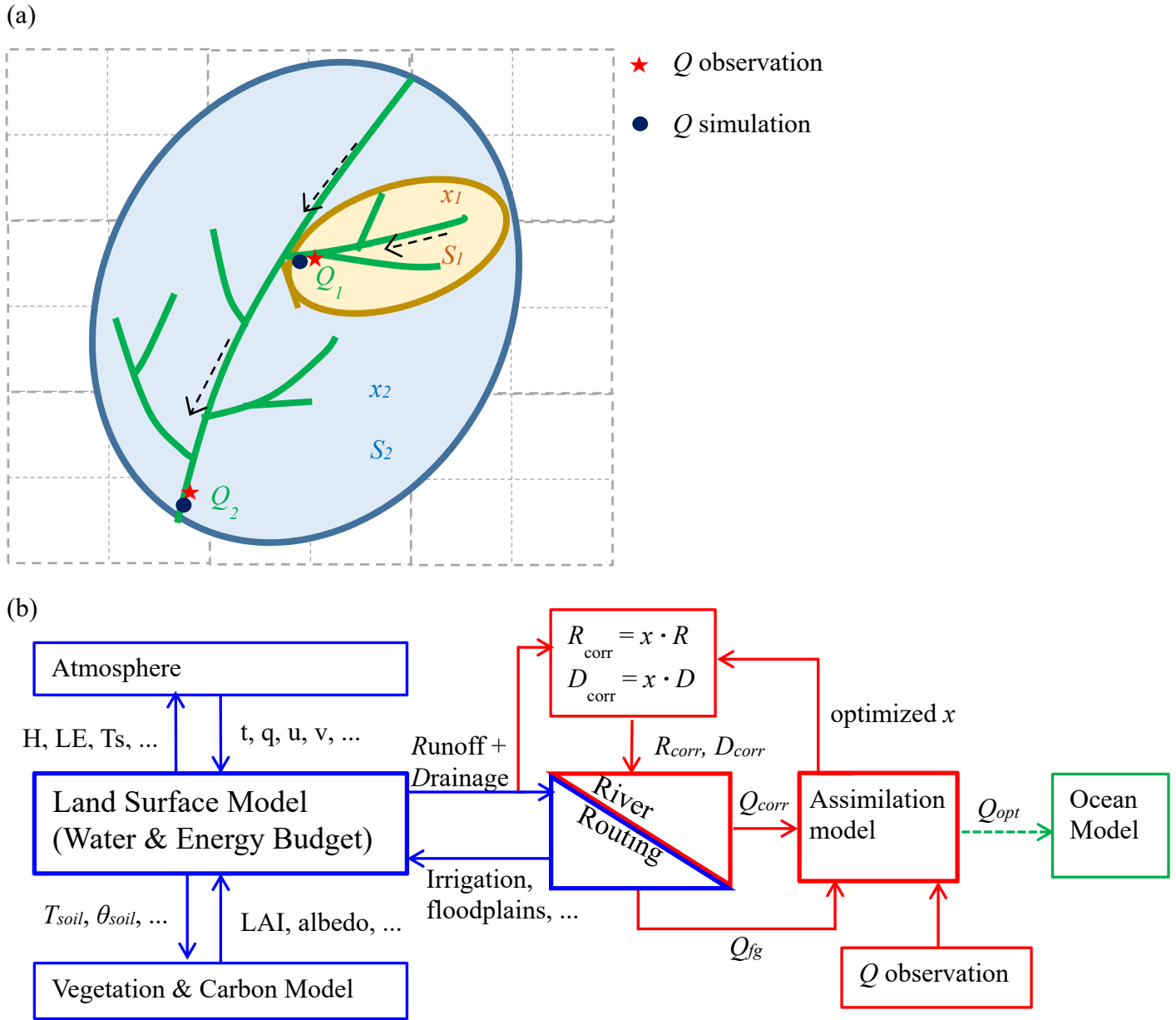


Figure 1. (a) The illustration of correcting river discharge (Q) simulation (simulation in blue solid dot, observation in red star) by applying correction factors (x) to runoff and drainage over different basins. The basin 1 and basin 2 are represented in yellow and blue, respectively. (b) The model framework of the river discharge assimilation. The blue and red parts are run for ‘First Guess’ and for assimilation, respectively.

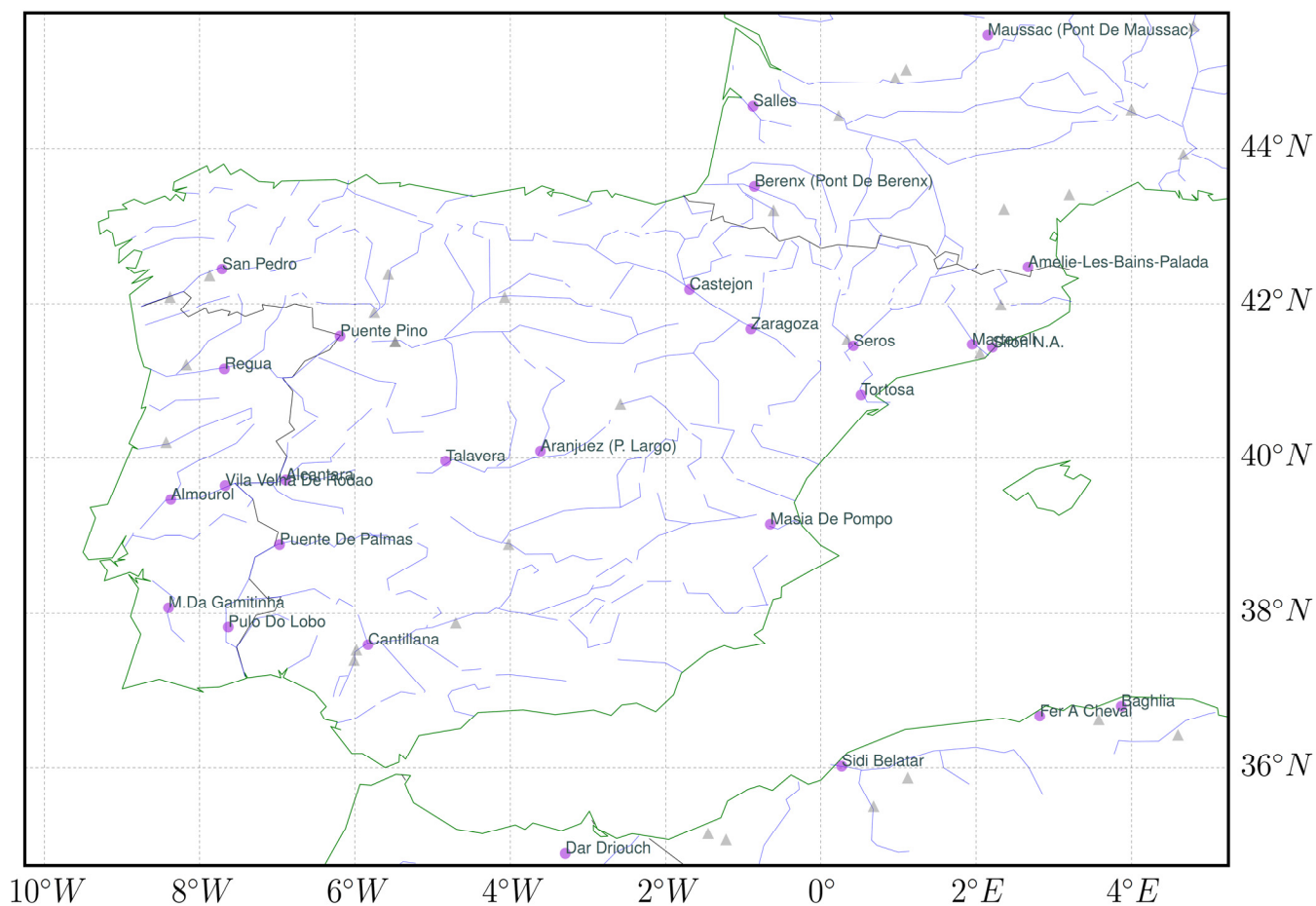


Figure 2. The river network (blue lines) and the GRDC stations (solid dots represent the 27 qualified stations and the gray triangles represent unqualified stations) over the study domain.

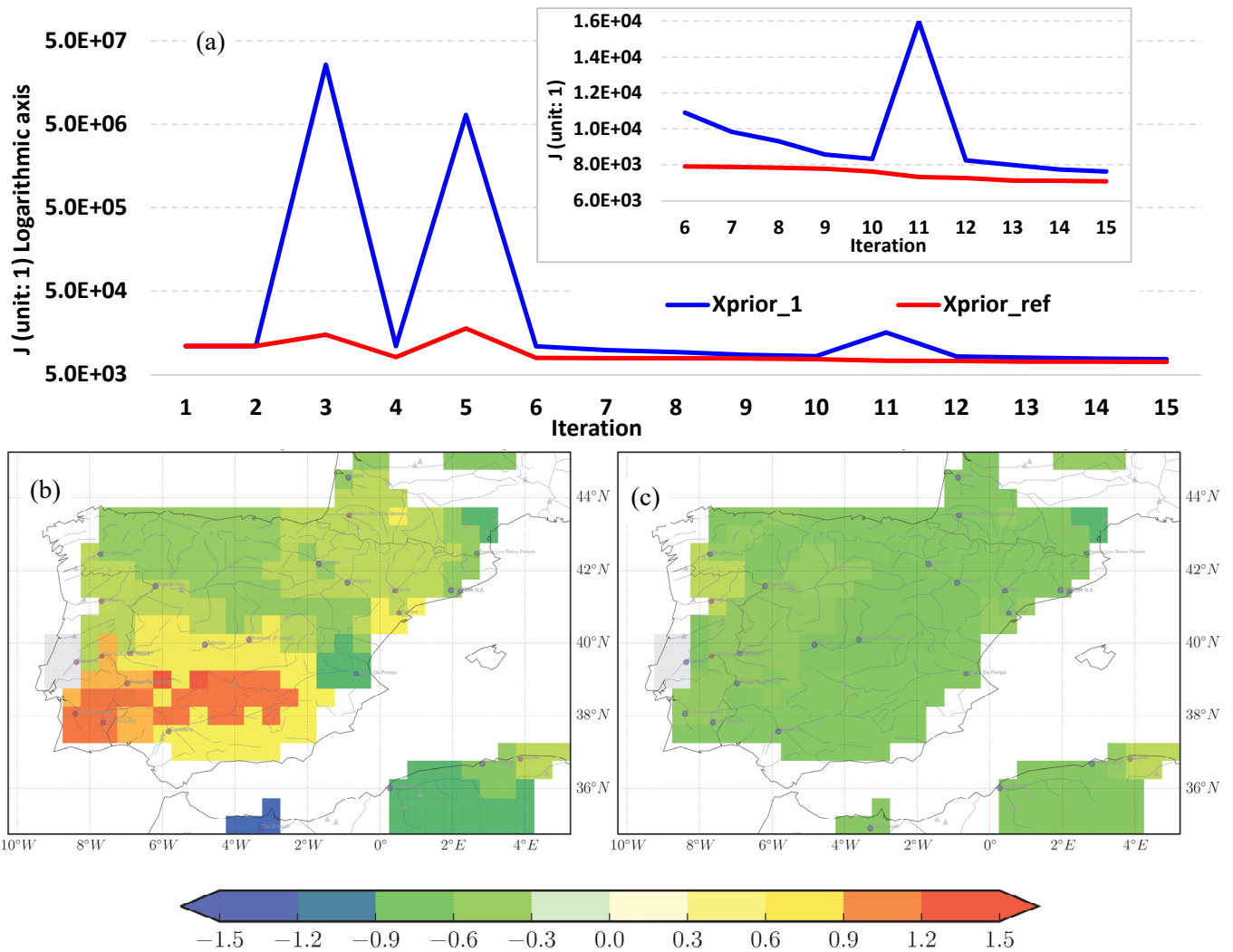
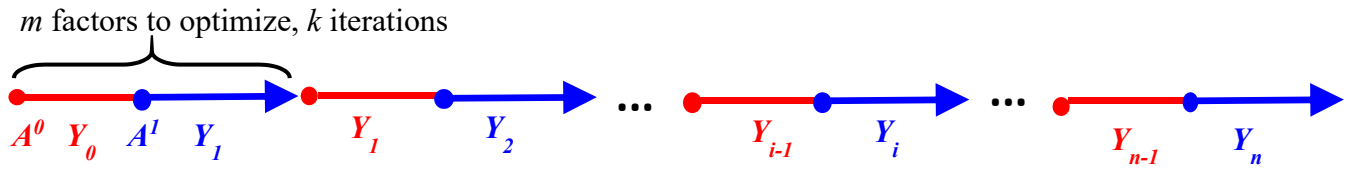
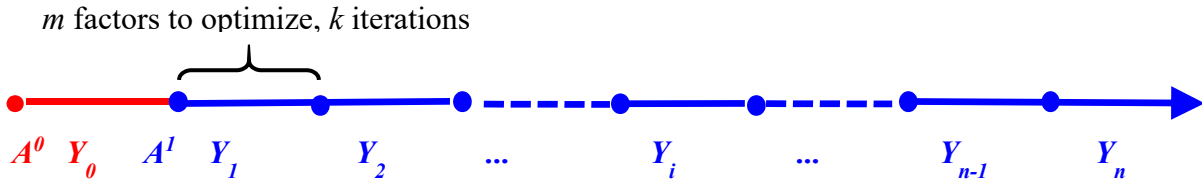


Figure 3. (a) The variation of cost function J (unit: 1; logarithmic y-axis) with iterations for x_{prior_1} ($x_{prior} = 1$, in blue) and for x_{prior_ref} ($x_{prior} = \text{pre-estimated-prior}$, in red). The iterations 6-15 are enlarged in the window (normal y-axis). The $Norm_BIAS$ of optimized river discharge after 7 iterations for x_{prior_1} (b) and for x_{prior_ref} (c).

(a) Y1SP1



(b) Y1SP0



(c) Y10C

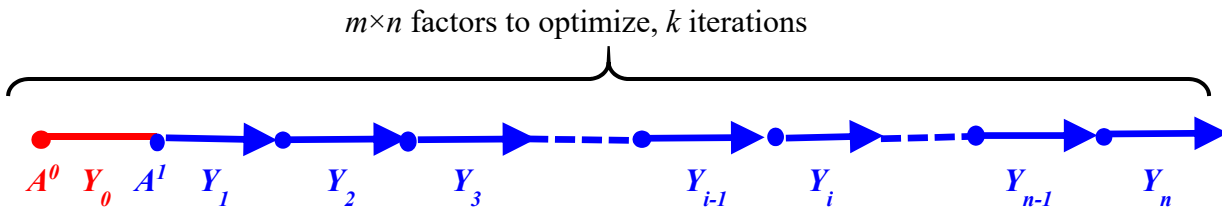


Figure 4. The set-up of assimilation experiments for n years ($n=10$, 1980-1989) and k iterations ($k=10$) with m ($m=27$) correction factors (x) each year (x is different over years). (a) The i th year (Y_i) optimization is initialized by the end of Y_{i-1} optimization; (b) the initial condition of Y_i optimization is got by running Y_{i-1} optimization fed with the same x as Y_i ; (c) optimizing n years together with one year spin-up at the beginning of n -year. The Y1SP0 and Y1SP1 divide the n -year optimization into n 1-year optimization periods. The blue and red colors mean optimization and spin-up simulations, respectively.

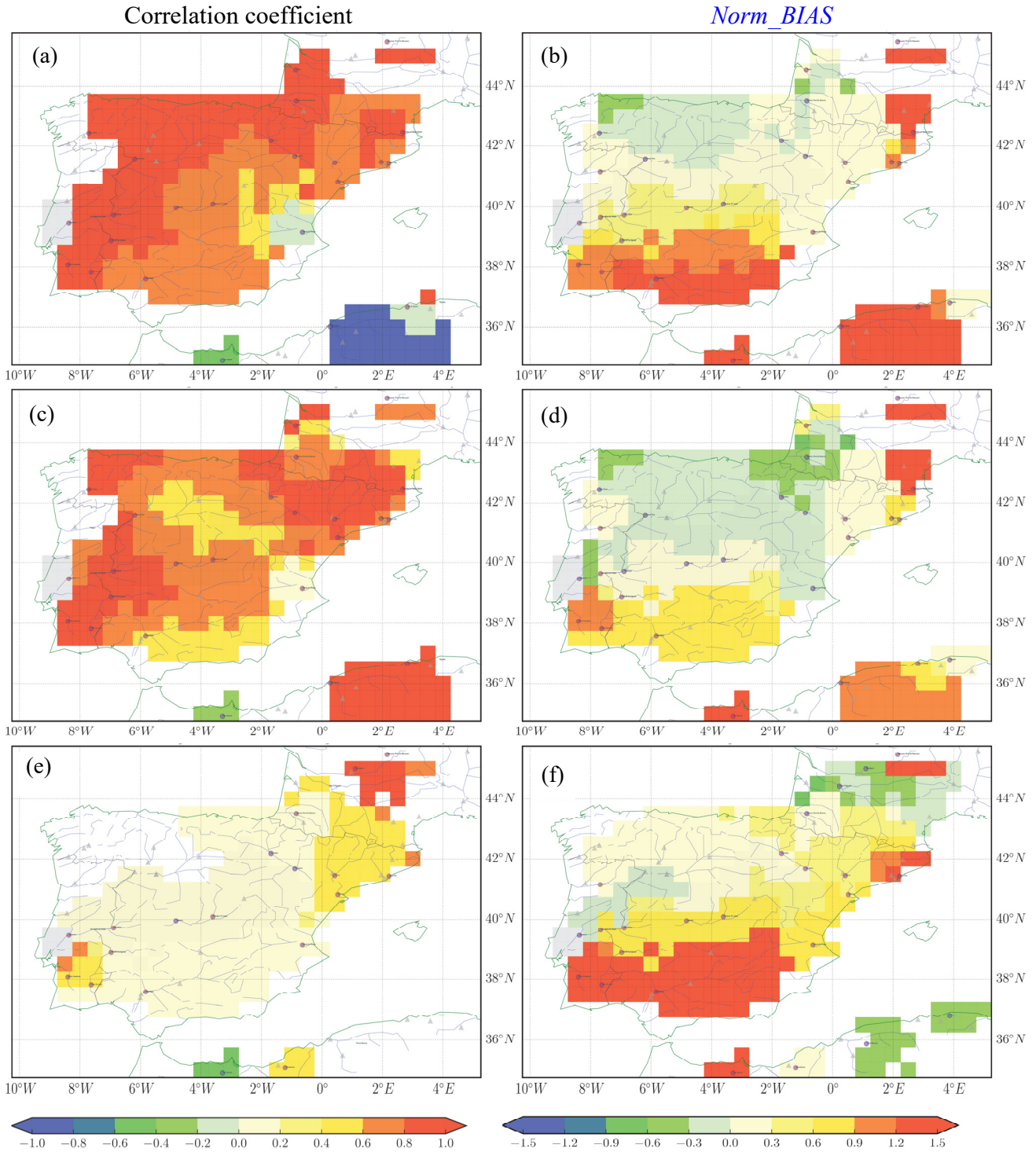


Figure 5. The river discharge simulations from 1980 to 1989 using WFDEI_GPCC (1st row), WFDEI_CRU (2nd row) and CRU_NCEP (3rd row) forcing. Left: the correlation coefficient of river discharge between observations and simulations; Right: the *Norm_BIAS* of simulated river discharge.

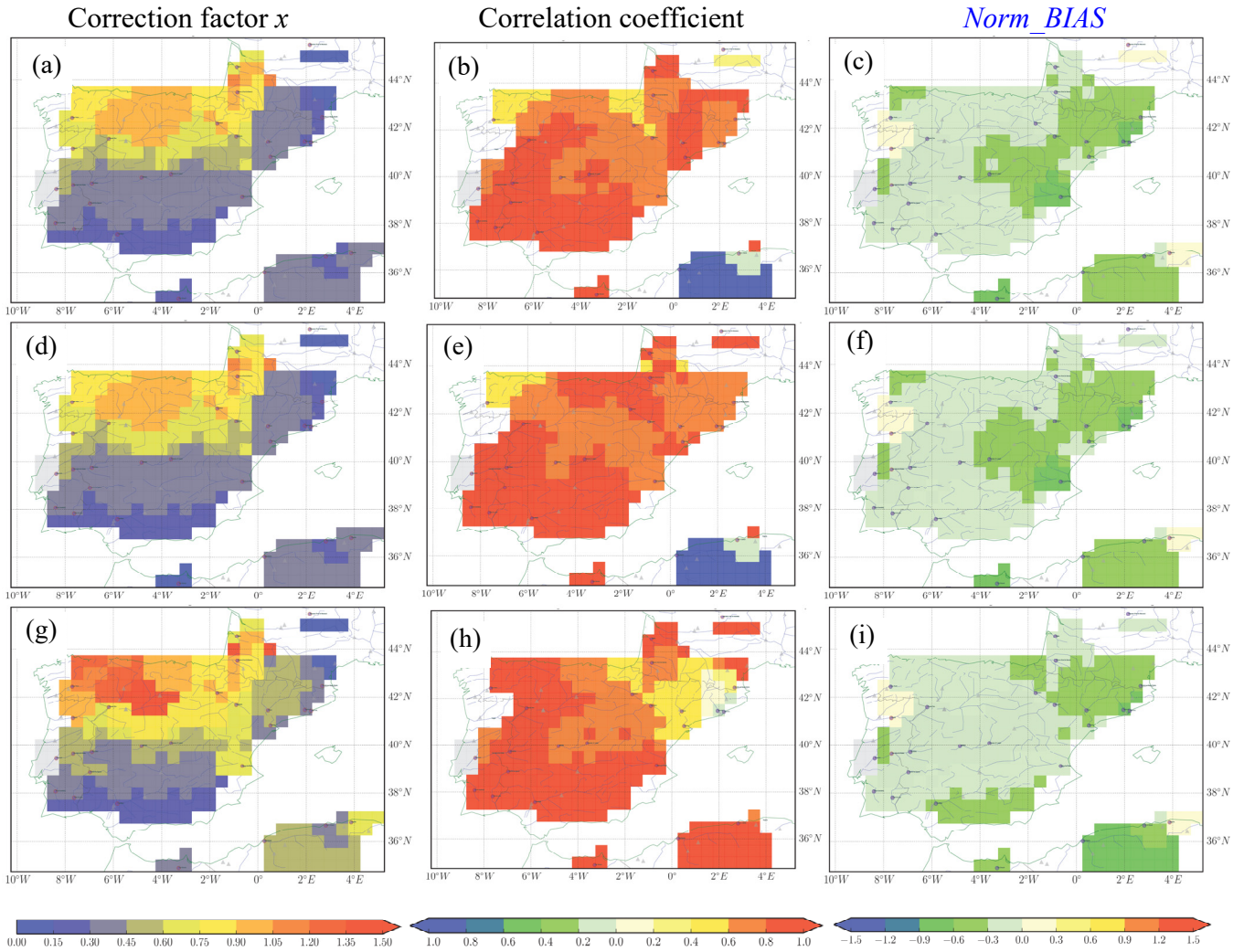


Figure 6. The optimization results from 1980 to 1989 using the three methods (1st row: Y1SP1; 2nd row: Y1SP0; 3rd row: Y10C) forced by WFDEI_GPCC. Left: the optimized correction factor x ; Middle: the correlation coefficient of river discharge between observations and optimizations; Right: the *Norm_BIAS* of optimized river discharge.

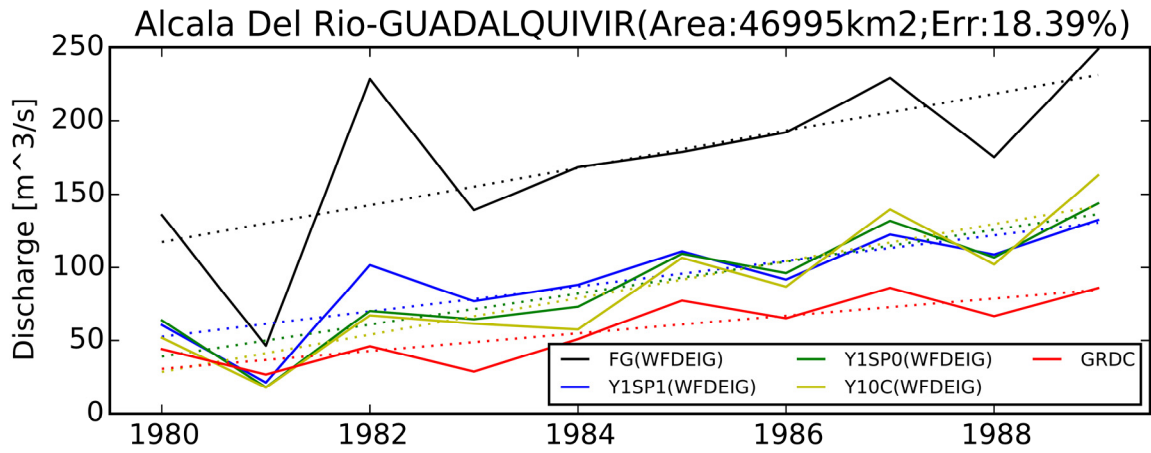


Figure 7. The annual cycles of river discharge for ‘First Guess’ (FG) forced by WFDEI-GPCC (black), Y1SP1 (blue), Y1SP0 (green), Y10C (yellow) and GRDC observations (red) over the Alcala Del Rio station (-5.98°W, 37.52°N) on the Guadalquivir river. The dotted lines mean the trend.

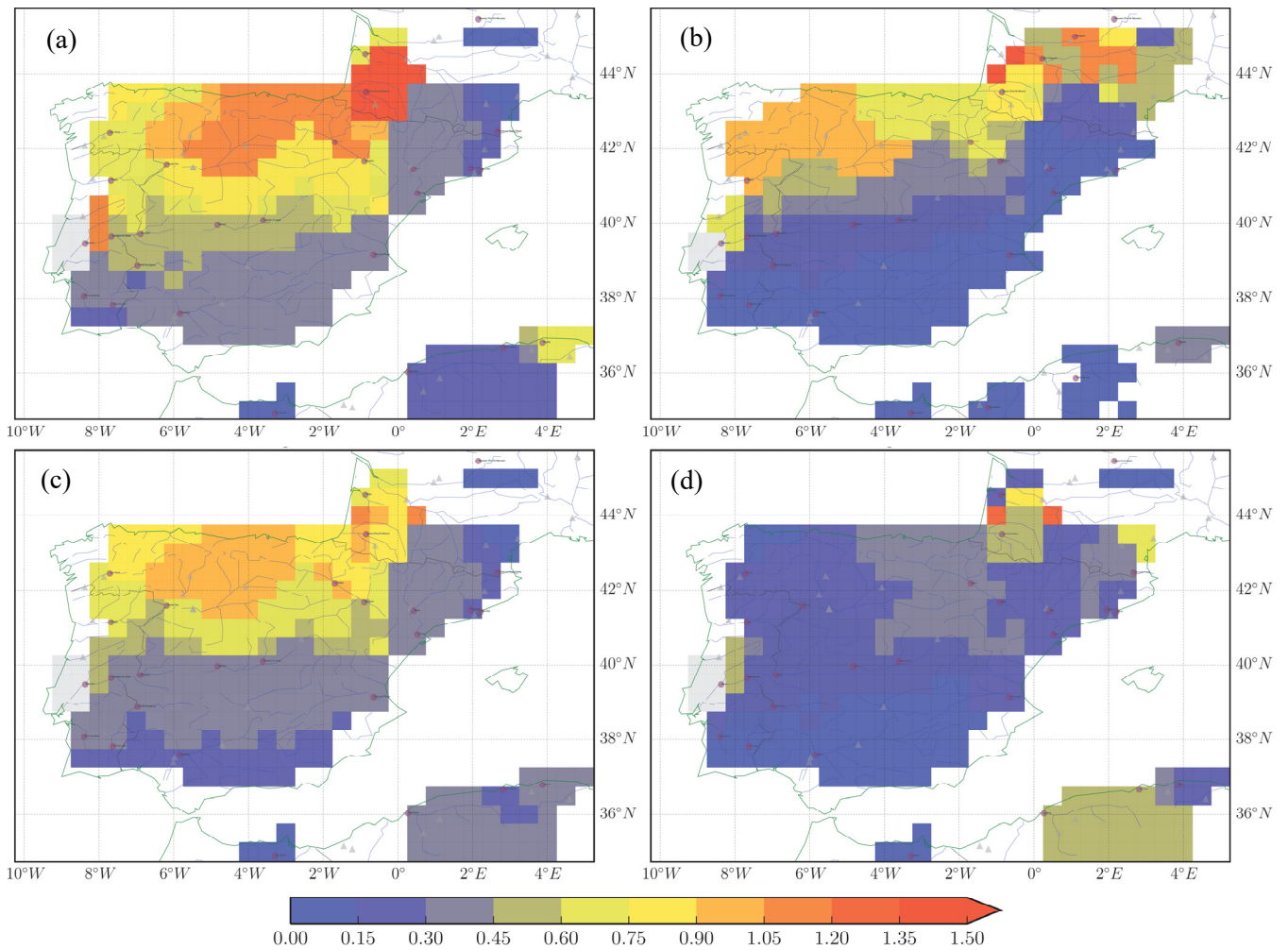


Figure 8. The correction factor x obtained from YISP0 forced by (a) WFDEI_CRU, (b) CRU_NCEP, (c) WFDEI_GPCC, and (d) the ‘Uncertainty’ (defined by Eq. 10) of x by different forcing. All values are averaged over 1980-1989.

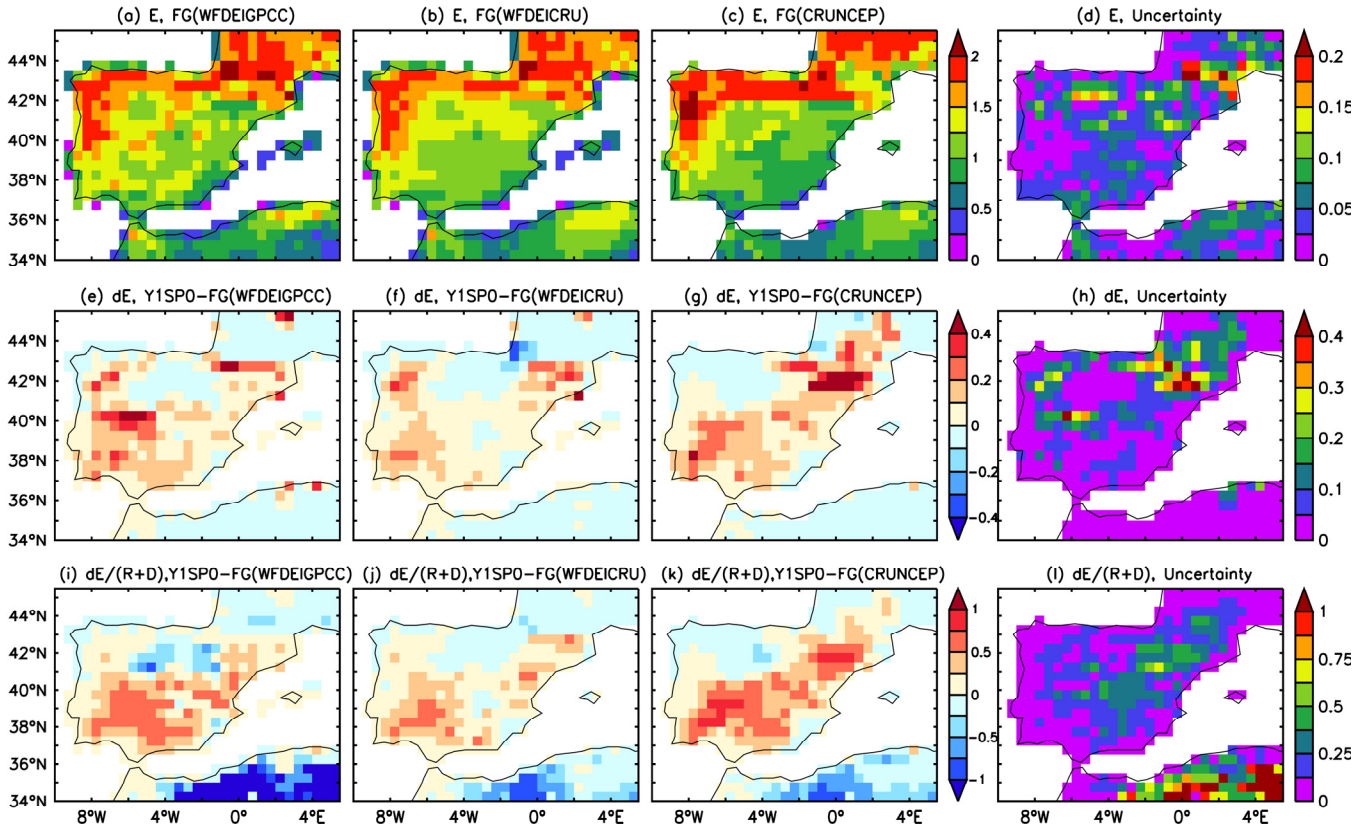


Figure 9. The evaporation (E , in mm/d) before assimilation (1st line), change of evaporation (dE , in mm/d) after and before assimilation (2nd line), and the ratio of dE and runoff + drainage (3rd line) for forcing WFDEI-GPCC (1st column), WFDEI-CRU (2nd column), CRU-NCEP (3rd column), and the ‘Uncertainty’ (defined by Eq. 10) in different forcing (4th column) averaged from 1980 to 1989.

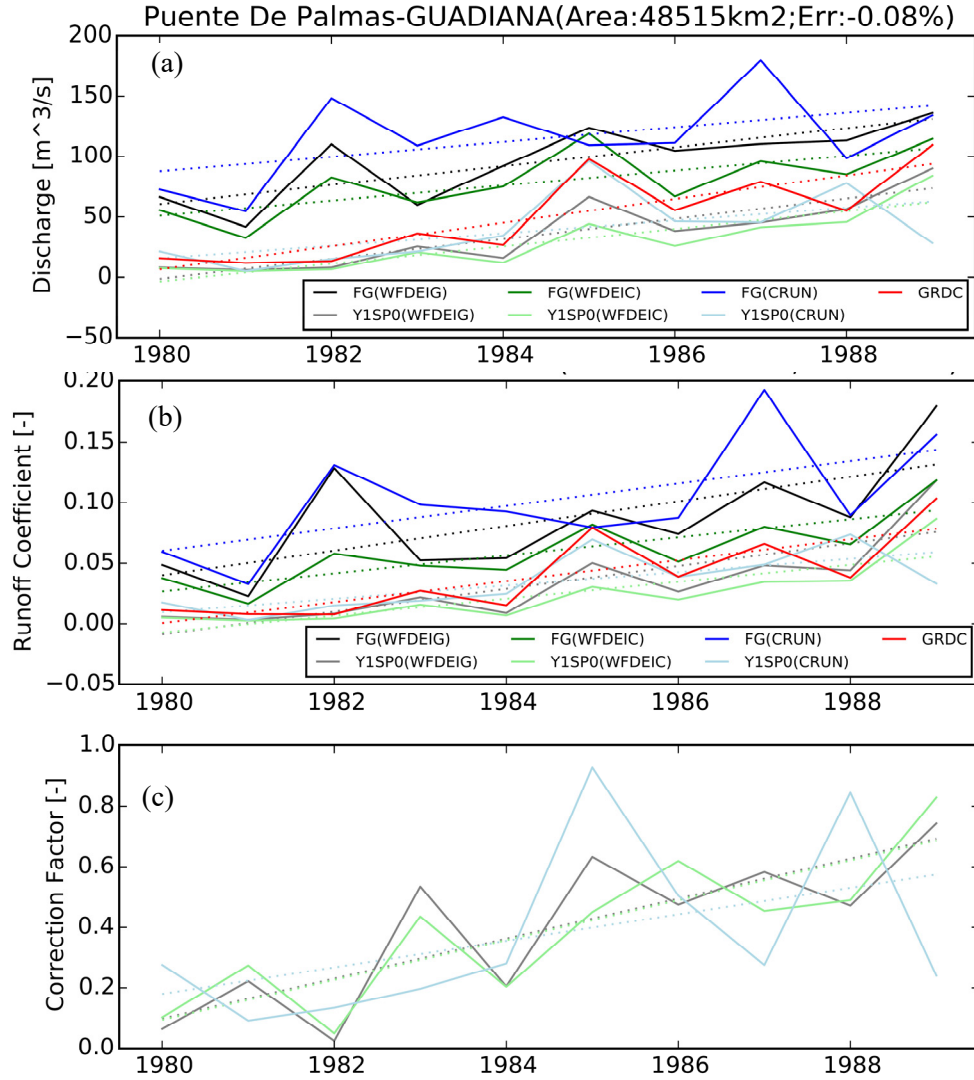


Figure 10. The optimization results by different atmospheric forcing (WFDEI-GPCC in black, WFDEI-CRU in green, and CRU-NCEP in blue) over the Puente De Palmas station on Guadiana River (a-d, -6.97°W, 38.88°N; 48515 km²) and over the Masia De Pompo station on the Jucar river (e-h, -0.65°W, 39.15°N; 17876 km²): (a, d) annual river discharges; (b, e) runoff coefficient; (e, f) optimized correction factor x for the simulated/assimilated river discharge (First Guess - FG in dark color, Y1SP0 in light color) with respect to GRDC observations (in red) from 1980 to 1989.

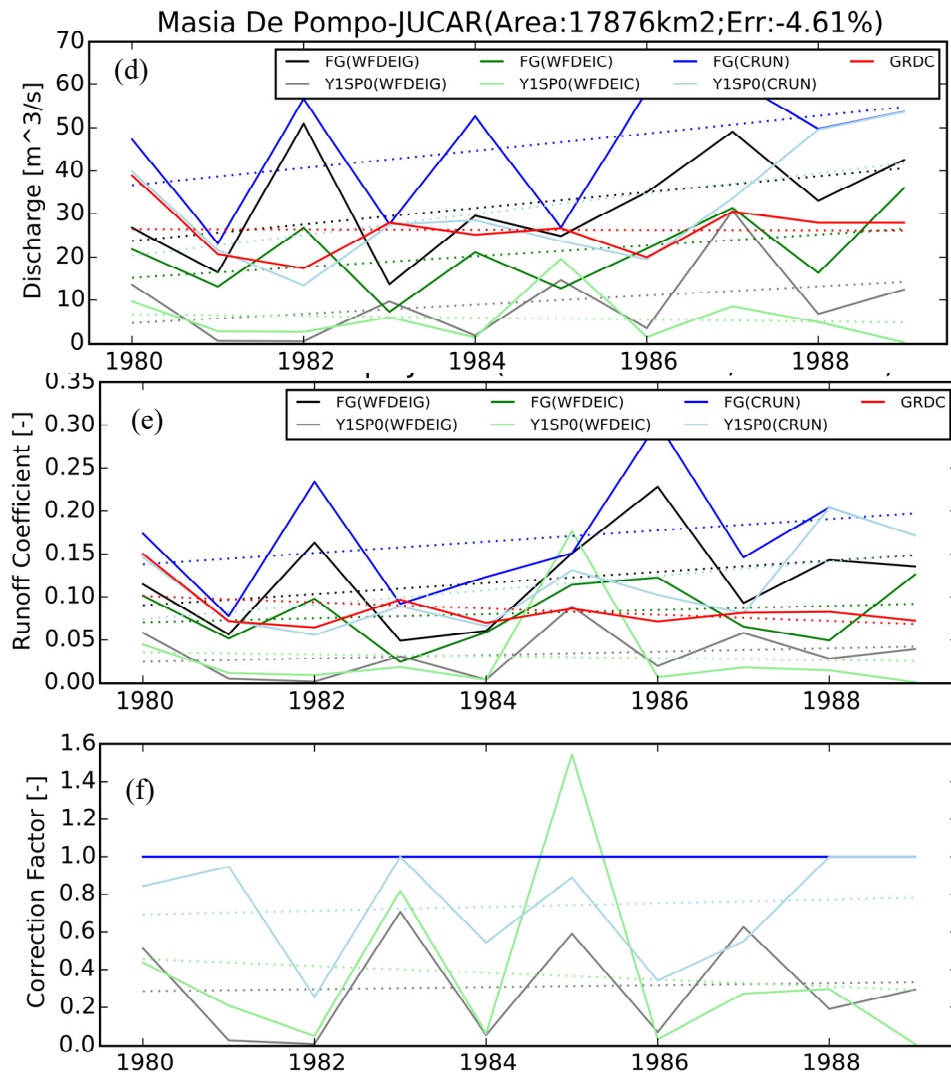


Figure 10. Continued.

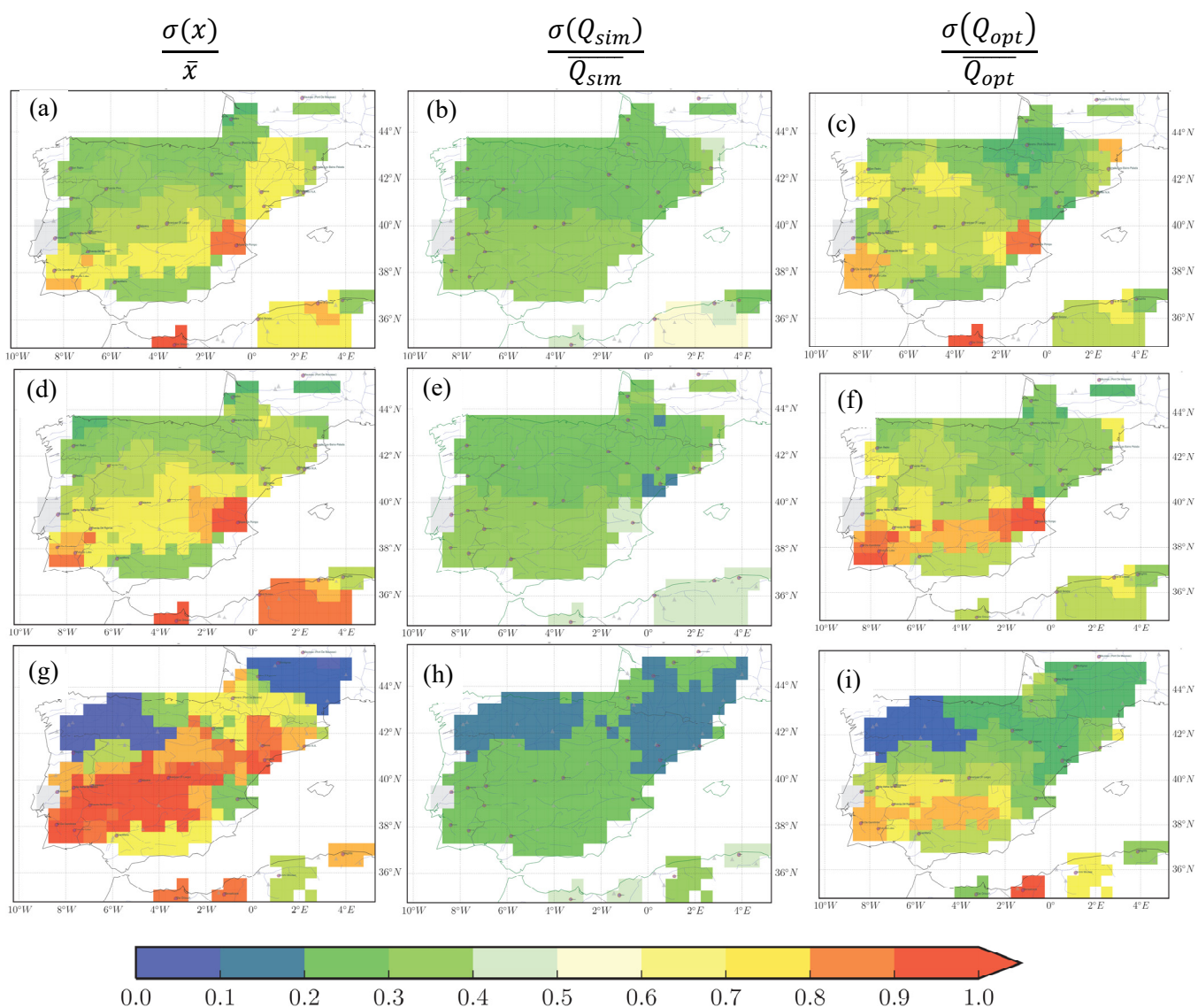


Figure 11. The inter-annual variation of correction factor x ($\frac{\sigma(x)}{\bar{x}}$; a, d, g), simulated river discharge without assimilation ($\frac{\sigma(Q_{sim})}{Q_{sim}}$; b, e, h) and optimized river discharge ($\frac{\sigma(Q_{opt})}{Q_{opt}}$; c, f, i) for Y1SP0_WFDEIGPCC (1st row), Y1SP0_WFDEICRU (2nd row) and Y1SP0_CRUNCEP (3rd row) averaged over 1980-1989.

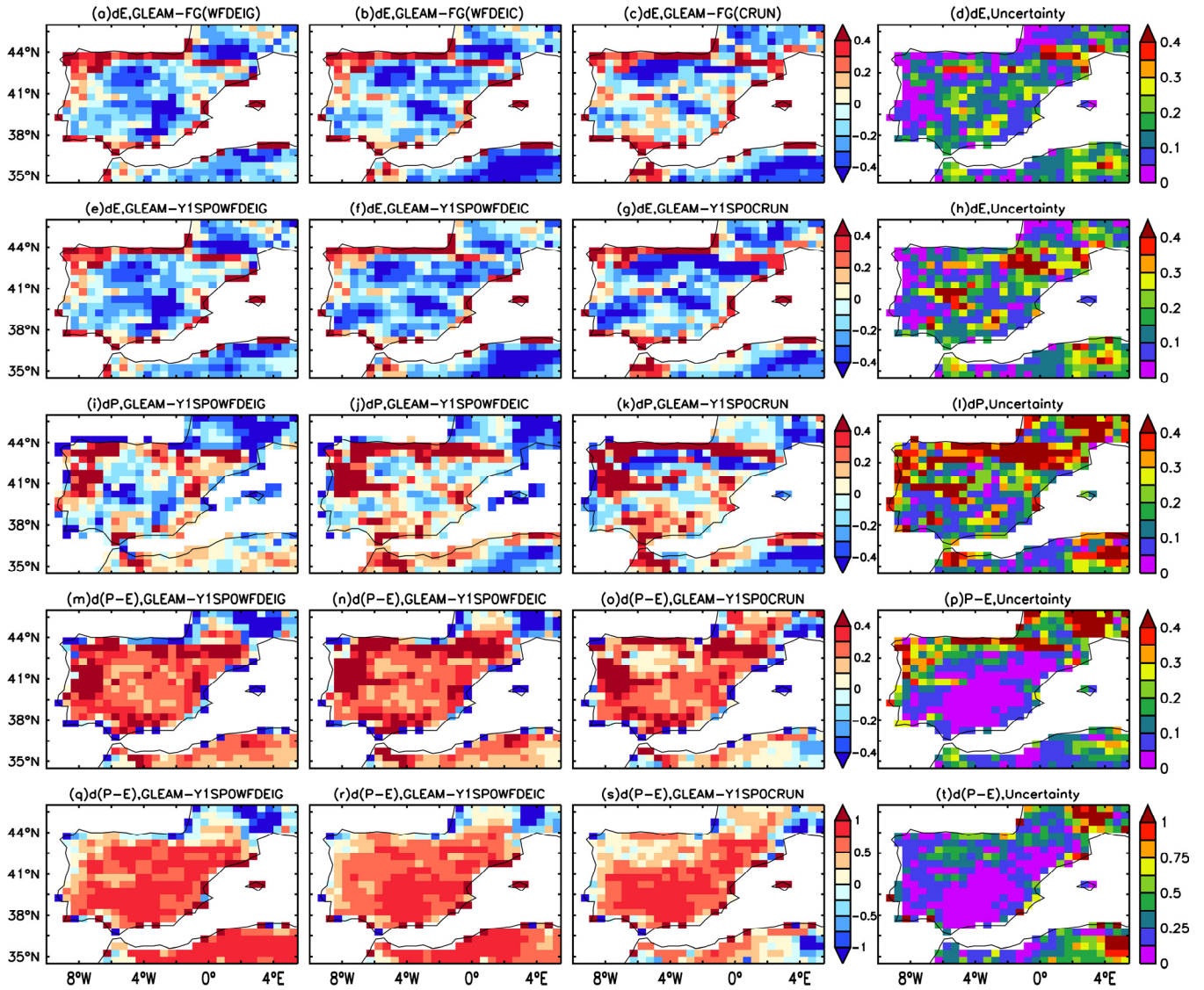


Figure 12. Comparison of evaporation (E , in mm/d, 1st line) between GLEAM (v3.1) and FG (First Guess), as well as E (2nd line), precipitation (P , in mm/d, 3rd line), $P-E$ (in mm/d, 4th line) and $P-E$ (relative value between 0-1, 5th line) between GLEAM (v3.1) and assimilated values using different forcing (1st column: WFDEI-GPCC; 2nd column: WFDEI-CRU; 3rd column: CRU-NCEP; 4th column: ‘Uncertainty’ (defined by Eq. 10) of using different forcing) averaged from 1980 to 1989.

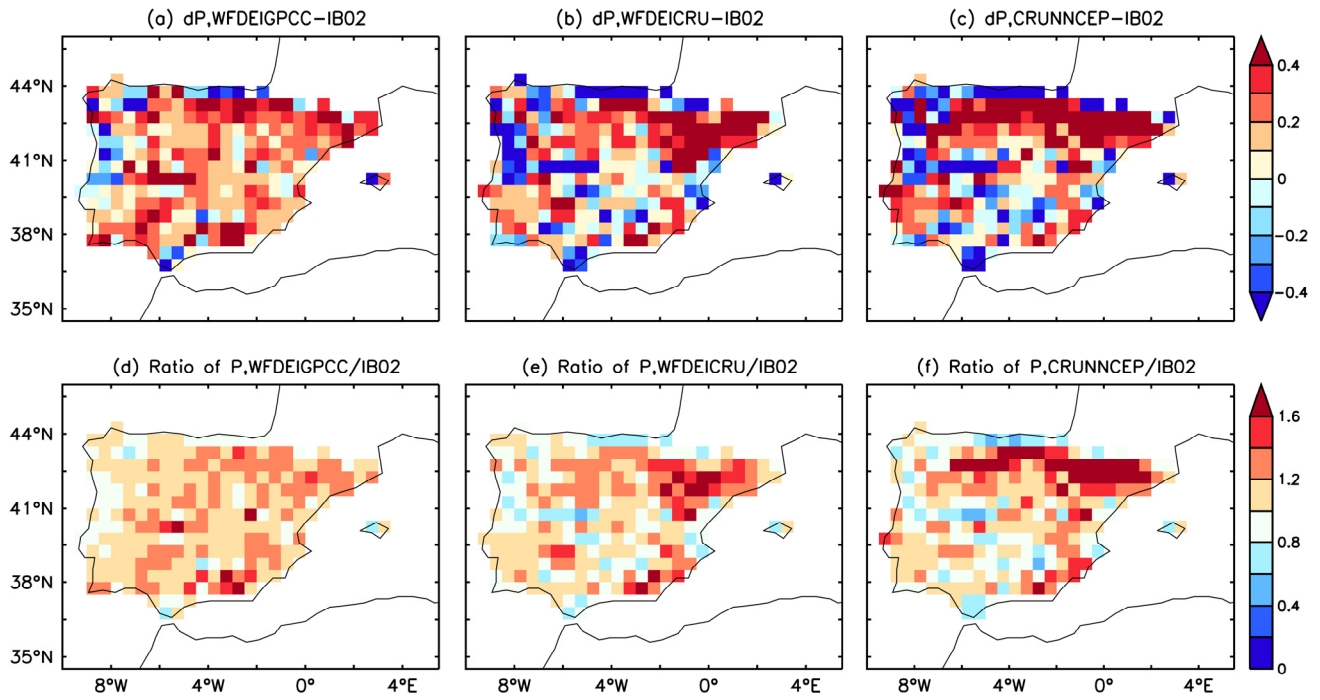


Figure S1. Comparison of precipitation (P , in mm/d) between IB02 and that used in the assimilation (a and d: WFDEI-GPCC; b and e: WFDEI-CRU; c and f: CRUNCEP) averaged from 1980 to 1989: 1st row for difference; 2nd row for ratio.