UNIVERSIDAD DEL BÍO-BÍO

Interactive Discussion: Author Response to Reviewer 01

SPatial Efficiency And Kmoments (SPEAK): Evaluating Spatial Consistency in (Semi)Distributed Rainfall–Runoff Models

Matías Moreno, Pablo Mendoza, Eduardo Muñoz-Castro, Mauricio Zambrano-Bigiarini, and Alonso Pizarro
HESS discussions/ EGUsphere, doi: 10.5194/egusphere-2026-2912

RC: ***Reviewer Comment***, AR: Author Response, □ Manuscript text

Dear Reviewer 01 (Anonymous Referee),

Thank you for taking the time and effort to review this manuscript. Surely, incorporating your suggestions will improve the manuscript's readability and overall quality. Please find our responses to your comments below. These should be considered as preliminary (part of the interactive discussion) since the actual implementation of changes depends on the editorial decision. Below, we list reviewer's comments verbatim in bold, followed by responses to these comments in blue.

Thank you again for your helpful feedback!

Kind regards,

Alonso Pizarro (on behalf of all co-authors)

RC 1: ***The paper is well-structured and proposes a novel approach to improve SPAEF metric. I have the following comments:***

-Did authors apply sensitivity analysis before model calibration? I couldnt find the details. Apparently all TUW parameters are included in the spatial calibration.
A robust calibration framework startes with a sensitivty analysis to reduce the parameter search space i.e. positively affecting the model runs and convergence to globally optimum metrics.

AR 1: Thank you for this comment. No parameter sensitivity analysis was performed before calibration, and all 15 TUW parameters were allowed to vary within the physically plausible bounds reported in Table 3 (the TUW model has a relatively low computational burden, so exploring the 15-parameter space was feasible). This was a deliberate methodological decision, since the purpose of the study is to compare SPEAK, SPAEF, and WSPAEF over an identical parameter space. Note that a sensitivity-based selection of parameters for each calibration metric has the potential to make the calibration experiments non-equivalent, since the suite of parameters may vary with the objective function. Moreover, hydroPSO (used R package) has previously been benchmarked against alternative optimisation algorithms and applied successfully to hydrological calibration problems. See also AR8, Reviewer 3.

RC 2: ***Line 250: "Fully distributed model configuration: Model inputs are provided at the grid-cell (0.05° x 0.05°) level across each catchment with uniform-in-space parameters' values (i.e., not depending on the spatial dimension)." I think the it would be excellent if the authors could apply pedotransfer functions to full distributed version of the model. Some models like mHM have this setting to gether with multi-parameter regionalization approach that help to create robust patterns instead of weak patterns due to uniform parameter values.***

AR 2: We decided to use spatially uniform parameter sets to simplify the calibration experiments, obtaining spatial variability in model states and fluxes only due to spatial heterogeneities in meteorological fields. In the revised manuscript, we will correct the model description, rename the 'fully distributed' case, and remove process claims involving spatial soil, vegetation, or land-use controls. The following sentences will be added to the main text:

a) Fully distributed model configuration will be renamed as 'gridded-forcing configuration':

> … In the gridded-forcing configuration, precipitation, temperature, and PET were supplied on the 0.10° spatial resolution (i.e., native GLEAM resolution), while one calibrated TUW parameter vector was applied uniformly to all grid cells within each catchment. Thus, this configuration represents spatial variability induced by the meteorological forcings and model states, but not spatial heterogeneity in soil-, vegetation-, or land-use-dependent parameters. Testing SPEAK in a model with pedotransfer functions or parameter regionalisation remains an important next step.…

b) Discussion section with related information and new research paths:

> … Future research will attempt to test the potential of SPEAK using a spatial parameter configuration that considers pedotransfer functions or other spatial regularization strategies.…

RC 3: ***Line 490: "Furthermore, the TUW model employed spatially uniform parameter values, potentially limiting its ability to represent local heterogeneity".***
Good that the authors mention this limitation in the text.

AR 3: Thank you. We will retain this limitation and make it explicit in the model description, not only in the final discussion. We will also narrow the interpretation so that spatial patterns are attributed only to mechanisms represented by the experiment. See also AR 2.

RC 4: **Line 65: "In recent years, the Spatial Efficiency (SPAEF; Koch et al., 2018)". In that paper the authors state that: "Following the multiple-component idea of KGE we present a novel spatial performance metric denoted SPAtial EFficiency (SPAEF), which was originally proposed by Demirel et al. (2018a, b)." https://gmd.copernicus.org/articles/11/1873/2018/**
This sentence can be helpful to find the origin of the metric.

AR 4: We agree and thank the reviewer for identifying the original attribution. We will revise the introduction and add the two Demirel et al. references cited by Koch et al. (2018):

> …The SPAtial EFficiency metric (SPAEF) was originally introduced and applied by Demirel et al. (2018a,b), and was subsequently tested by Koch et al. (2018)…

Reference:

Demirel, M. C., Mai, J., Mendiguren, G., Koch, J., Samaniego, L., and Stisen, S.: Combining satellite data and appropriate objective functions for improved spatial pattern performance of a distributed hydrologic model, Hydrol. Earth Syst. Sci., 22, 1299–1315, https://doi.org/10.5194/hess-22-1299-2018, 2018a.

Demirel, M. C., Stisen, S., and Koch, J.: SPAEF: SPAtial EFficiency, https://doi.org/10.5281/ZENODO.1158890, 2018b.

Koch, J., Demirel, M. C., and Stisen, S.: The SPAtial EFficiency metric (SPAEF): multiple-component evaluation of spatial patterns for optimization of hydrological models, Geosci. Model Dev., 11, 1873–1886, https://doi.org/10.5194/gmd-11-1873-2018, 2018.

RC 5: Sample size (number of grid cells) seems small in Fig4-5 and 6 as compared to a basin with 100x100 grids for example.
The reader can be cruious why the authors selected small catchments? Or why didnt they give the gridded maps of continental Chile (99 near-natural catchments)?

AR 5: In this study, the original catchment selection was conducted based on streamflow data completeness and near-naturalness criteria (see section '3.1 Study area and data', page 7, lines 183 – 191). Nevertheless, we agree that the number of native GLEAM pixels, rather than the nominal number of 0.05° model grid cells, determines the effective spatial information available for evaluating spatial patterns. As a consequence, the new catchment selection criteria apply a size filter to the 99 original catchments, retaining only those with 15 or more pixels at 0.10° spatial resolution. The latter implies reducing the number of catchments from 99 to 43.

We already recalibrated the TUW model at the native spatial resolution of GLEAM (i.e., 0.10°). Three objective functions were considered (see also AR 1 from Anonymous Reviewer 2, and AR 3 from Reviewer 3 – Juraj Parajka): i) SPEAK; ii) SPAEF; and, iii) WSPAEF. Some results as follows:

a) Table R1.1 shows the 43 catchments with 15 or more pixels at a spatial resolution of 0.10°:

Table R1.1. Selected catchments for analysis at 0.10° spatial resolution (i.e., the native GLEAM resolution).

Catchment ID	N° of pixels at 0.10° spatial resolution
1310002	18
2101001	30
3414001	29
3430003	93
4314002	22
4513001	20
4523002	47
4703002	20
7112001	16
8117005	18
8123001	16
8124002	20
8132001	27
8135002	66
8317001	104
8334001	150
8342001	22

9123001	24
9129002	46
9135001	29
9140001	81
9404001	25
9414001	22
9420001	40
9437002	110
10111001	72
10134001	30
10304001	30
10328001	39
10340001	19
10356001	37
10364001	80
11141001	18
11143001	33
11302001	33
11308001	71
12284003	33
12284006	19
12284007	71
12582001	17
12802001	20
12805001	18
12876001	55

b) Reproduction of Figures 4, 5, and 6 at the new 0.10° spatial resolution:

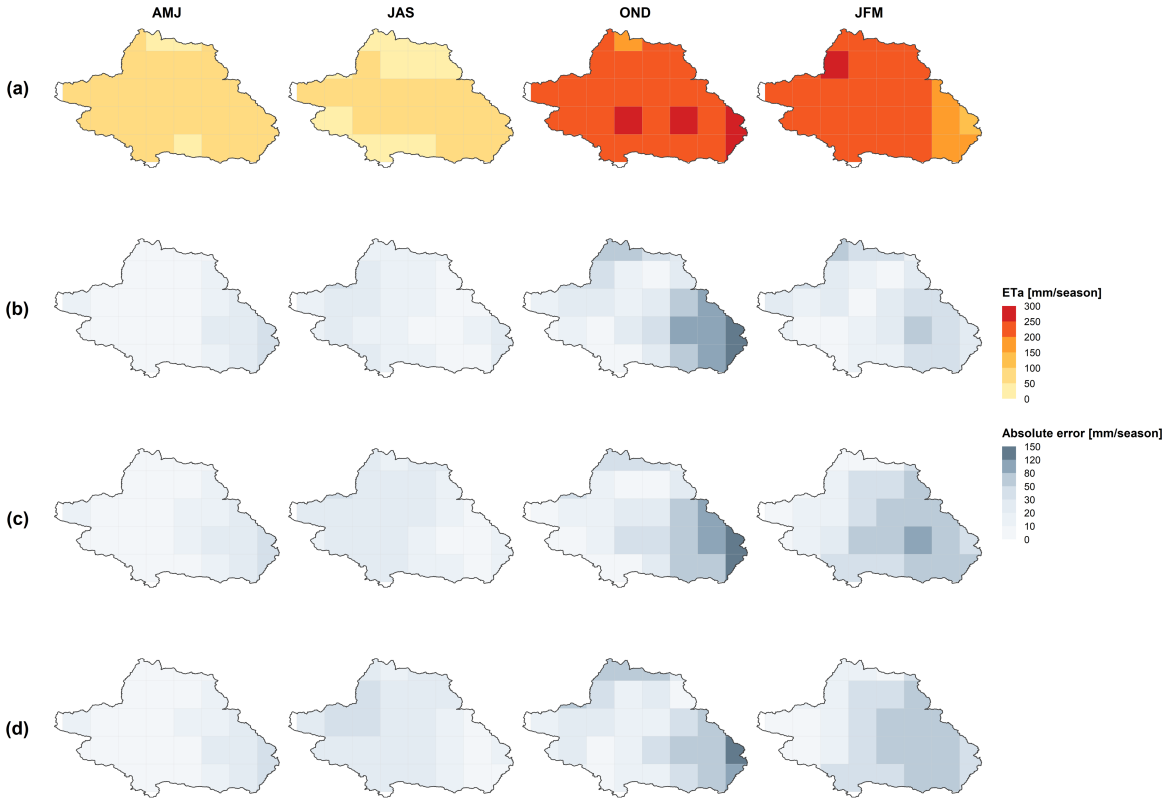

Figure R1.1 (Figure 4 in the main text): (a) Reference ETa from GLEAM for each three-month season (AMJ, JAS, OND, JFM), and spatial distribution of relative errors in simulated ETa with the TUW model with (b) the best SPEAK, (c) the best SPAEF, and (d) the best WSPAEF calibrations for catchment ID: 11302001 (period 2001-2020). The red colour bar indicates the ETa in (mm/season), whereas the grey colour bar indicates the absolute error (in mm/season) in simulated ETa with respect to the reference GLEAM data.

Figure R1.2 (Figure 5 in the main text): FSS for SPEAK-, SPAEF-, and WSPAEF-based mean annual ETa simulation (catchment ID: 11302001; period: 2001-2020). From a) to d): 75th, 90th, 95th, and 99th ETa percentiles taken as thresholds for FSS computation.

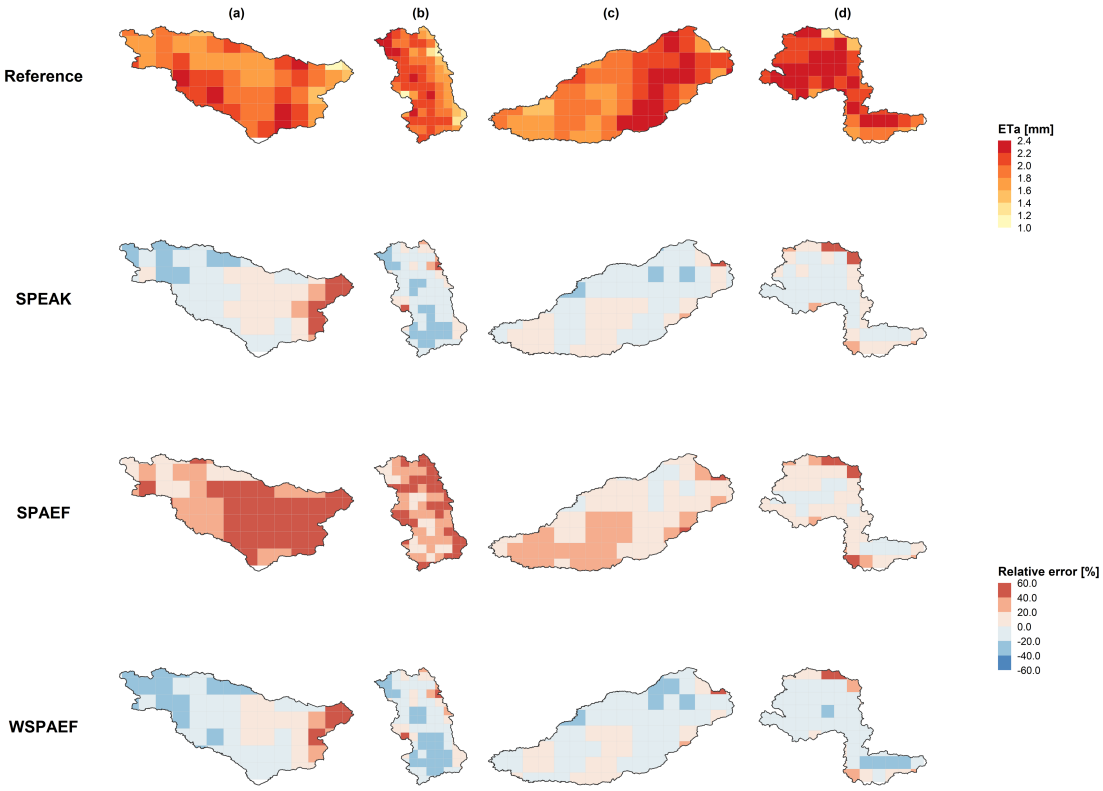

Figure R1.3 (Figure 6 in the main text): Comparison between GLEAM actual evapotranspiration fields (annual mean per pixel) and simulated ETa using parameters obtained by calibrating with SPEAK, SPAEF, and WSPAEF against GLEAM reference data. Each column corresponds to a specific catchment: (a) Río Itata at Balsa Nueva Aldea (ID: 8135002); (b) Río Biobío at Rucalhue (ID: 8317001); (c) Río Cautín at Almagro (ID: 9140001); and (d) Río Trancura Antes Río Llafenco (ID: 9414001).

c) Moran index and closer agreement with reference data:

Table R1.2. Moran index and closer agreement with GLEAM data.

		MORAN Index				
Catchment ID	**N° pixels at 0.10°**	**GLEAM**	**SPEAK**	**SPAEF**	**WSPAEF**	**Closer to GLEAM**
1310002	18	0.47	0.64	0.64	0.64	**draw**
2101001	30	0.42	0.77	0.57	0.77	**SPAEF**
3414001	29	0.35	0.39	0.40	0.69	**SPEAK**
3430003	93	0.70	0.87	0.91	0.87	**SPEAK, WSPAEF**
4314002	22	0.16	0.35	0.31	0.37	**SPAEF**
4513001	20	0.17	0.10	0.26	0.09	**SPEAK**
4523002	47	0.23	0.56	0.57	0.57	**SPEAK**
4703002	20	0.15	0.57	0.52	0.52	**SPAEF, WSPAEF**
7112001	16	0.64	0.14	0.83	0.09	**SPAEF**
8117005	18	0.32	0.83	0.75	0.86	**SPAEF**
8123001	16	0.38	0.36	0.26	0.42	**SPEAK**
8124002	20	0.31	0.61	0.40	0.62	**SPAEF**
8132001	27	0.48	0.61	0.35	0.70	**SPEAK**
8135002	66	0.50	0.87	0.75	0.89	**SPAEF**
8317001	104	0.63	0.53	0.22	0.53	**SPEAK**
8334001	150	0.69	0.66	0.60	0.72	**SPEAK, WSPAEF**
8342001	22	0.52	0.46	0.42	0.54	**WSPAEF**
9123001	24	0.31	0.25	0.08	0.19	**SPEAK**
9129002	46	0.39	0.75	0.65	0.73	**SPAEF**
9135001	29	0.58	0.67	0.61	0.67	**SPAEF**
9140001	81	0.52	0.82	0.79	0.84	**SPAEF**
9404001	25	0.20	0.08	-0.01	0.08	**SPEAK, WSPAEF**
9414001	22	0.31	0.20	0.21	0.30	**WSPAEF**
9420001	40	0.37	0.13	0.14	0.37	**WSPAEF**
9437002	110	0.43	0.85	0.84	0.86	**SPAEF**
10111001	72	0.31	0.56	0.57	0.61	**SPEAK**
10134001	30	0.06	0.55	0.34	0.51	**SPAEF**
10304001	30	0.38	0.36	0.25	0.29	**SPEAK**
10328001	39	0.44	0.77	0.84	0.79	**SPEAK**
10340001	19	0.17	0.04	0.41	0.09	**WSPAEF**
10356001	37	0.55	0.18	0.18	0.26	**WSPAEF**
10364001	80	0.61	0.84	0.87	0.86	**SPEAK**
11141001	18	0.37	0.50	0.65	0.67	**SPEAK**
11143001	33	0.51	0.85	0.83	0.86	**SPAEF**

11302001	33	0.29	0.85	0.86	0.85	**SPEAK, WSPAEF**
11308001	71	-0.03	0.89	0.90	0.92	**SPEAK**
12284003	33	0.26	0.48	0.65	0.47	**WSPAEF**
12284006	19	0.35	0.64	0.63	0.63	**SPAEF, WSPAEF**
12284007	71	0.51	0.78	-0.09	0.77	**WSPAEF**
12582001	17	0.07	0.61	0.70	0.72	**SPEAK**
12802001	20	0.27	0.66	0.67	0.66	**SPEAK, WSPAEF**
12805001	18	0.33	0.45	0.69	0.45	**SPEAK, WSPAEF**
12876001	55	0.53	0.55	0.82	0.55	**SPEAK, WSPAEF**
AVERAGE	**41.63**	**0.38**	**0.55**	**0.53**	**0.58**	
MIN	**16**	**-0.03**	**0.04**	**-0.09**	**0.08**	
MAX	**150**	**0.70**	**0.89**	**0.91**	**0.92**	

d) Summary of results presented in AR 5.c:

Table R1.3. Hits (closer agreement with GLEAM data) as a function of SPEAK, SPAEF, and WSPAEF.

Objective function	Hits	Percentage
SPEAK	22	41%
SPAEF	15	28%
WSPAEF	17	31%