Supporting information of

Same Streamflow, Different Water Stories: The Hidden Impacts of Streamflow-Only Calibration in Distributed Hydrological Modeling

Nicolás A. Vásquez[1,*], Pablo A. Mendoza[1,2], Wouter Knoben[3], Martyn Clark[3], Tricia Stadnyk[3], and Naoki Mizukami[4]

[1]Department of Civil Engineering, Universidad de Chile, Santiago, Chile
[2]Advanced Mining Technology Center (AMTC), Universidad de Chile, Santiago, Chile
[3]Schulich School of Engineering, University of Calgary, Calgary, Alberta, Canada
[4]NSF National Center for Atmospheric Research (NCAR), Boulder, Colorado, United States

This document shows the results for all the basins and objective functions and presents additional analyses. The different sections of this document are summarized as follows:

1. **Missing fSCA data**
   This section provides a brief explanation about filling in missing fSCA data from MOD10 and MYD10 products.

2. **Principal Component Analysis**
   We use the first principal component (PC_1) to derive the *a priori* spatial distribution of VIC parameters. Hence, we show how much of the total variance is explained by PC_1 for all basins, as well as its relationship with the original soil attributes.

3. **Average seasonalities**
   The average seasonalities for streamflow (Q), evapotranspiration (ET), soil moisture (SM_1), land surface temperature (LST), and fractional snow-covered area (fSCA) are shown for five calibration objective functions and nine parameter configurations.

4. **Average deeper soil layers' moisture seasonalities**
   The manuscript shows the average (dimensionless) soil moisture for three basins. This document extends the figure to all the basins.

5. **Verifying equifinality**
   Main results only consider the "best" parameter set, chosen by maximizing each objective function. In this document, we verify if including more parameter sets (i.e., behavioral parameters) may modify the average seasonality of Q, ET, SM_1, LST, and fSCA.

6. **Grid cell performance**
   The grid cell performance for ET, SM_1, LST, and fSCA are shown for each basin, objective function, and parameters' configuration.

7. **Multivariate calibration**

We calibrated the Variable Infiltration Capacity (VIC) model Liang et al. (1994) based on objective functions computed only with streamflow data. We expand our analysis by also including ET and SM_1 in the objective function.

S.1 Filling in missing fSCA data

For a specific grid cell and missing day t, the fSCA is estimated as the average between fSCA from the previous $(t-1)$ and the next day $(t+1)$. After filling as many days as possible for individual grid cells, we fill the remaining grid cells using the mean fSCA among the eight nearest neighboring grid cells for a specific day t, where the weighted mean is computed considering a similarity index (between the grid cell to be filled and its neighbors) based on elevation and aspect. The temporal and spatial filling is repeated until all grid cells and days have an estimated fSCA. (see Cornwell et al. (2016) for details).

S.2 Principal component analysis

To derive *a priori* parameters, we conduct a Principal Component Analysis between (i) elevation, (ii) slope, (iii) clay and sand contents, and (iv) bulk density. Figure S.1a illustrates the fraction of the total variance explained by each component, while Figure S.1b contrasts the PC_1 against the attributes.

S.3 Grid cell performance

The simulated performance at the grid cell scales is displayed in Figures S.2-S.7 for ET, Figures S.8-S.13 for SM_1, Figures S.14-S.19 for LST and Figures S.20-S.25 for fSCA. The performance of each regularized parameter case is contrasted against the case with spatially constant parameters.

S.4 Average seasonalities

Figures S.26-S.32 display the average seasonalities of streamflow (Q), evapotranspiration (ET), soil moisture (SM_1), land surface temperature (LST), and fractional snow-covered area (fSCA) for different streamflow-based objective functions and nine different parameters' configuration: (i) all parameters are spatially constant and (ii) one parameter is regularized while the rest remain spatially constant.

S.5 Deeper soil layers' moisture

Different parameters' spatial configurations may yield very similar simulated streamflow response and seasonality. Nevertheless, ET-averaged seasonality can be different among those configurations, with important temporal shifts, mostly explained by different soil moisture seasonalities, feeding the vegetation's transpiration. Figure S.33 illustrates the average soil moisture seasonality for the second (SM_2) and third (SM_3) soil layers.

S.6 Verifying equifinality, LAI and root fraction

For each different spatial parameter's configurations and objective functions (OF), we selected the "best" parameter based on the OF value. However, the results can be sensitive to such criteria. Hence, we evaluate the average seasonalities for the best 1% parameter sets considering two criteria:

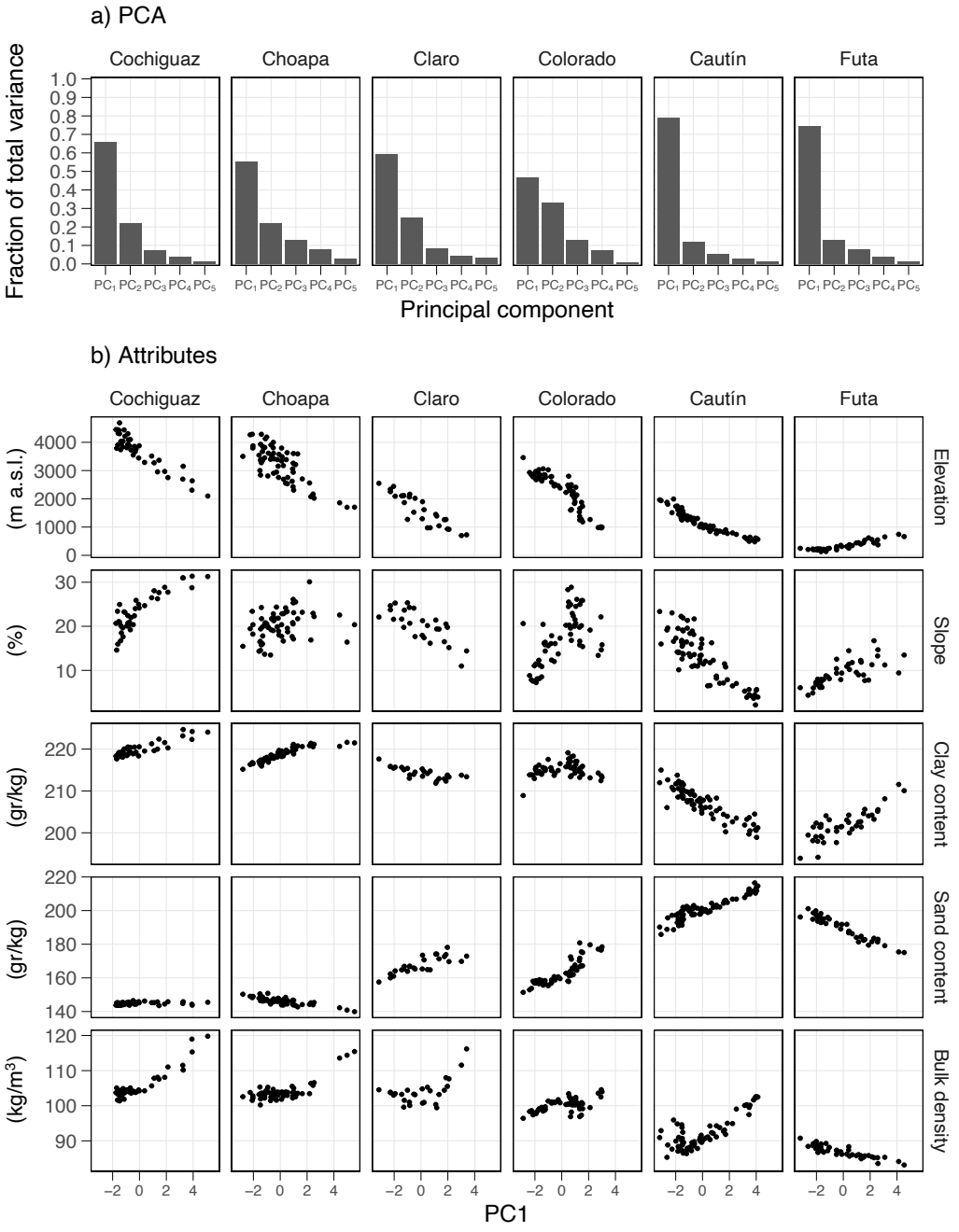

Figure S.1. Principal component analysis to derive *a priori* parameters. a) Fraction of total variance explained by each component. b) Relationship between PC_1 and the attributes. Each dot represents a grid cell (0.05° horizontal resolution).

1. Choose the best 1% OF value.

2. Choose all the parameter sets with an OF higher than the best OF value minus a threshold ($OF' > (OF_{best} - OF_{threshold})$). Among those parameters meeting that criterion, select the same number of parameter sets as in criteria 1). The selected $OF_{threshold} = 0.02$. For example, if the maximum $KGE(Q)$ is 0.7, the parameter sets to randomly select the "behavioral" parameters are those with a $KGE(Q) > 0.7 - 0.02 = 0.68$.

Figure S.34 illustrates the seasonalities for the best parameters according to criteria 1) and 2). Although there is dispersion among each parameter configuration, the average seasonalities remain mostly the same. Interestingly, for the Cautín River basin, different streamflow-behavioral parameter sets yield different ET' seasonalities, even for the same spatial configuration (e.g., parameters spatially constant). Similarly, for the Colorado River basin, all the behavioral parameter sets yield very similar Q seasonalities, while regularizing $Depth_2$ impacts the ET' seasonality.

For the Claro and Colorado River basins, the ET average seasonalities show large biases. In VIC, to compute ET, the Leaf Area Index (LAI) is required for each vegetation type for each grid cell. To verify if ET's seasonality is coherent with the MOD16 product, we compute the basin-scale LEAF area index per vegetation type. Figure S.35 displays the monthly LAIs values for each basin. These values show that the highest values are reached during summer (DJF), similar to the ET derived from the MOD16 and GLEAM products. Hence, VIC simulations for the Claro and Colorado River basins may not adequately represent the observed ET seasonality.

The VIC model estimates the vegetation's transpiration as a weighted sum considering the root fraction and vegetation type within each soil layer. However, VIC outputs are aggregated for each grid cell. We estimate the contribution from each soil layer to the total vegetative transpiration based on the average root fraction per grid cell and then aggregate the results at the basin scale. Figure S.36 illustrates the relative soil layer contribution to total transpiration and the vegetative transpiration amount. Since the root fraction is constant in time, the relative contribution is the same among the parameter configurations, although with different amounts of transpiration.

References

Cornwell, E., Molotch, N. P., and McPhee, J.: Spatio-temporal variability of snow water equivalent in the extra-tropical Andes Cordillera from distributed energy balance modeling and remotely sensed snow cover, Hydrology and Earth System Sciences, https://doi.org/10.5194/hess-20-411-2016, 2016.

Liang, X., Lettenmaier, D. P., Wood, E. F., and Burges, S. J.: A simple hydrologically based model of land surface water and energy fluxes for general circulation models, Journal of Geophysical Research, https://doi.org/10.1029/94jd00483, 1994.

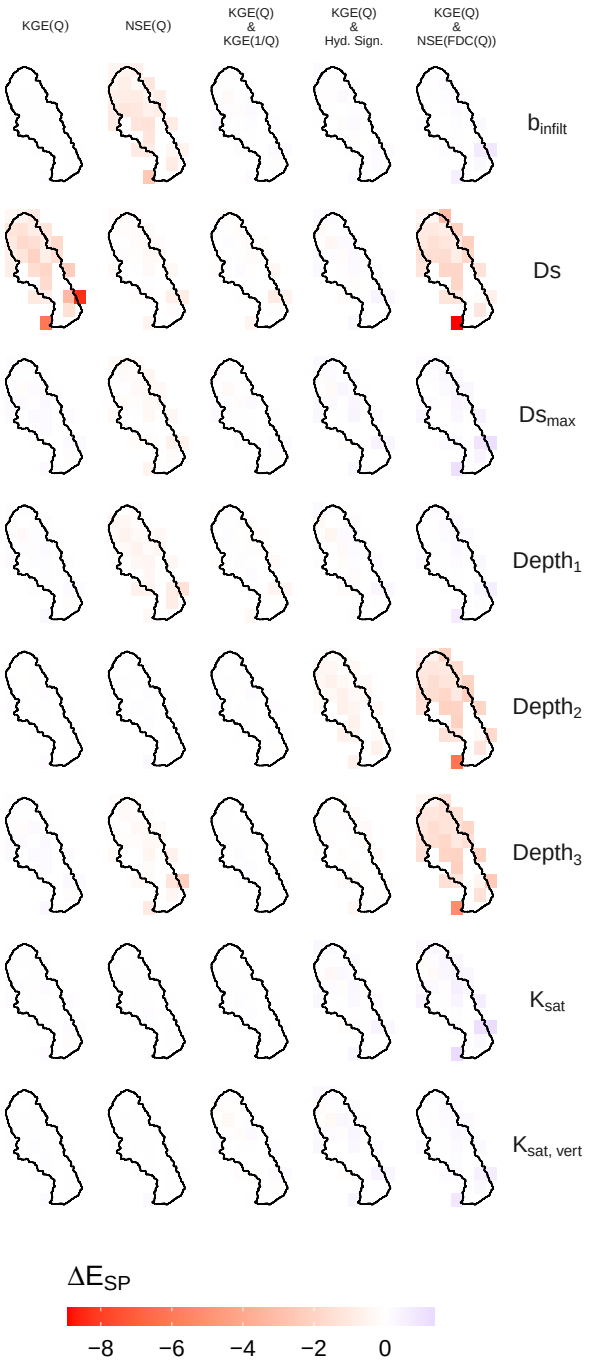

Figure S.2. ET performance change (with respect to the case where parameters are spatially constant) at the grid cell scale for basin Cochiguaz at El Peñón during the calibration period (2005-2018). Notice that the benchmark changes among objective functions. Blue colors represent an improvement in the simulated performance, while red colors a decrease.

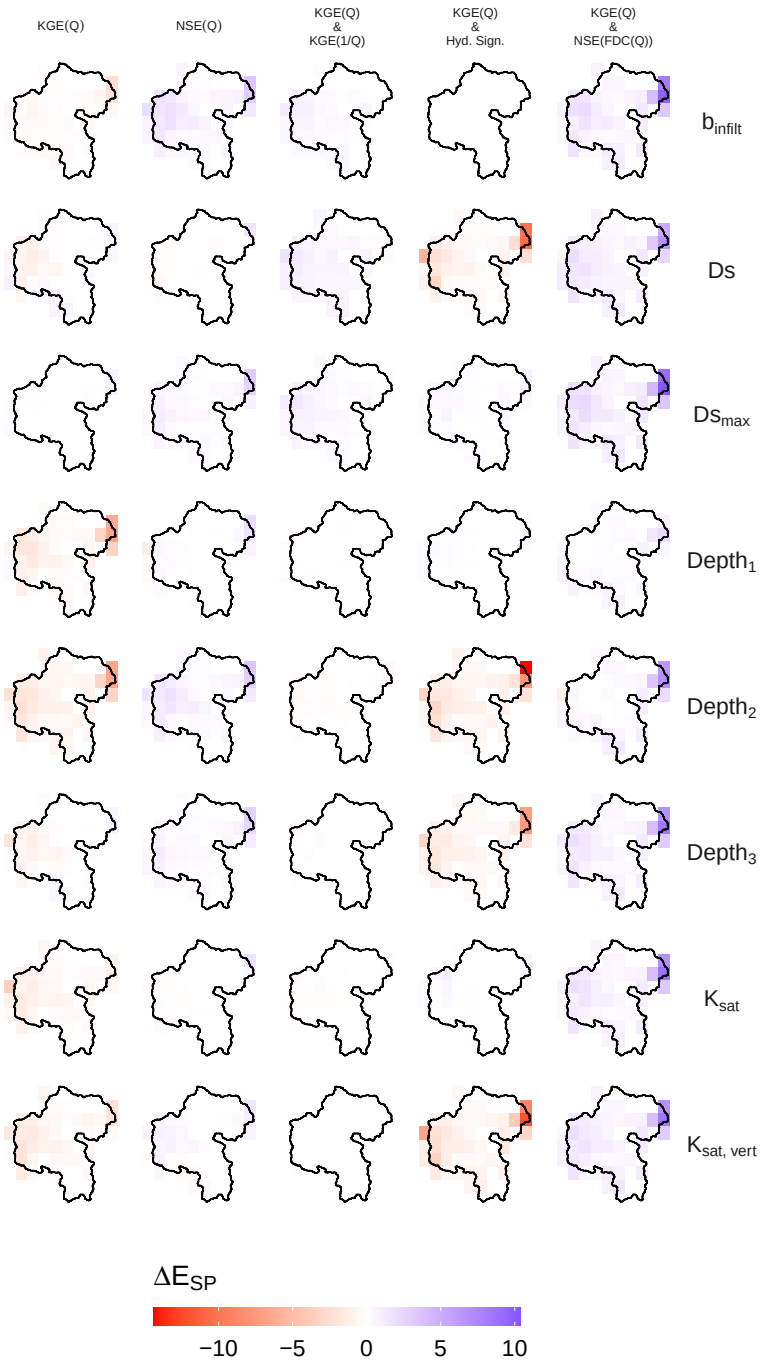

Figure S.3. Same as in Figure S.2, but for the Choapa at Cuncumén basin.

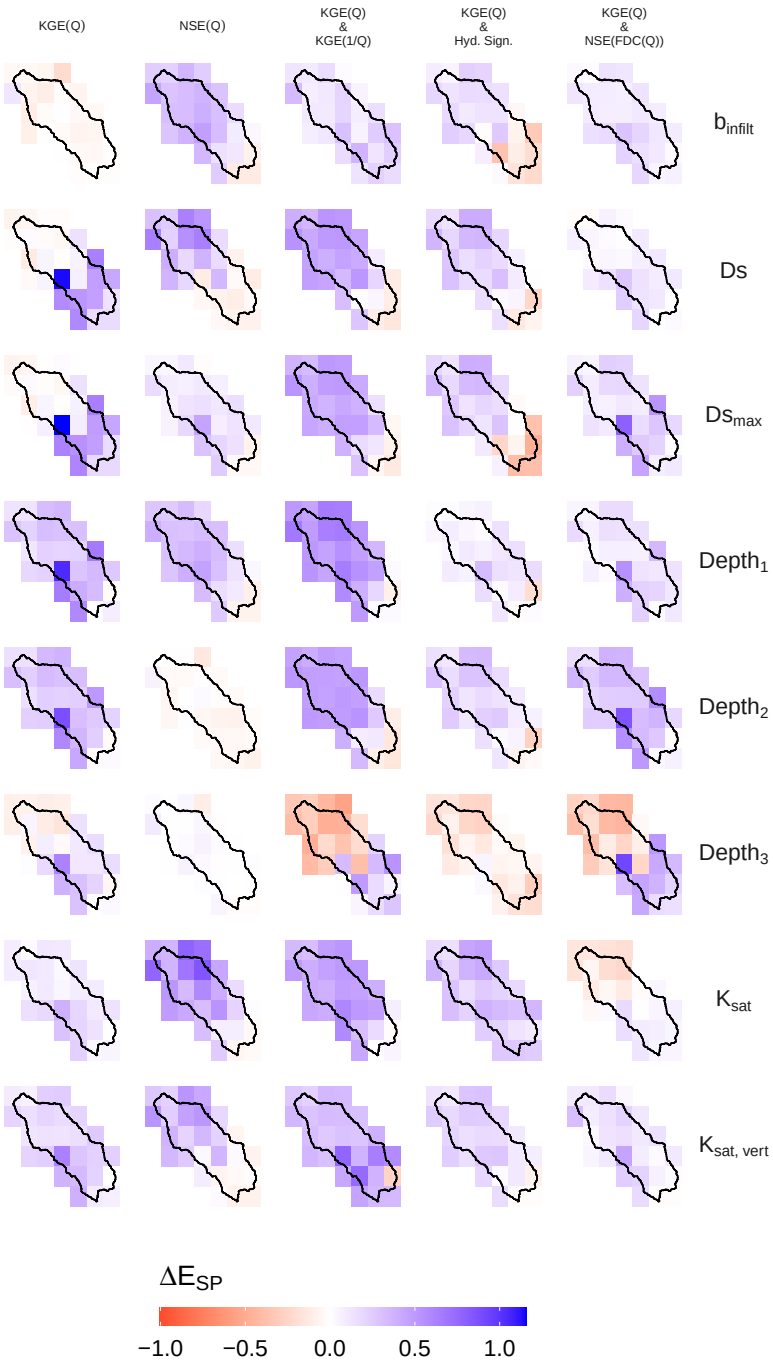

Figure S.4. Same as in Figure S.2, but for the Claro at El Valle basin.

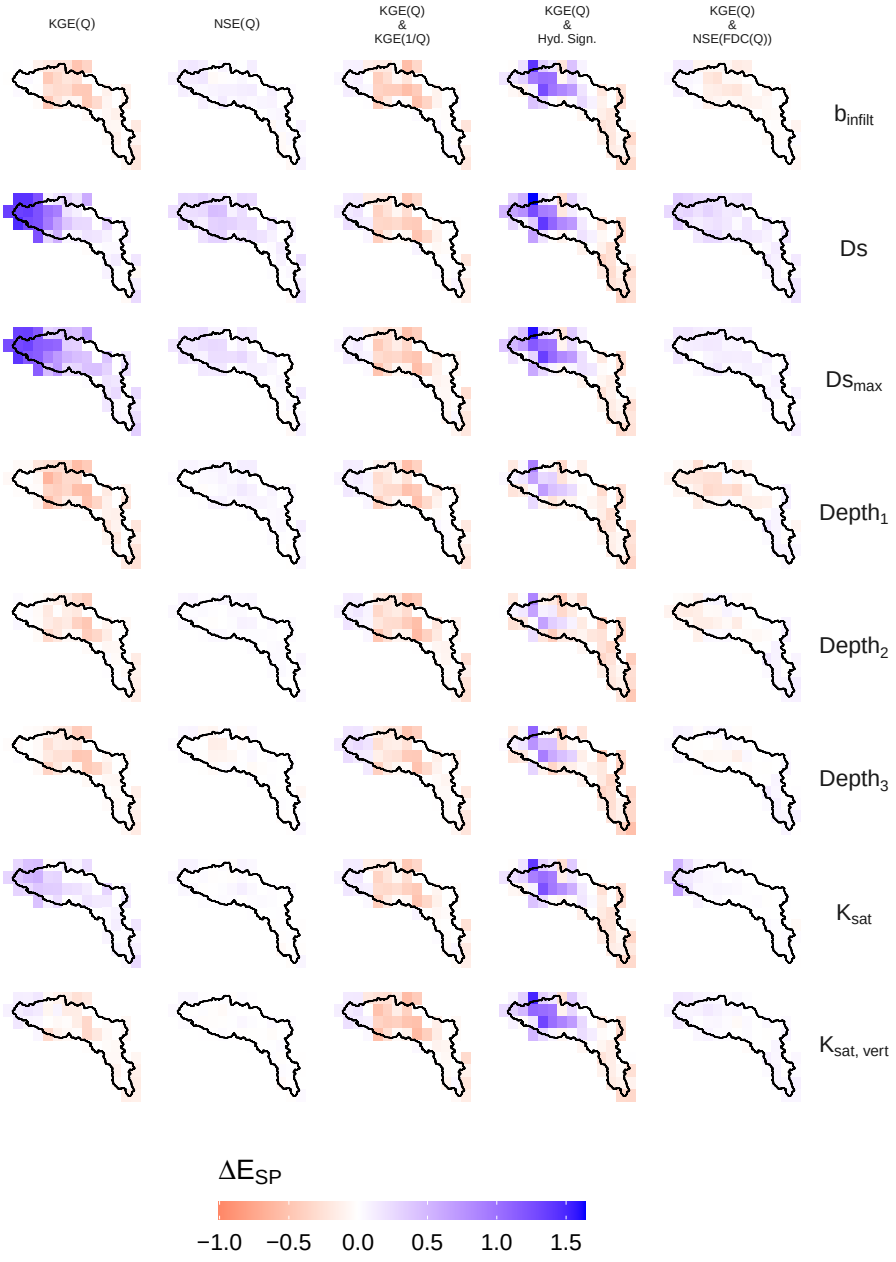

Figure S.5. Same as in Figure S.2, but for the Colorado at Los Palos basin.

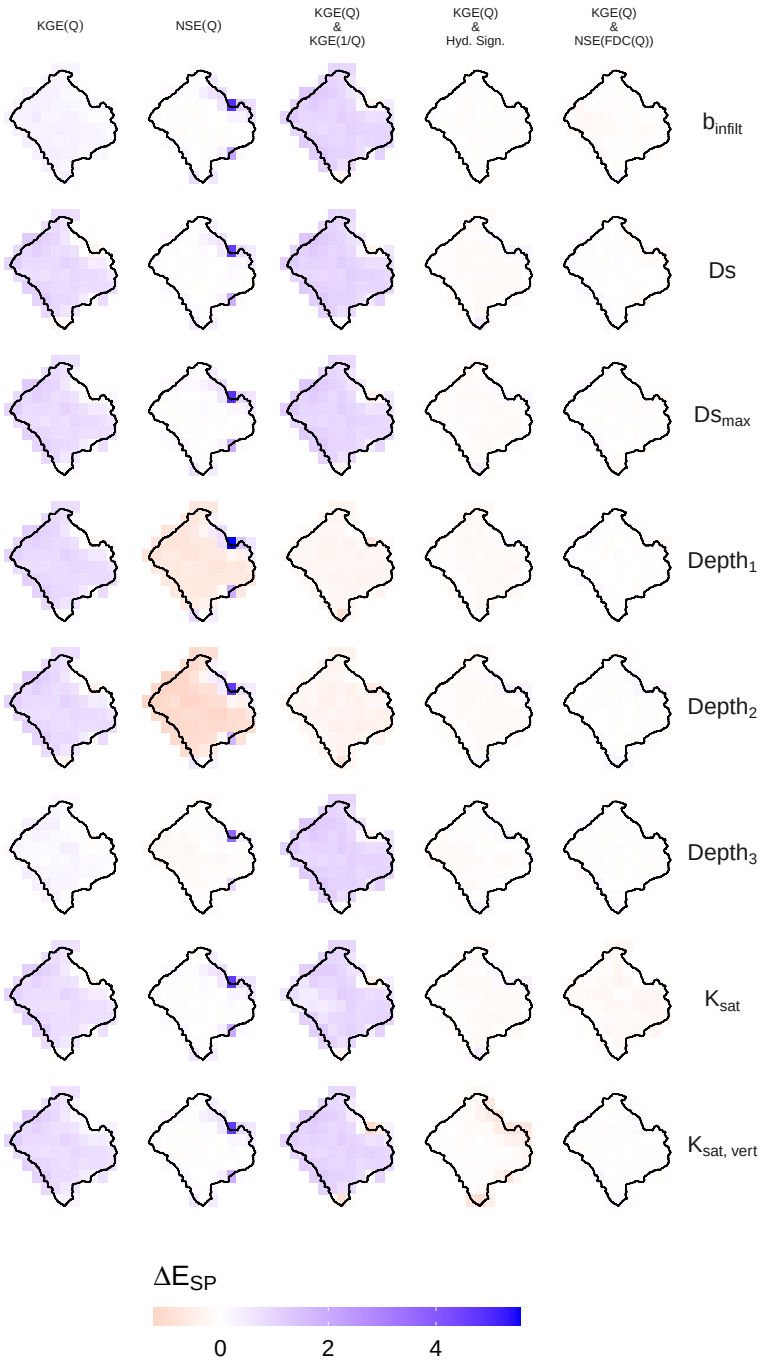

Figure S.6. Same as in Figure S.2, but for the Cautín at Rari-Ruca basin.

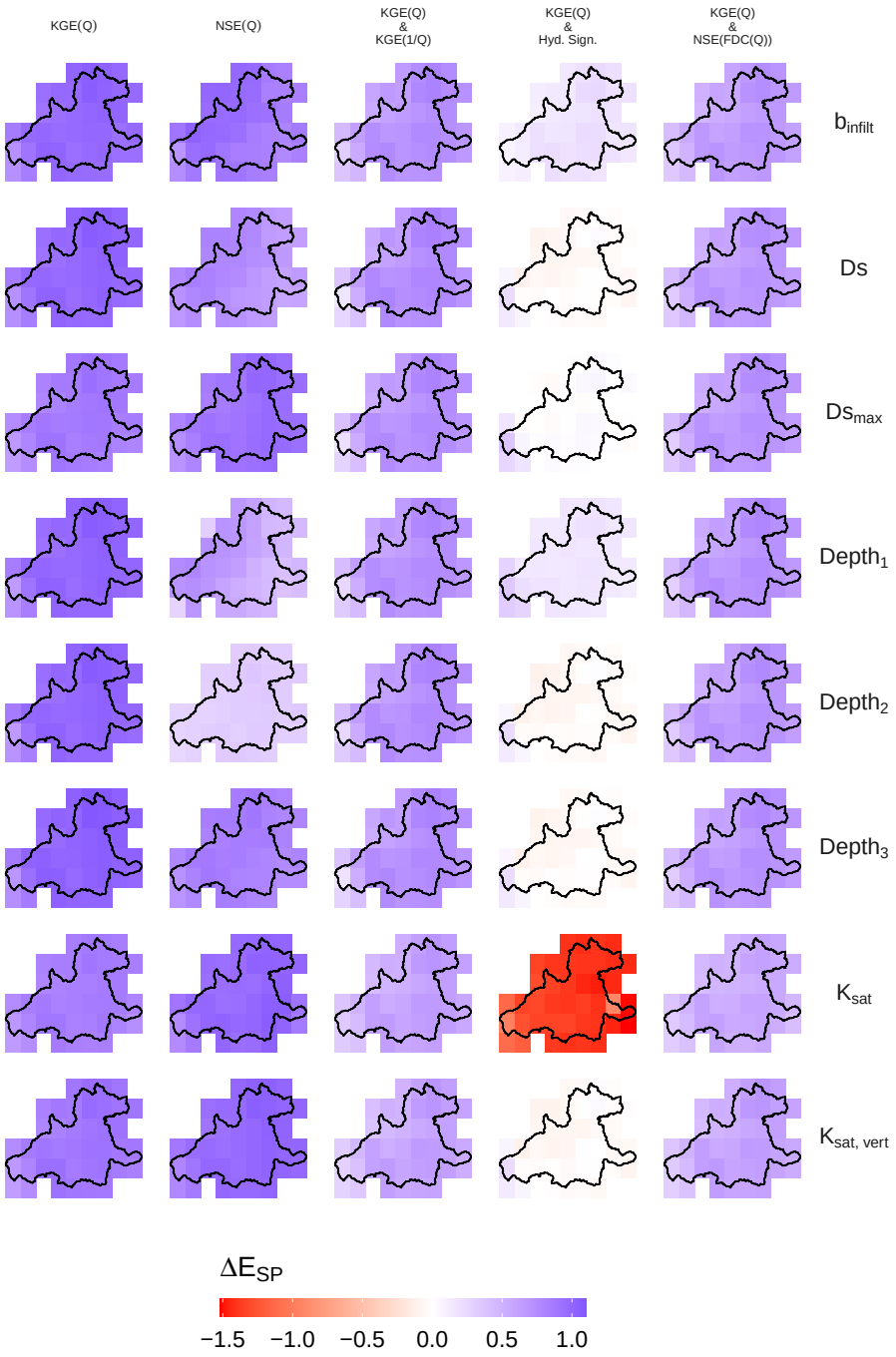

Figure S.7. Same as in Figure S.2, but for the Futa at Tres Chiflones basin.

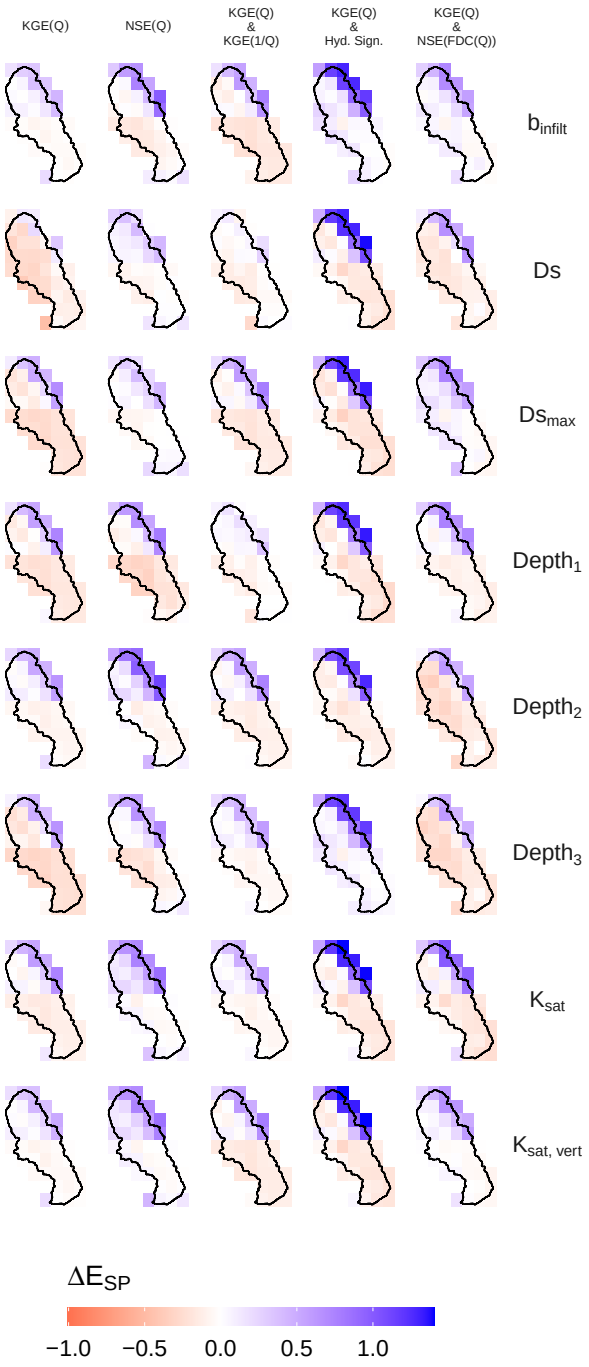

Figure S.8. SM_1 performance change (with respect to the case where parameters are spatially constant) at the grid cell scale for basin Cochiguaz at El Peñón during the calibration period (2005-2018). Notice that the benchmark changes among objective functions. Blue colors represent an improvement in the simulated performance, while red colors a decrease.

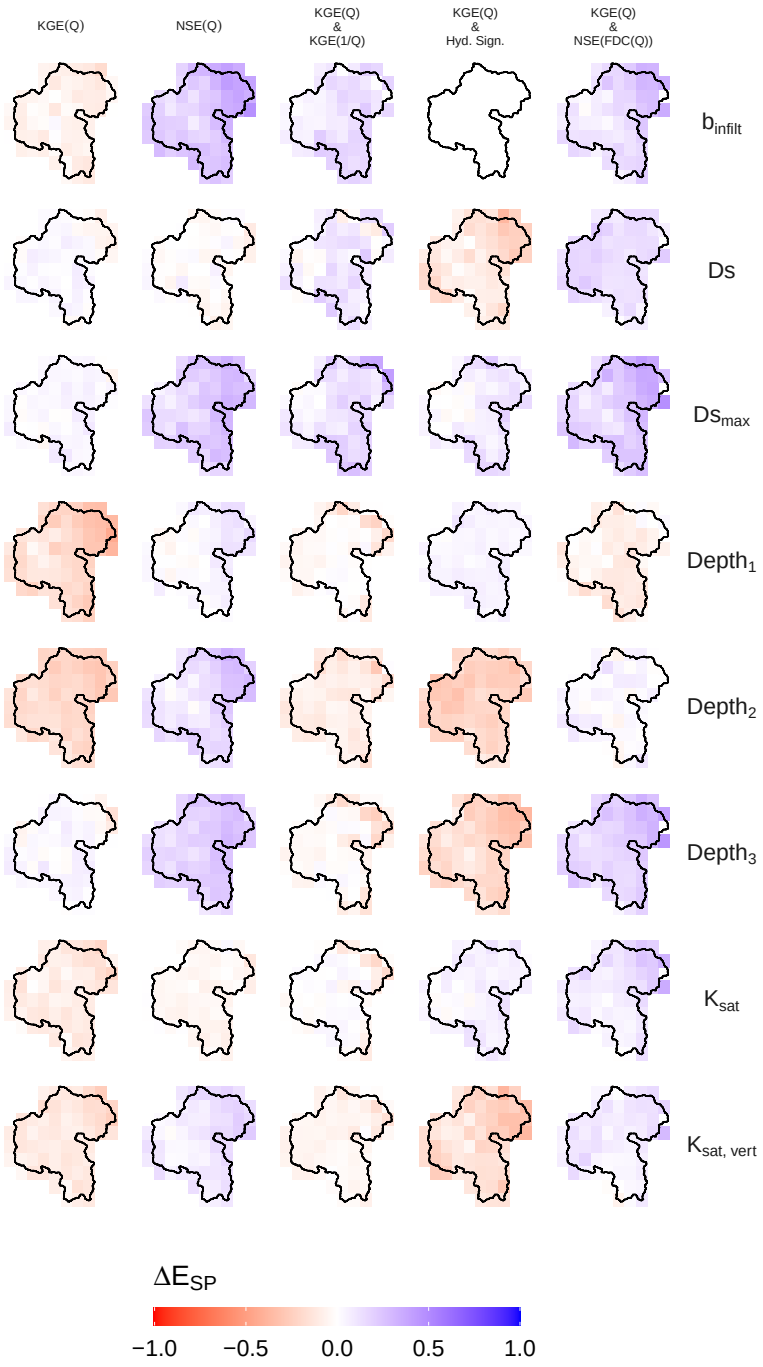

Figure S.9. Same as in Figure S.8, but for the Choapa at Cuncumén basin.

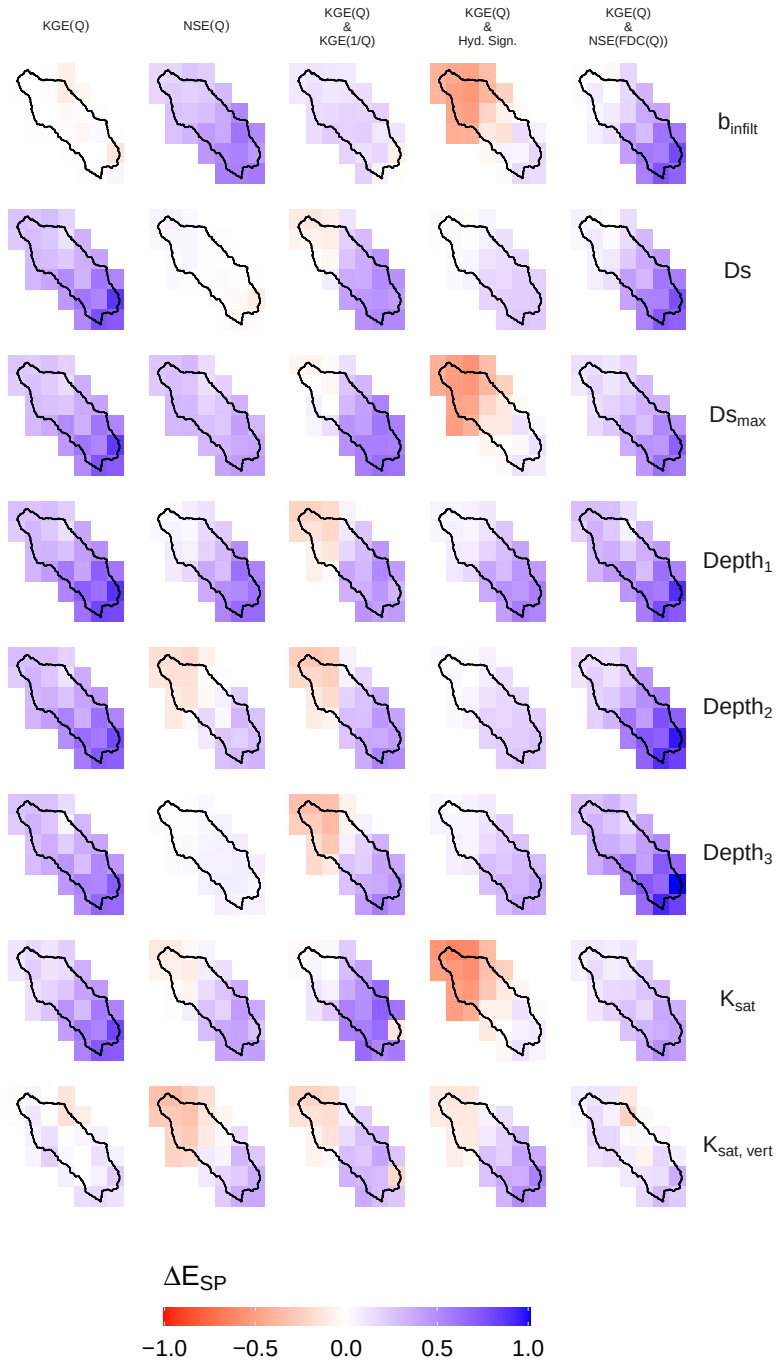

Figure S.10. Same as in Figure S.8, but for the Claro at El Valle basin.

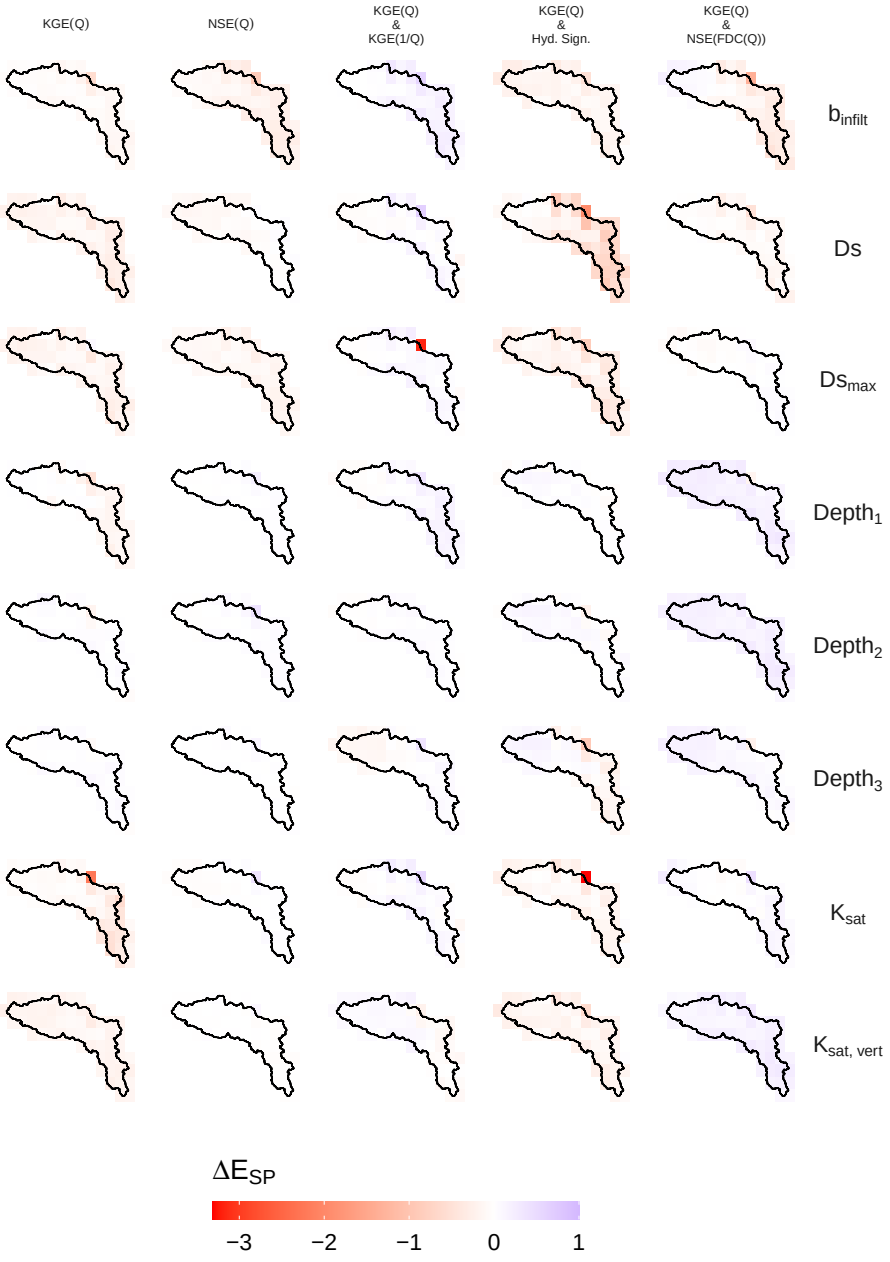

Figure S.11. Same as in Figure S.8, but for the Colorado at Los Palos basin.

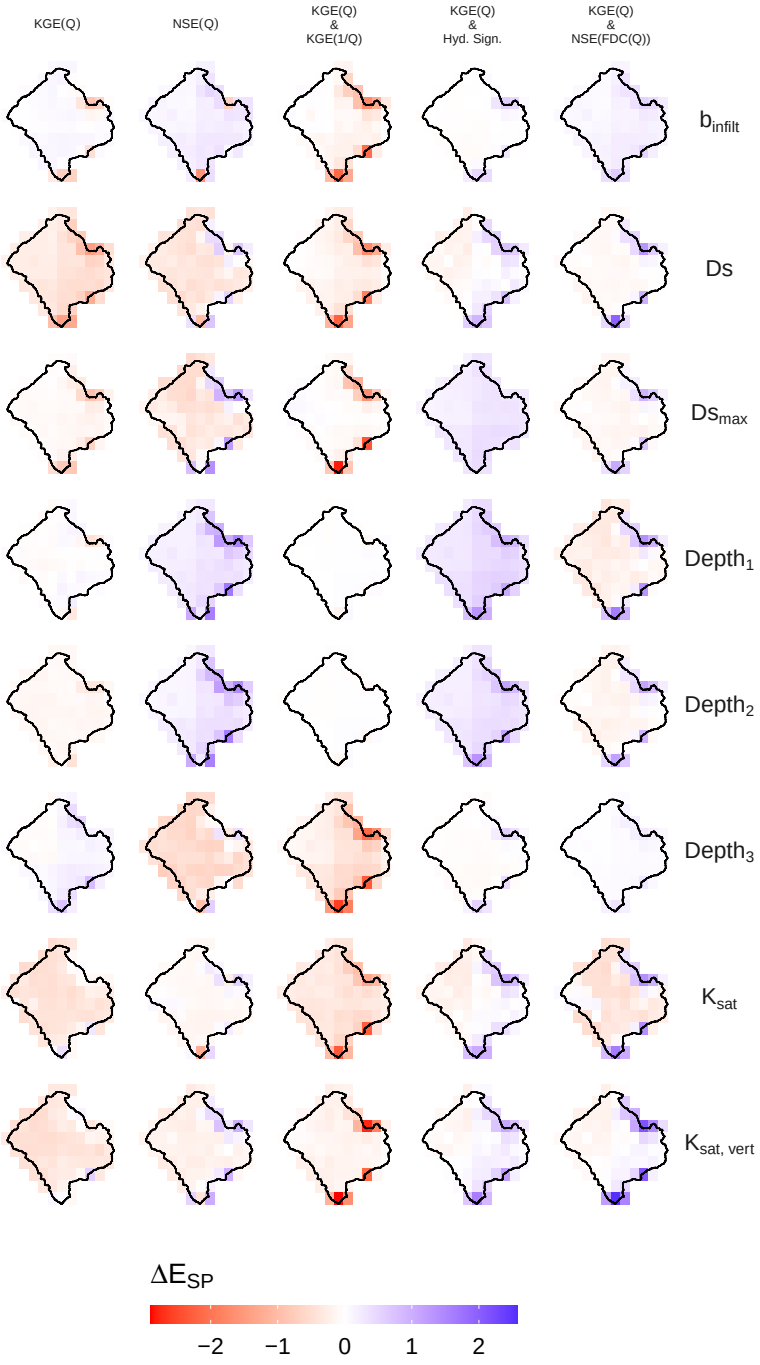

Figure S.12. Same as in Figure S.8, but for the Cautín at Rari-Ruca basin.

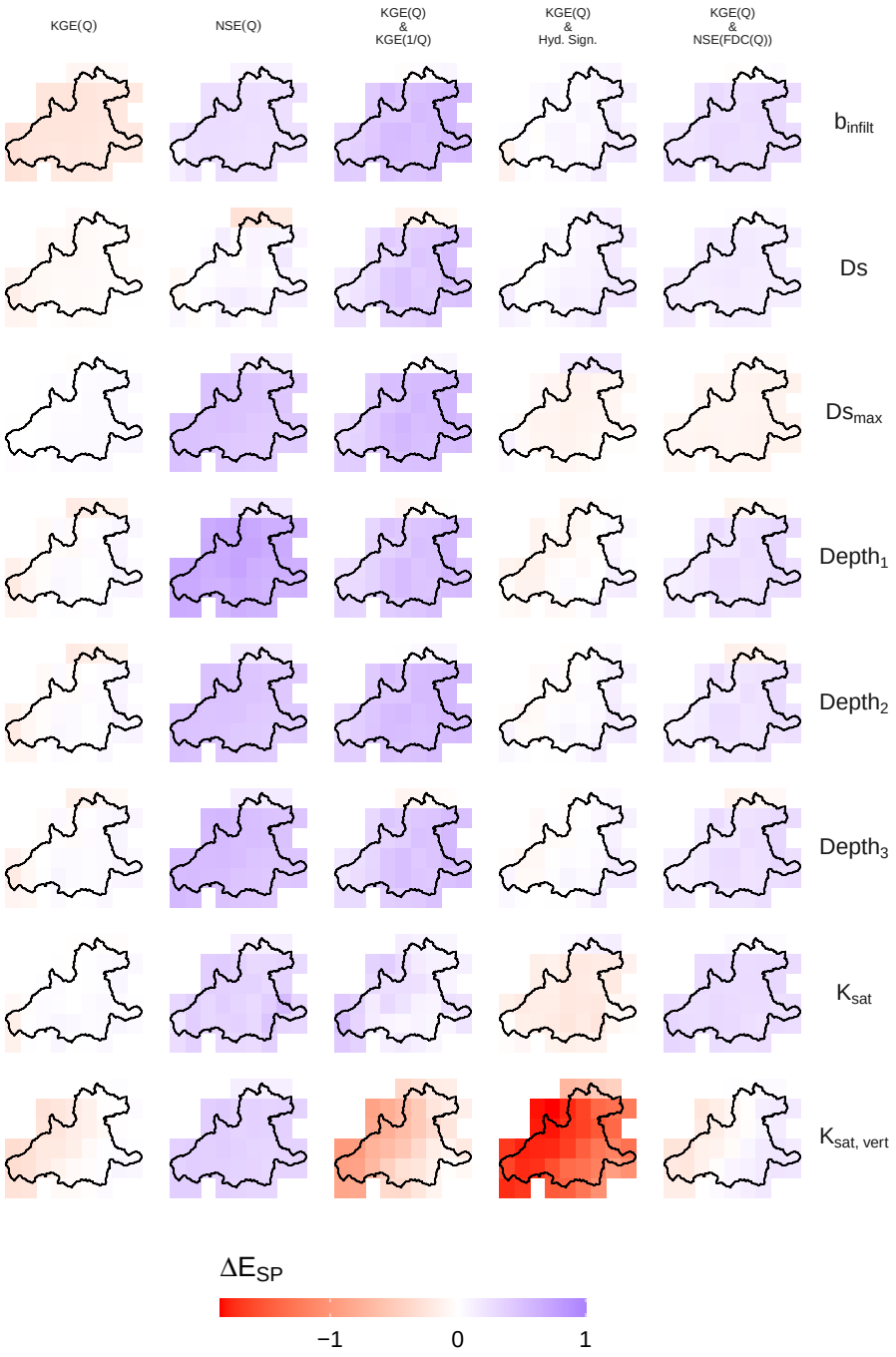

Figure S.13. Same as in Figure S.8, but for the Futa at Tres Chiflones basin.

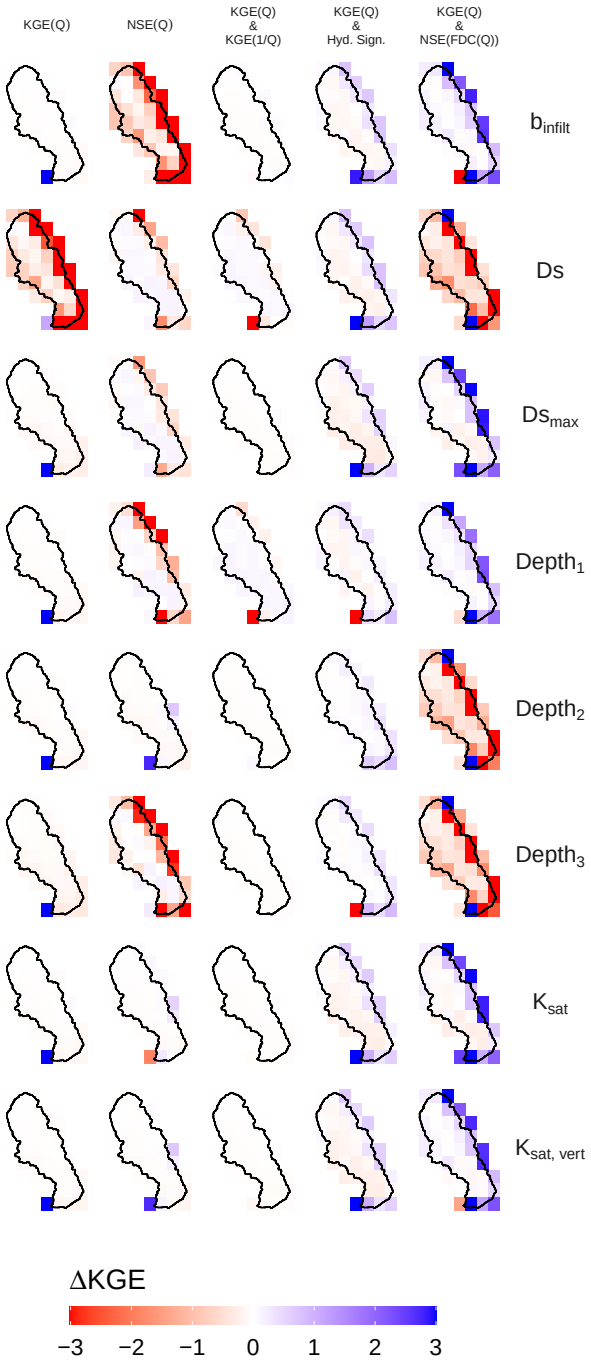

Figure S.14. LST performance change (with respect to the case where parameters are spatially constant) at the grid cell scale for basin Cochiguaz at El Peñón during the calibration period (2005-2018). Notice that the benchmark changes among objective functions. Blue colors represent an improvement in the simulated performance, while red colors a decrease.

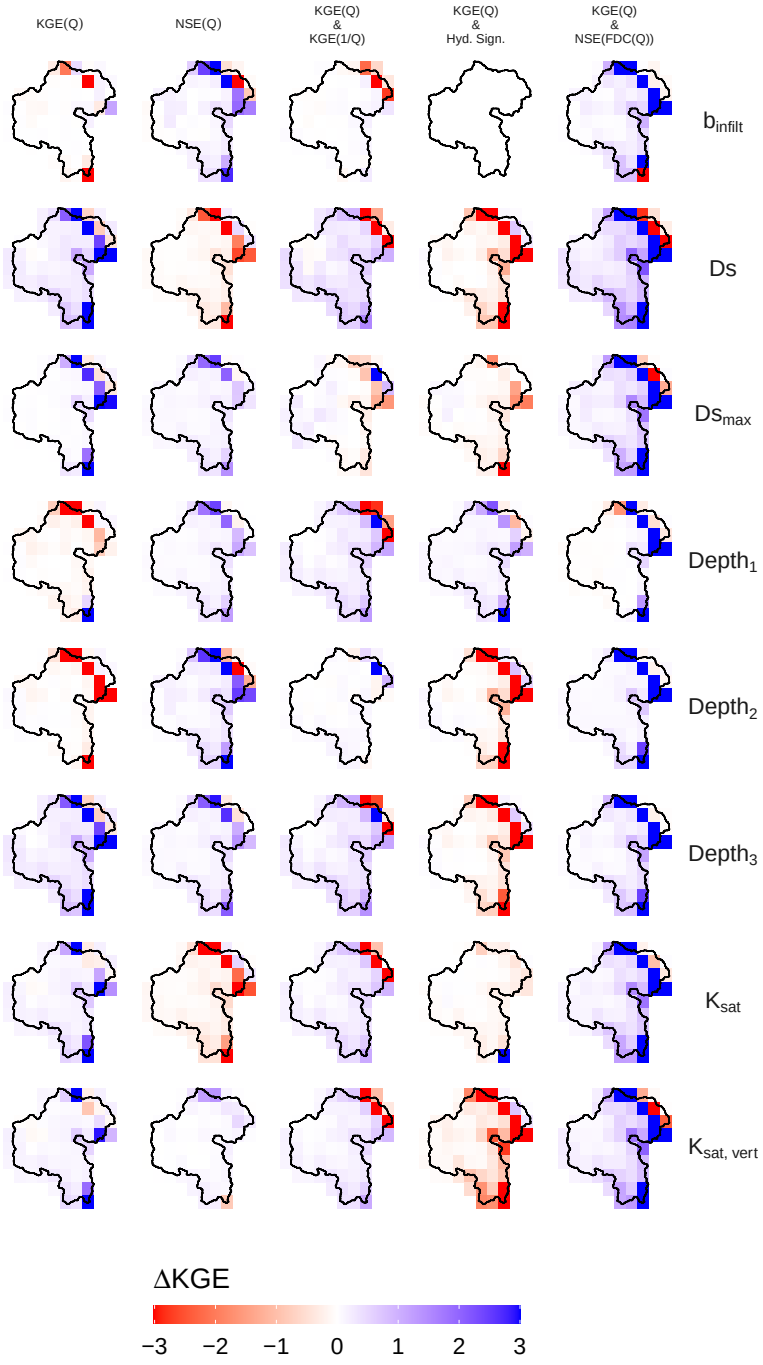

Figure S.15. Same as in Figure S.14, but for the Choapa at Cuncumén basin.

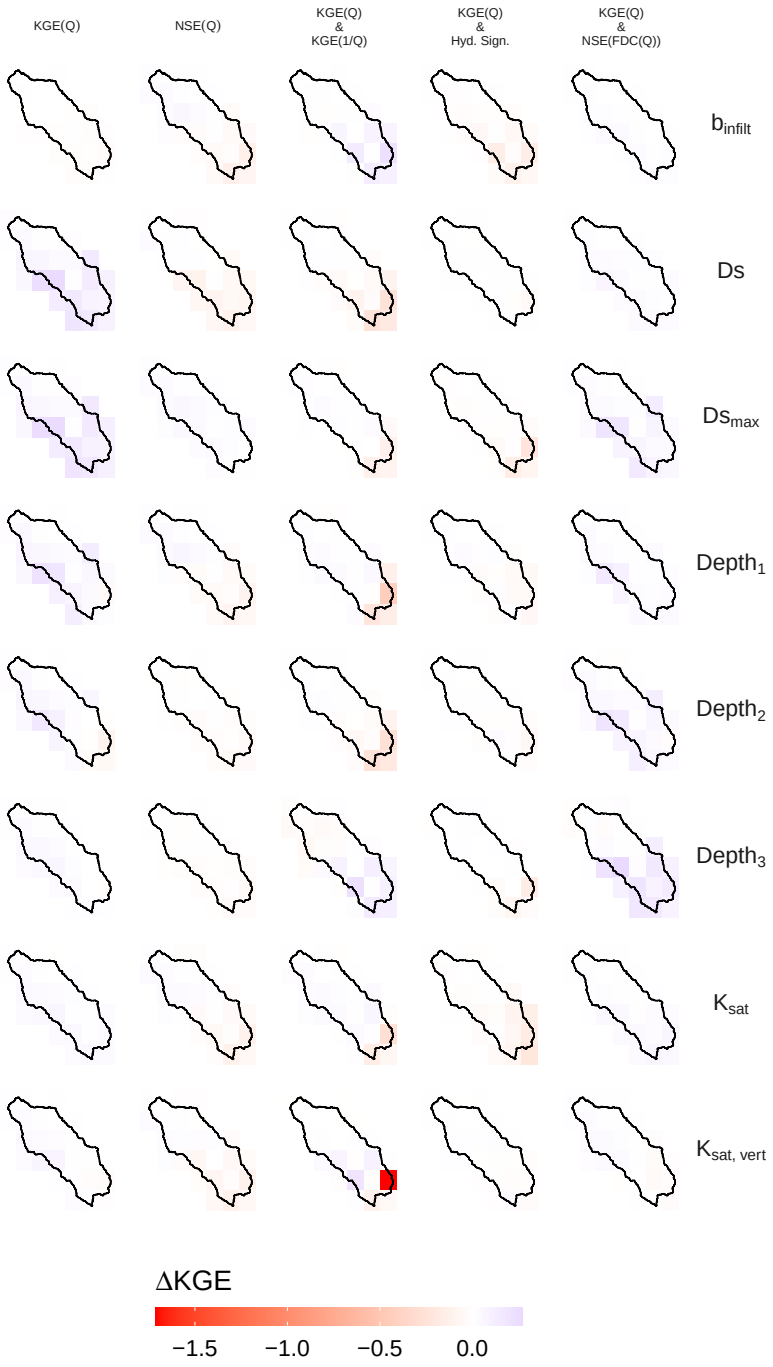

Figure S.16. Same as in Figure S.14, but for the Claro at El Valle basin.

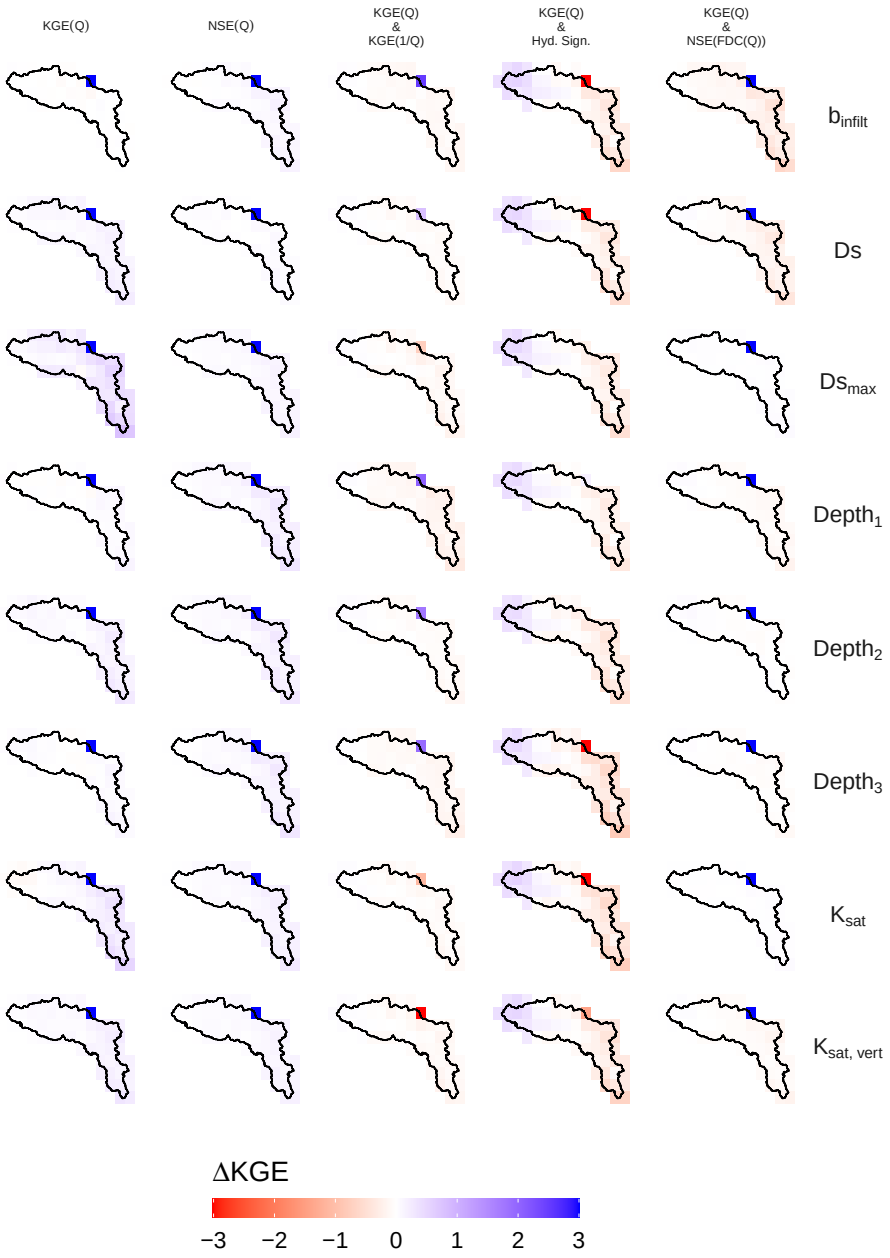

Figure S.17. Same as in Figure S.14, but for the Colorado at Los Palos basin.

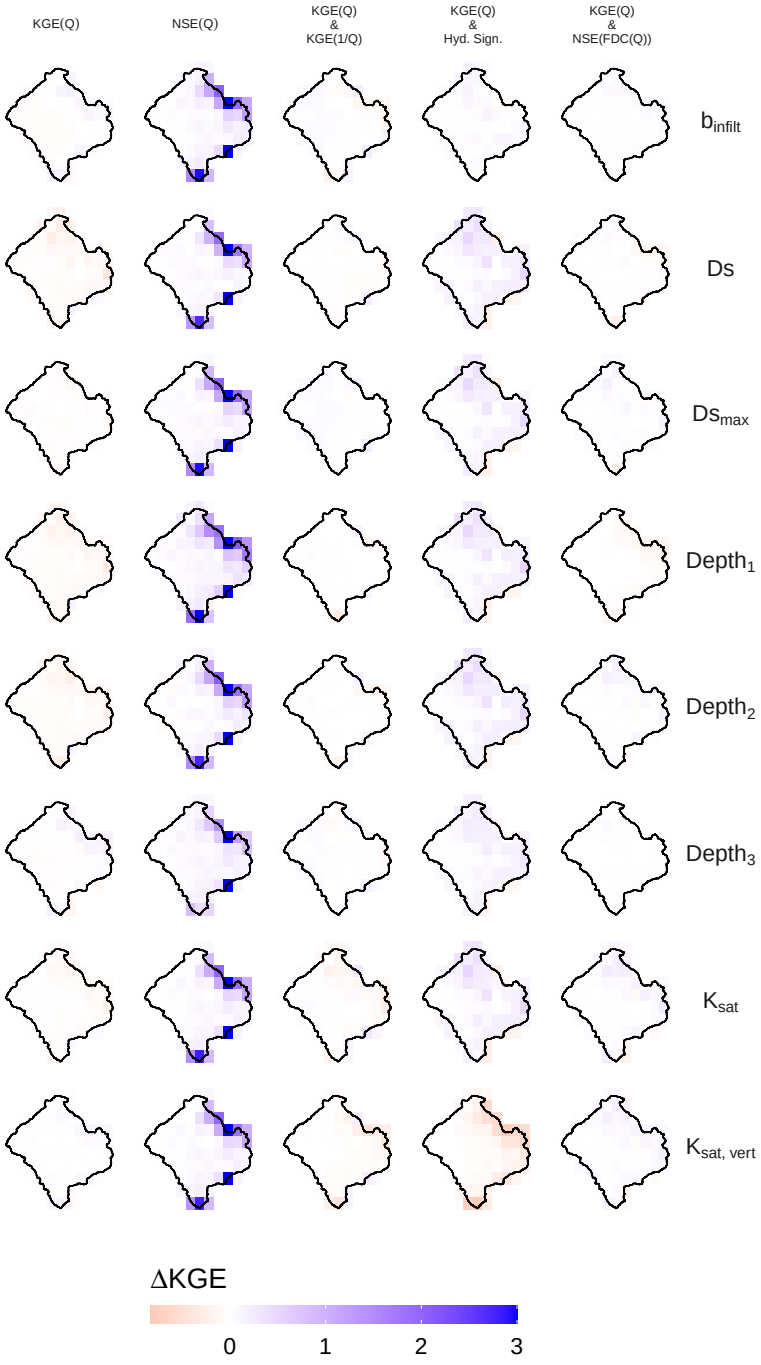

Figure S.18. Same as in Figure S.14, but for the Cautín at Rari-Ruca basin.

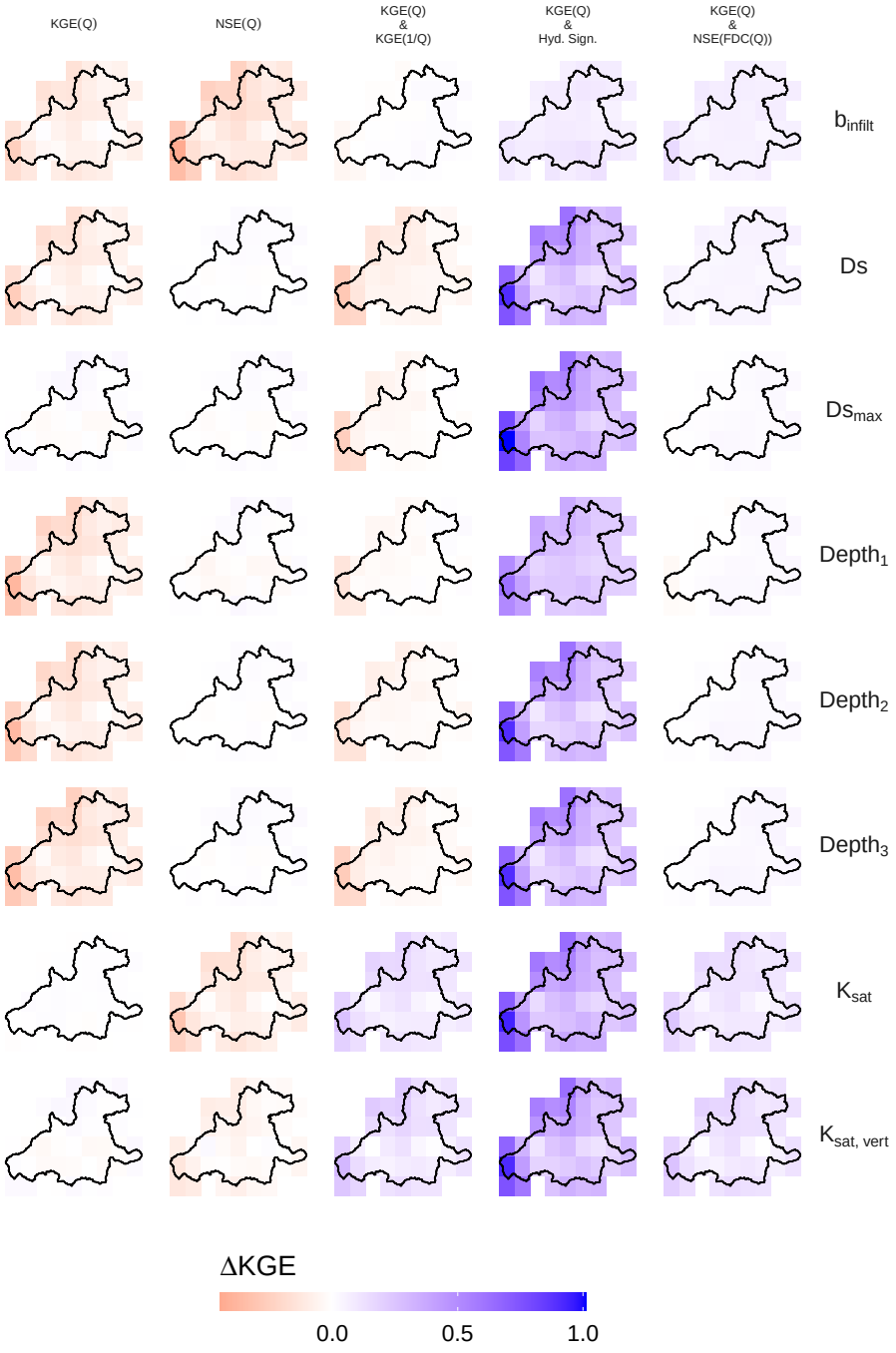

Figure S.19. Same as in Figure S.14, but for the Futa at Tres Chiflones basin.

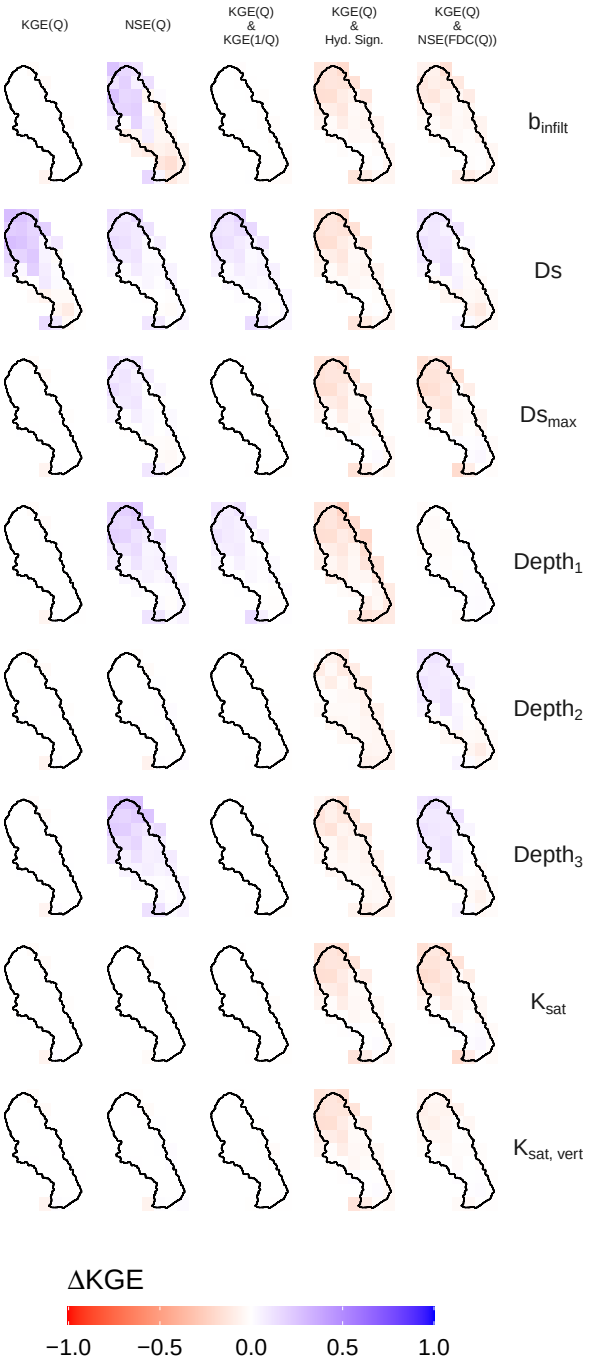

Figure S.20. fSCA performance change (with respect to the case where parameters are spatially constant) at the grid cell scale for basin Cochiguaz at El Peñón during the calibration period (2005-2018). Notice that the benchmark changes among objective functions. Blue colors represent an improvement in the simulated performance, while red colors a decrease.

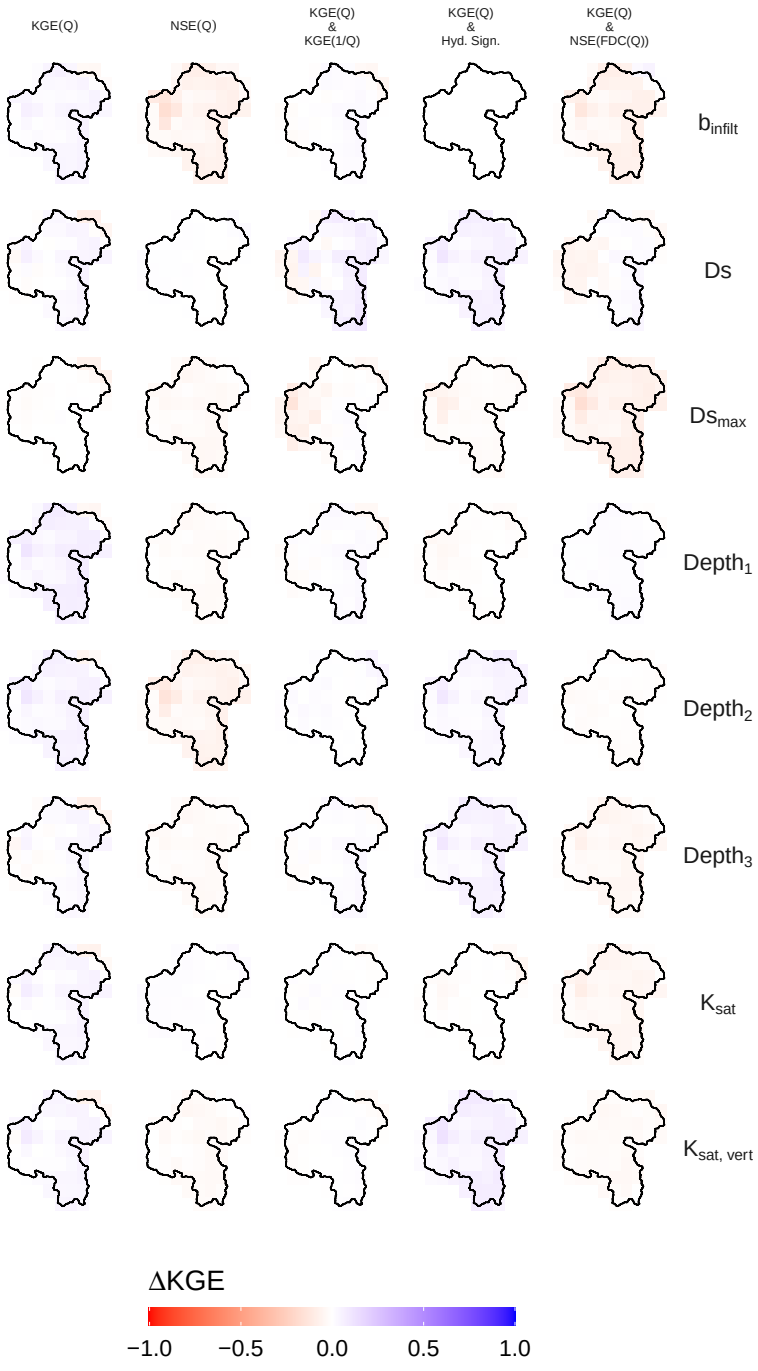

Figure S.21. Same as in Figure S.20, but for the Choapa at Cuncumén basin.

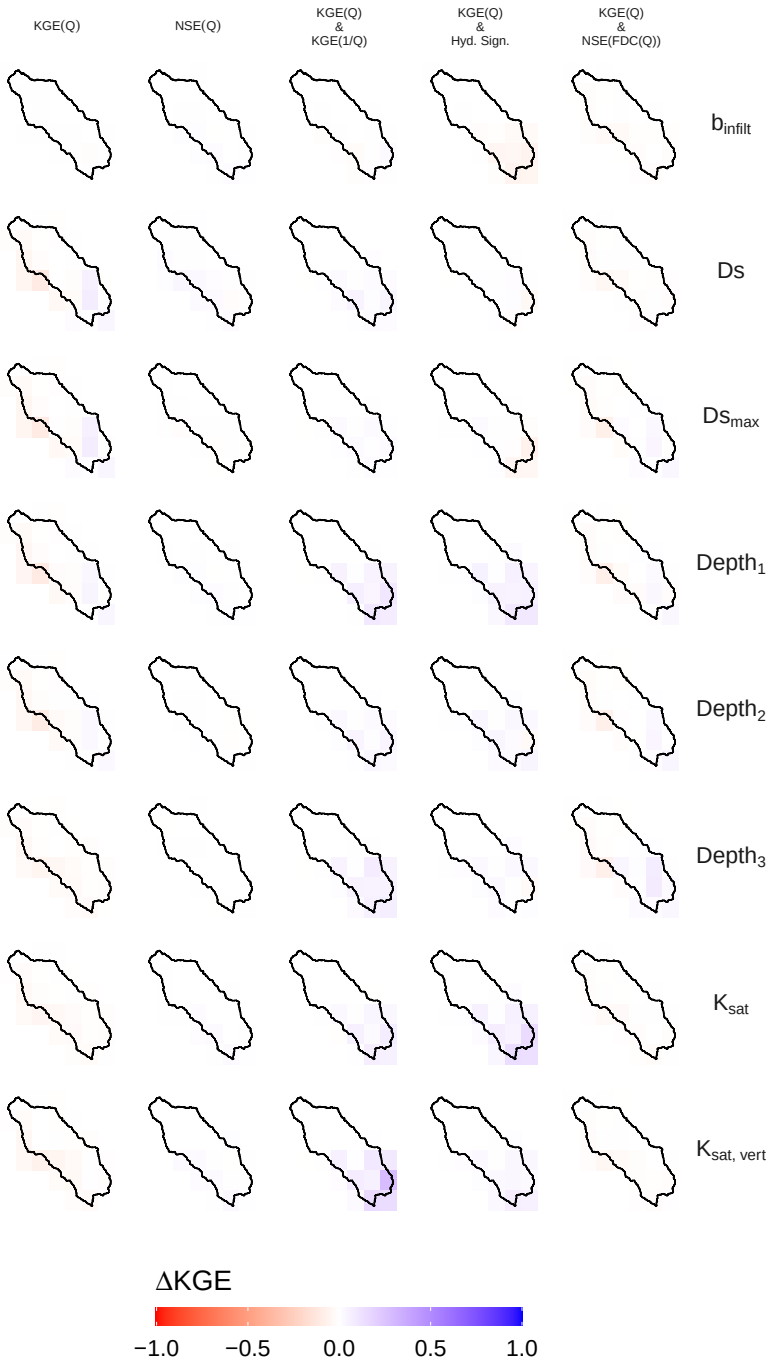

Figure S.22. Same as in Figure S.20, but for the Claro at El Valle basin.

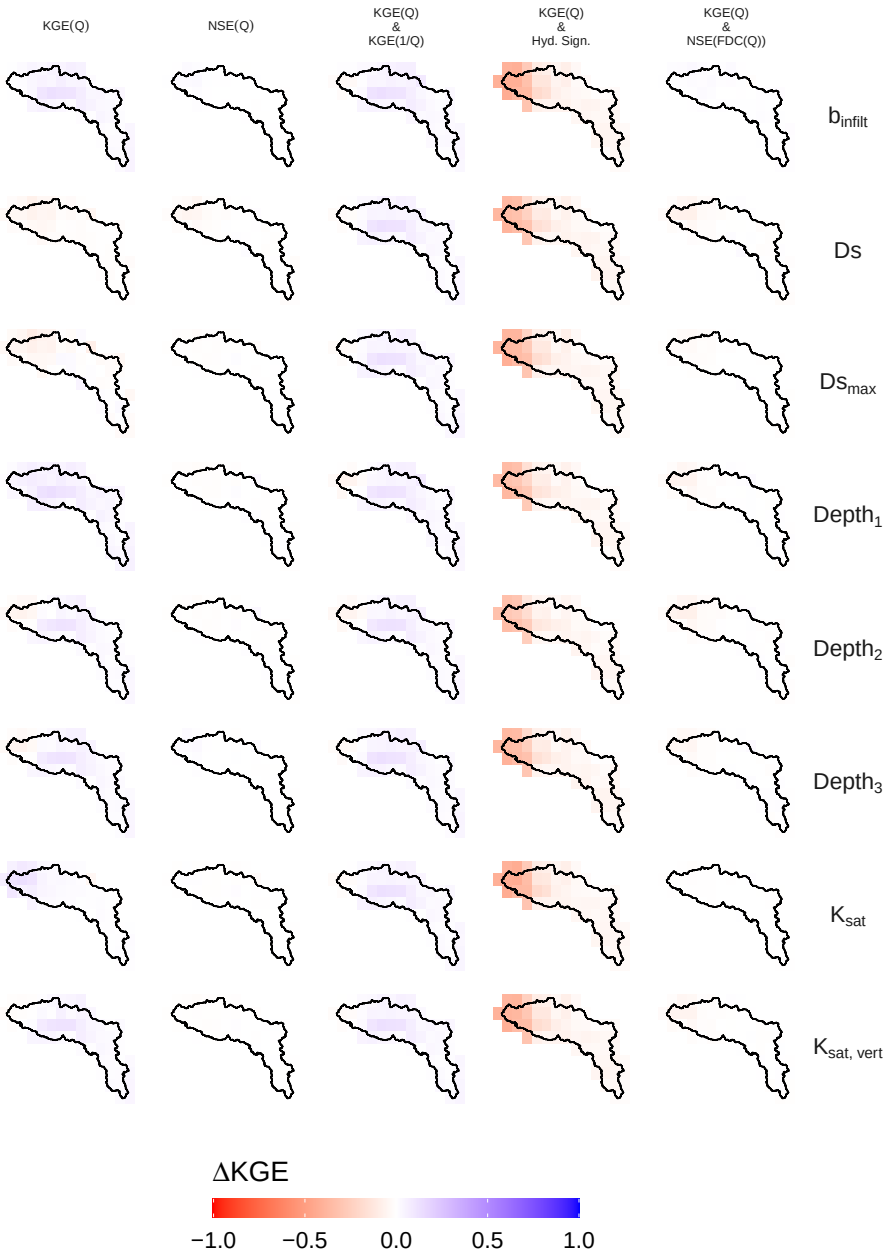

Figure S.23. Same as in Figure S.20, but for the Colorado at Los Palos basin.

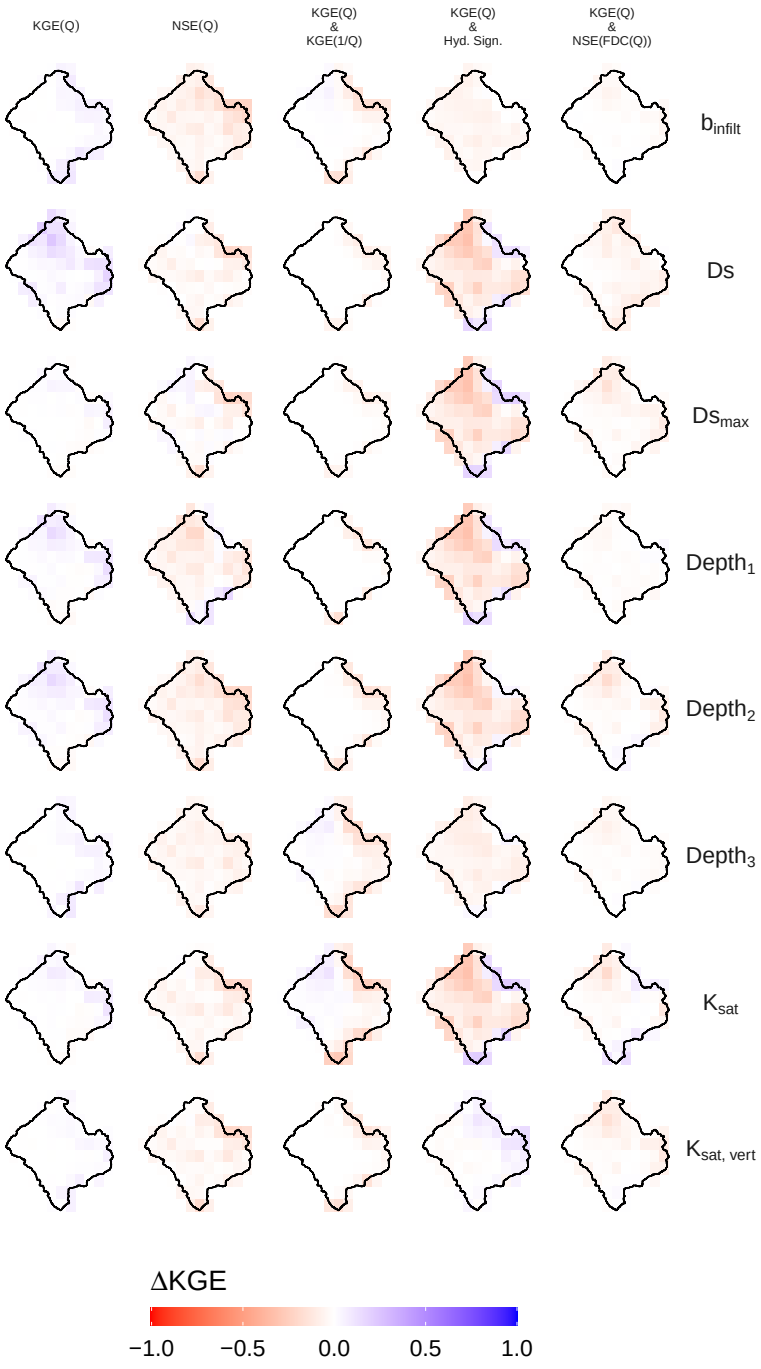

Figure S.24. Same as in Figure S.20, but for the Cautín at Rari-Ruca basin.

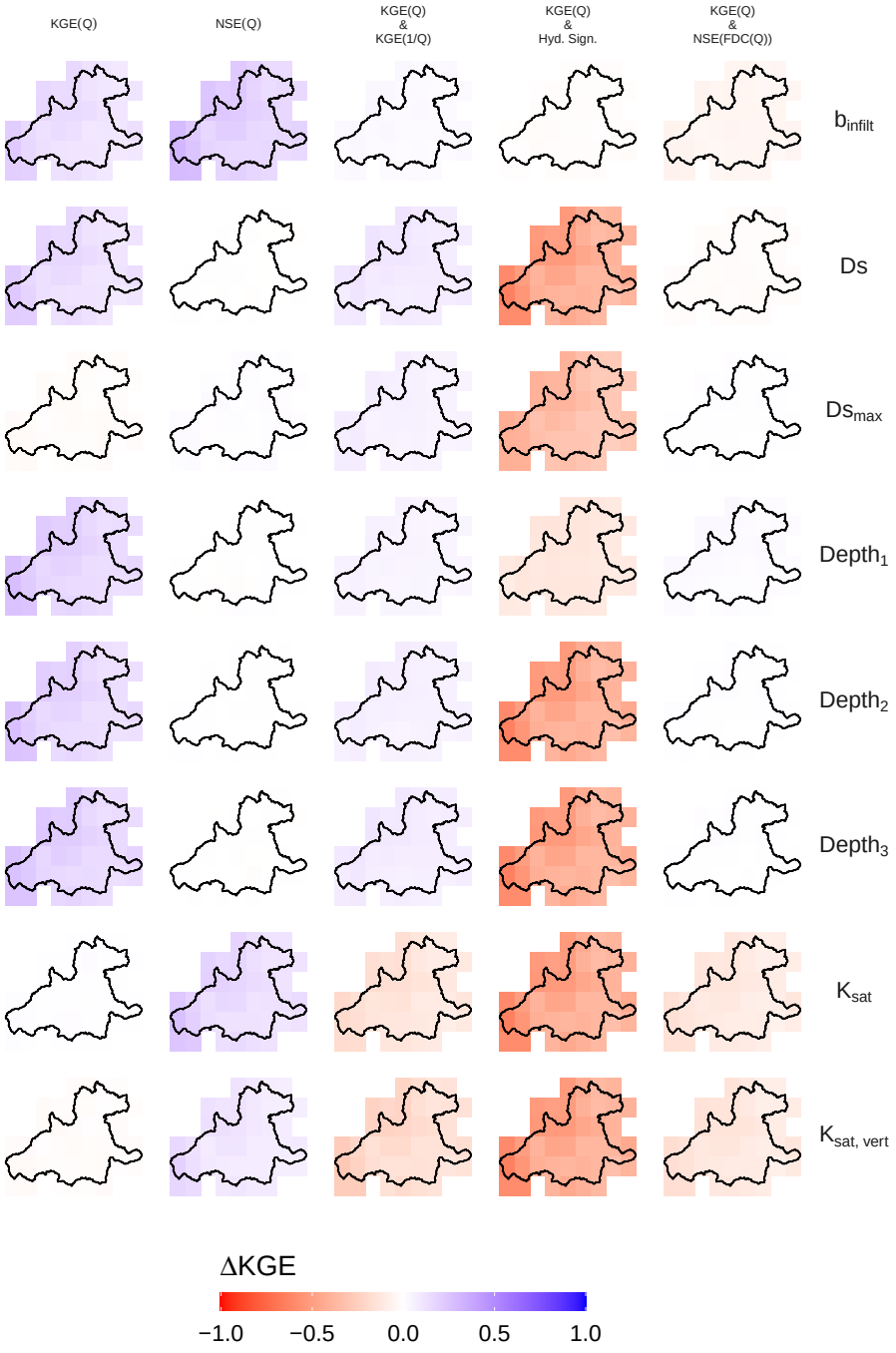

Figure S.25. Same as in Figure S.20, but for the Futa at Tres Chiflones basin.

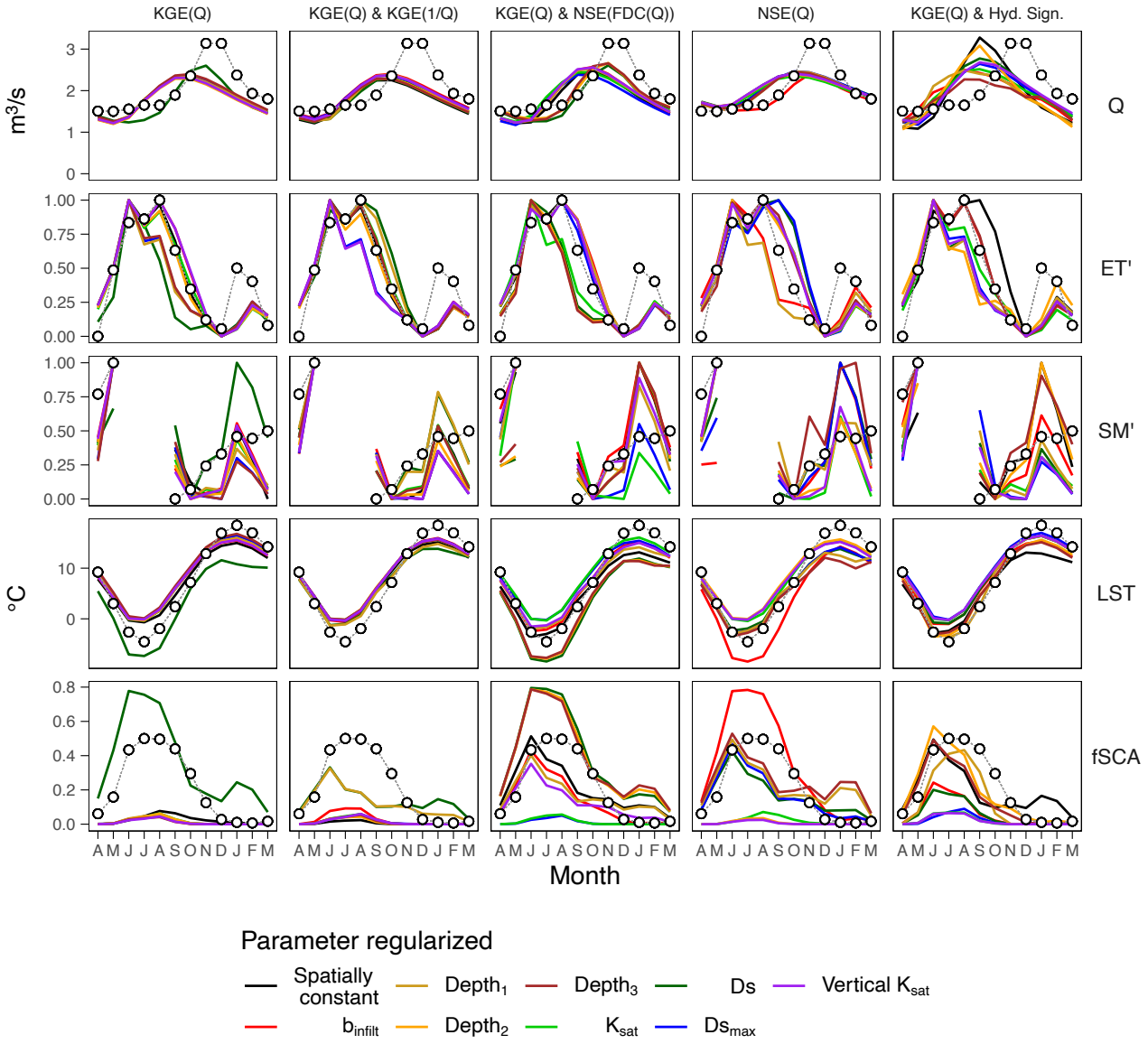

Figure S.26. Mean seasonality at the basin scale of streamflow (Q), normalized evapotranspiration (ET'), normalized soil moisture (SM'_1), land surface temperature (LST), and fractional snow-covered area (fSCA) for the Cochiguaz at El Peñón. Mean monthly values are computed only if there are at least 50 days with information and consider the same days during the calibration period (2005-2018). Results correspond to the best parameter set for each objective function (columns). Notice that winter corresponds to JJA, while summer to DJF. Streamflow observations and remotely sensed variables are referred to as "reference" and symbolized with white dots. The normalization of ET and SM_1 is computed as $X' = (X - X_{min})/(X_{max} - X_{min})$.

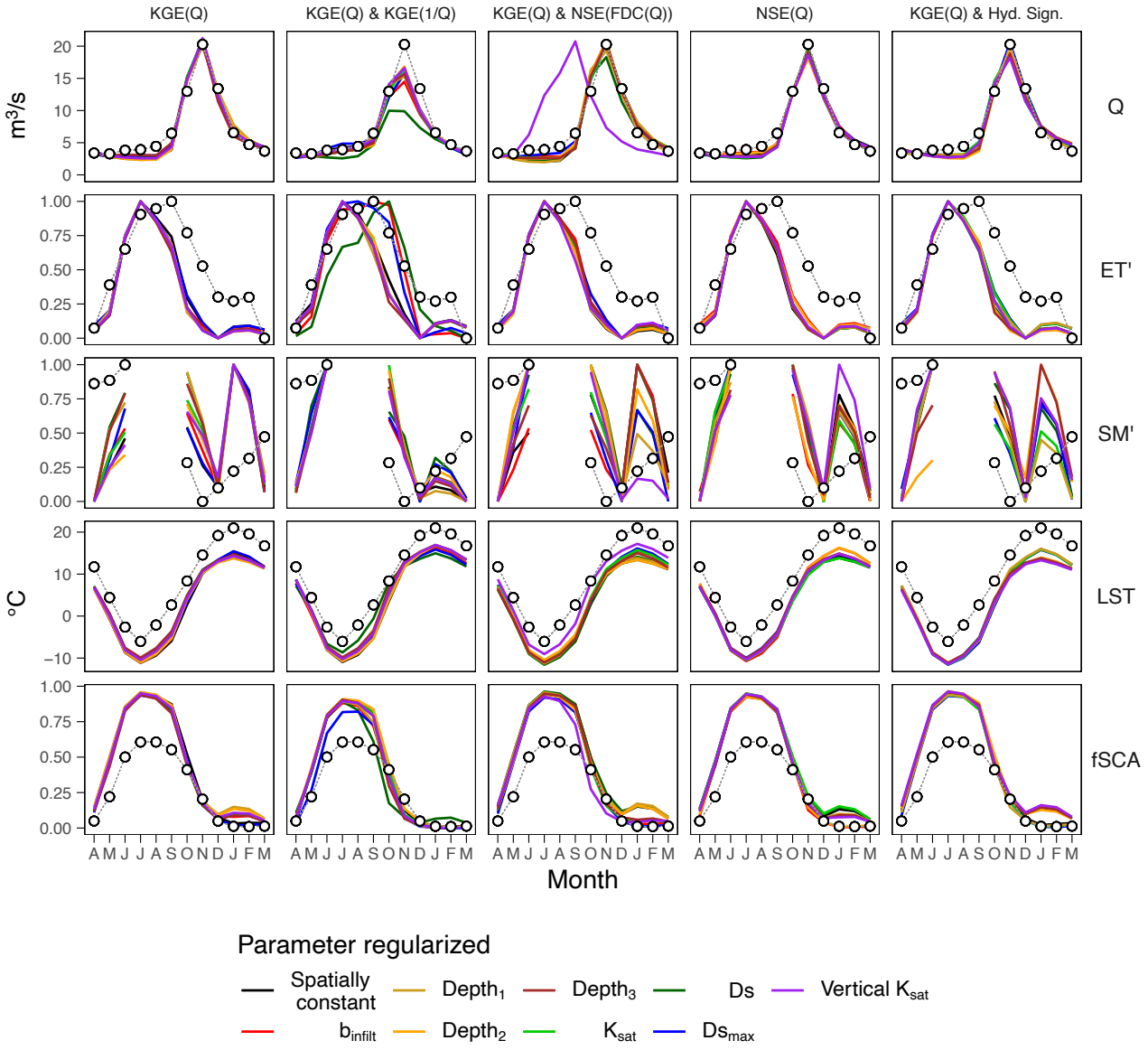

Figure S.27. Same as in Figure S.26, but for the Choapa at Cuncumén basin.

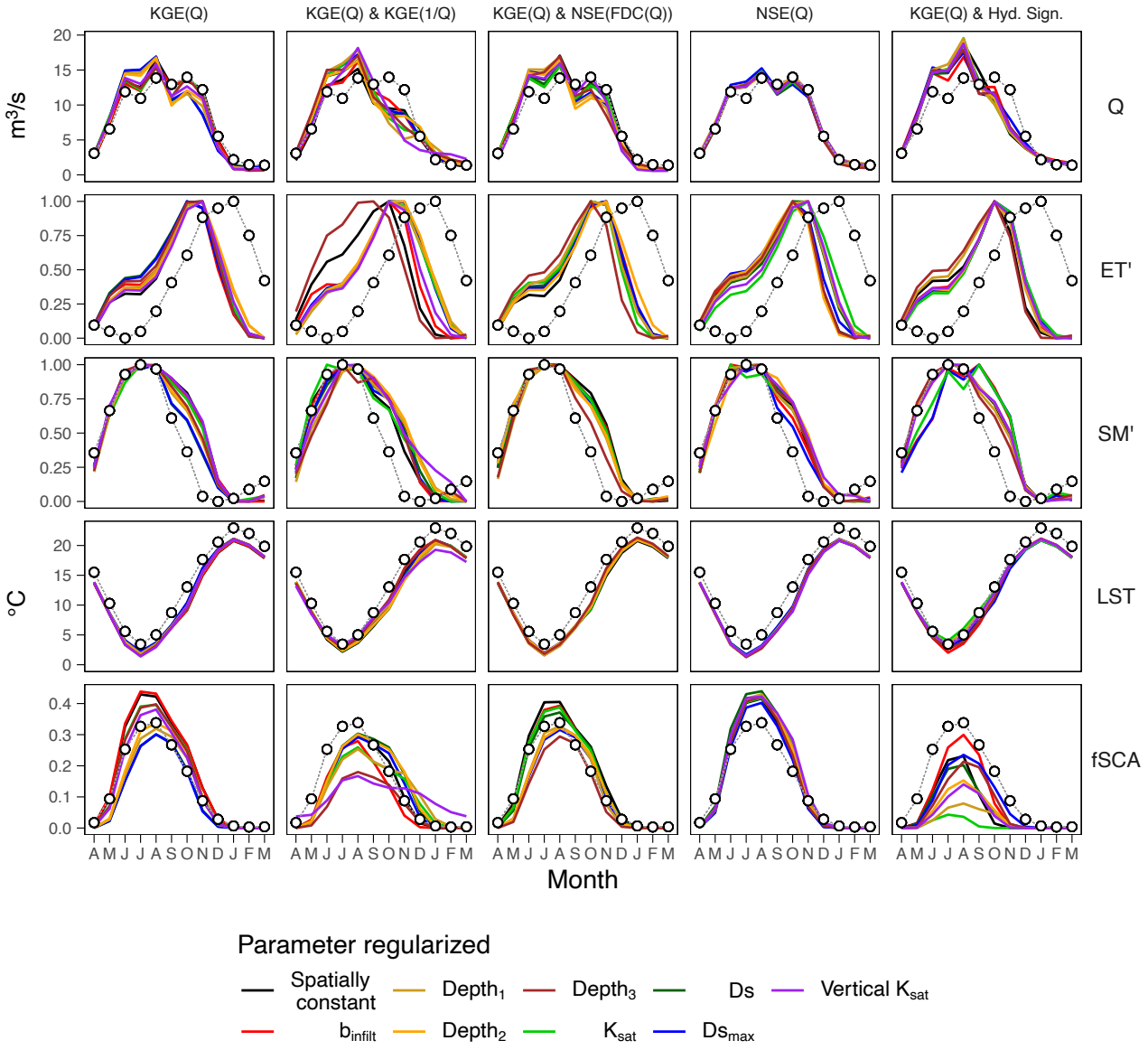

Figure S.28. Same as in Figure S.26, but for the Claro at El Valle basin.

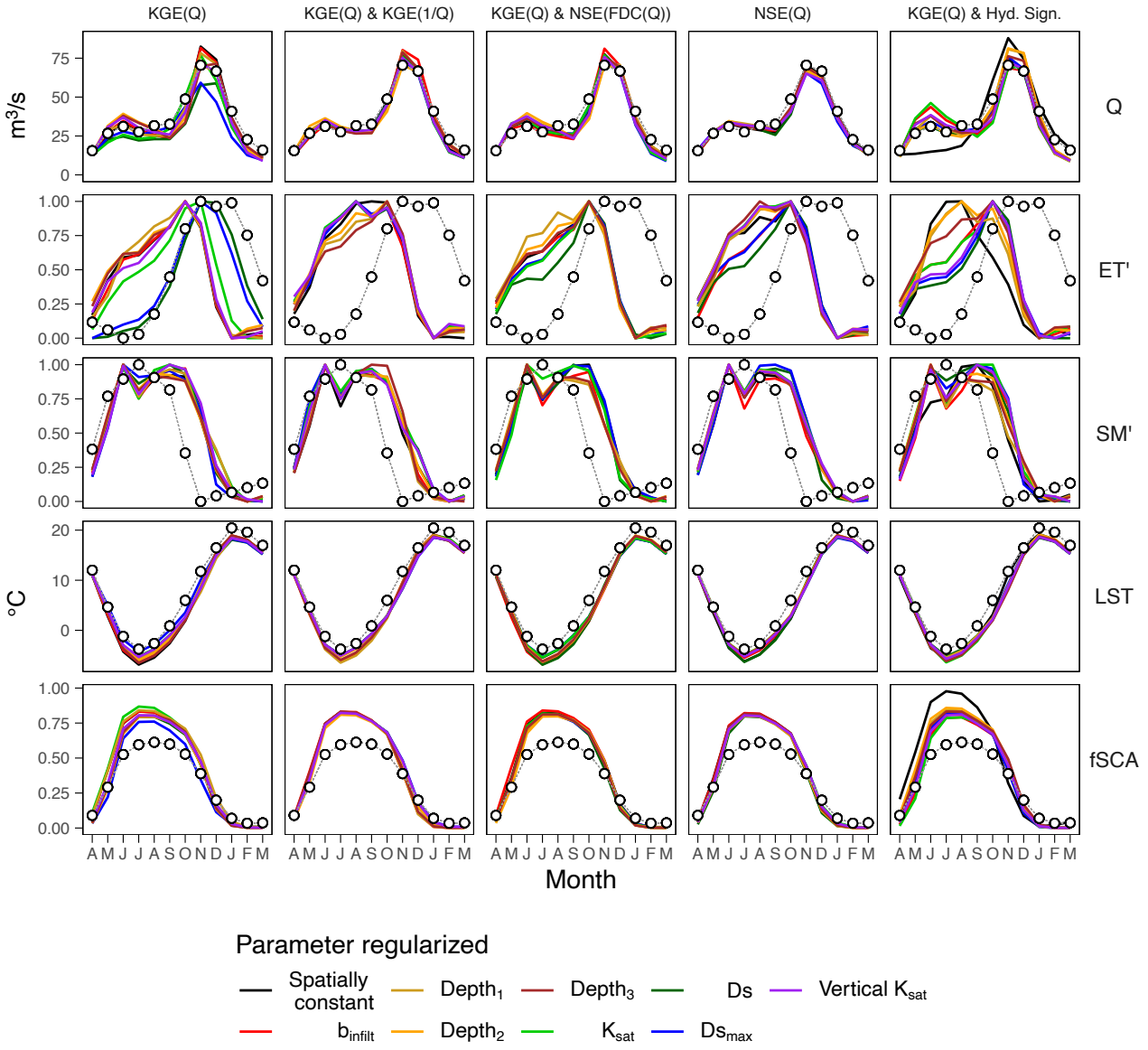

Figure S.29. Same as in Figure S.26, but for the Colorado at Los Palos basin.

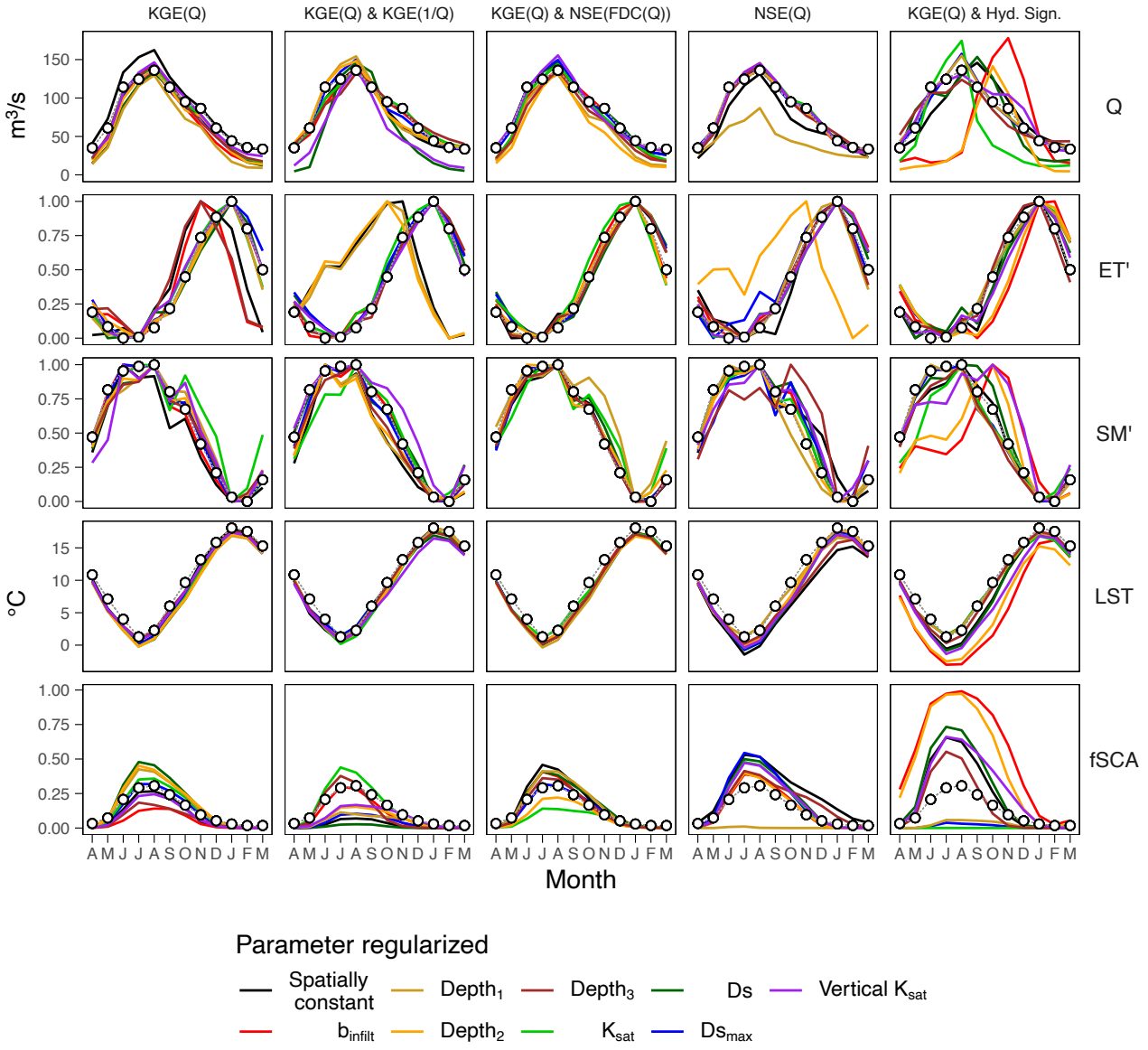

Figure S.30. Same as in Figure S.26, but for the Cautín at Rari-Ruca basin.

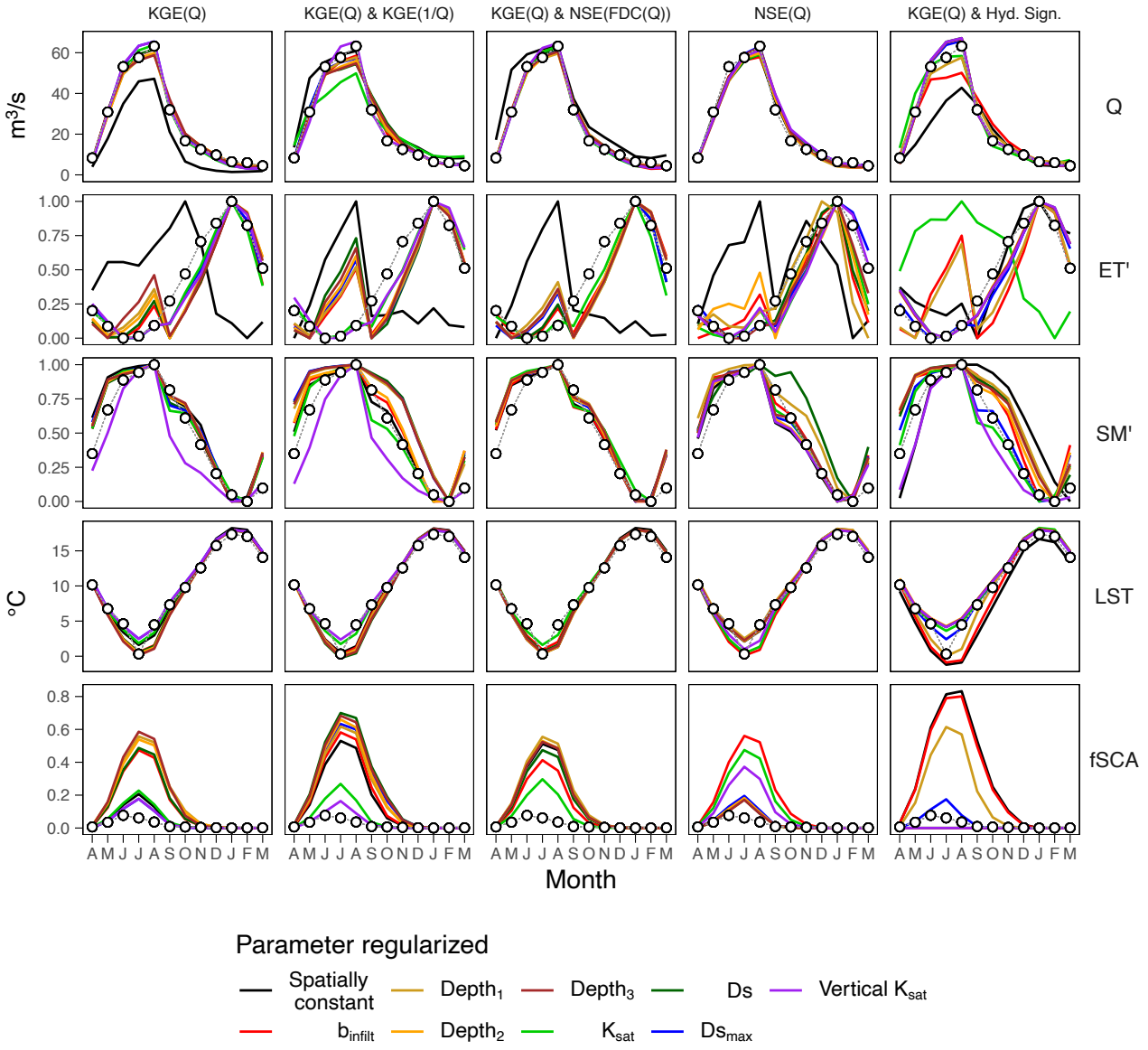

Figure S.31. Same as in Figure S.26, but for the Futa at Tres Chiflones basin.

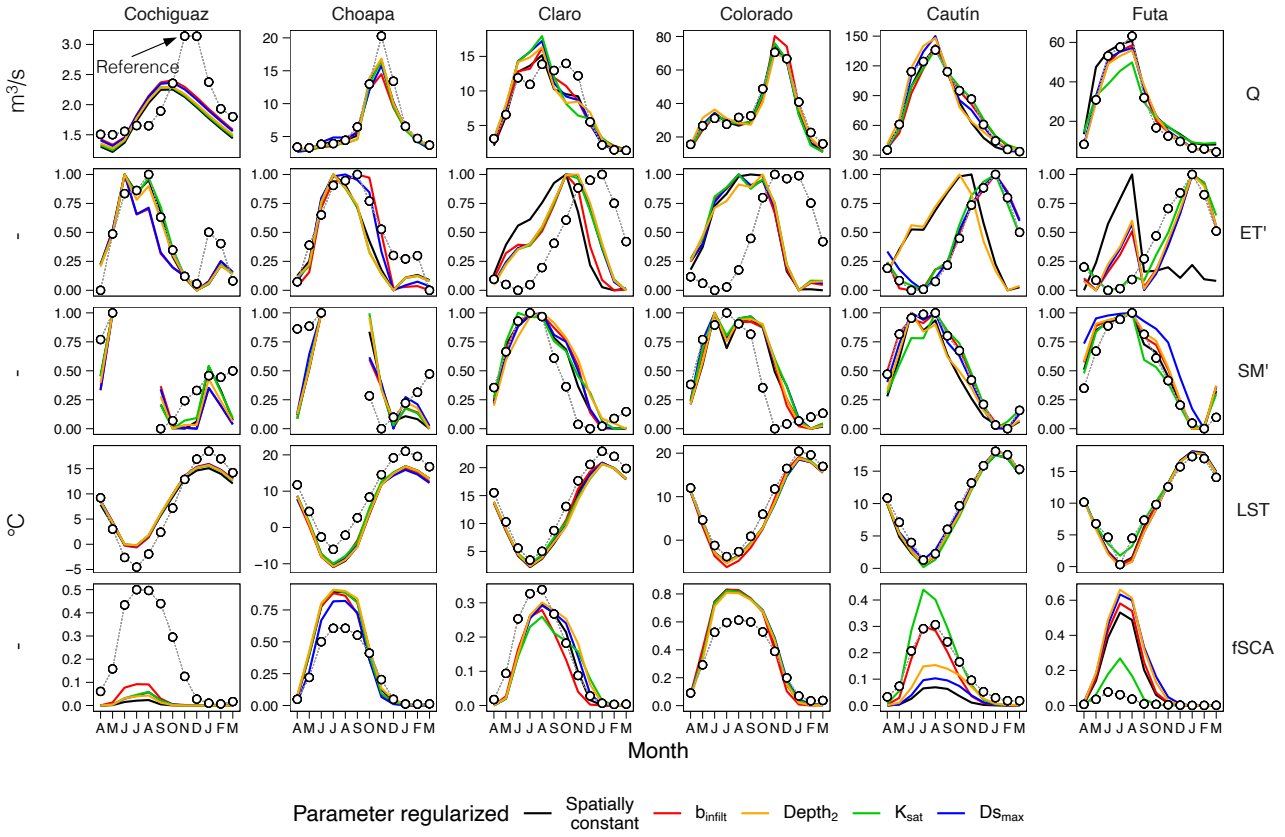

Figure S.32. Summary for the mean seasonality at the basin scale of streamflow (Q), normalized evapotranspiration (ET'), normalized soil moisture (SM'_1), land surface temperature (LST), and fractional snow-covered area (fSCA) for all the basins. Mean monthly values are computed only if there are at least 50 days with information and consider the same days during the calibration period (2005-2018) for the objective function $KGE(Q)$ & $KGE(1/Q)$. Results correspond to the best parameter set for each objective function (columns). Notice that winter corresponds to JJA, while summer to DJF. Streamflow observations and remotely sensed variables are referred to as "reference" and symbolized with white dots. The normalization of ET and SM_1 is computed as $X' = (X - X_{min})/(X_{max} - X_{min})$.

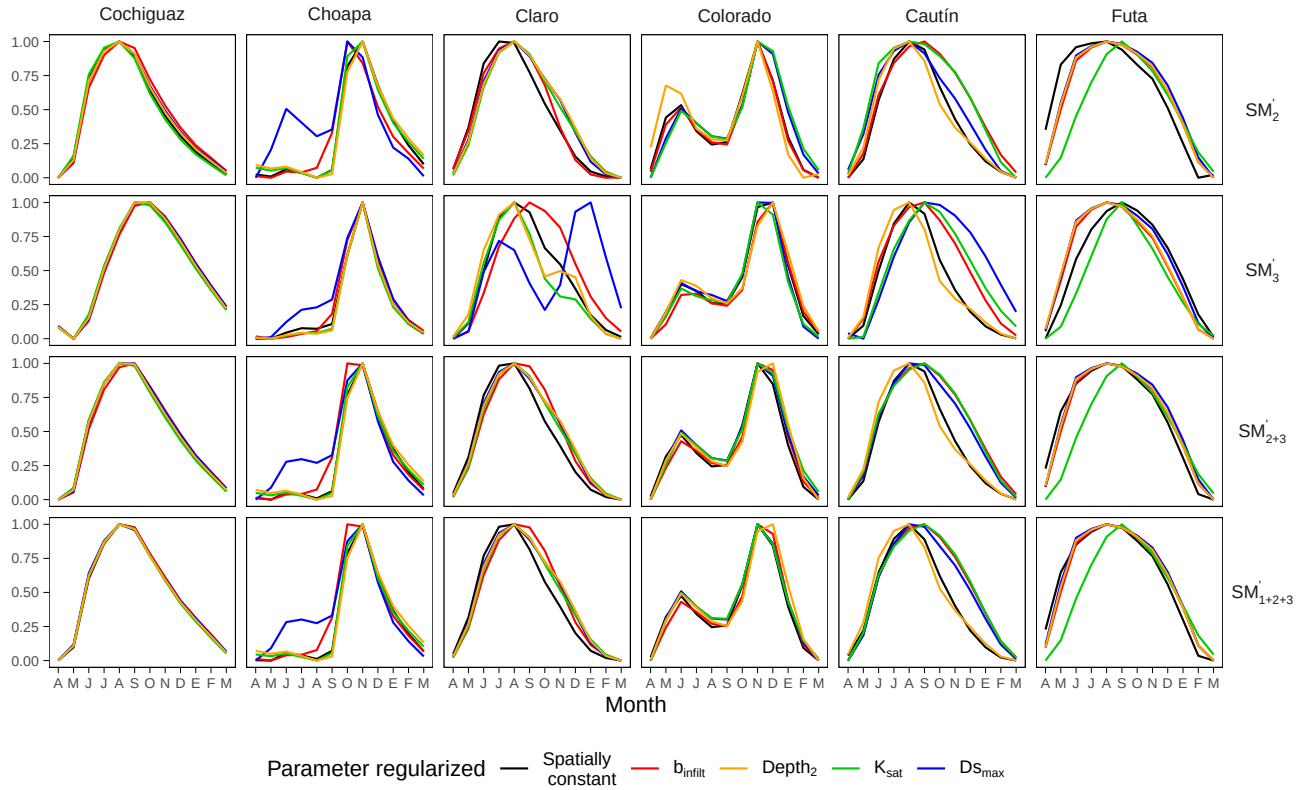

Figure S.33. Mean seasonality at the basin scale of normalized soil moisture (SM'_1) for all the river basins. Mean monthly values are computed for the calibration period (2005-2018). Results correspond to the best parameter set for objective function KGE(Q) & KGE(1/Q). Notice that winter corresponds to JJA, while summer to DJF. The normalization of SM_1 is computed as $X' = (X - X_{min})/(X_{max} - X_{min})$.

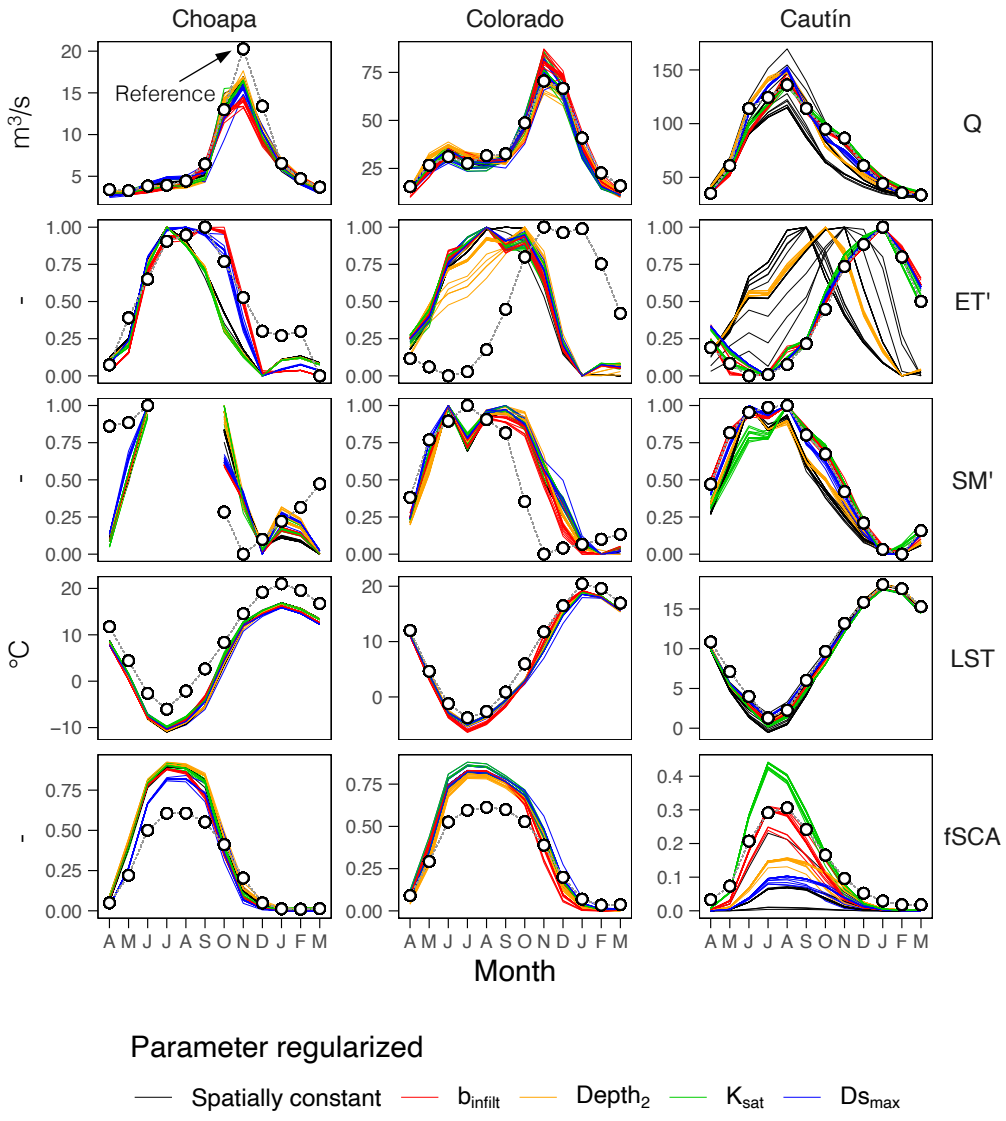

Figure S.34. Catchment-scale annual cycles of streamflow (Q), normalized evapotranspiration (ET'), normalized soil moisture (SM'_1), land surface temperature (LST), and fractional snow-covered area (fSCA) for the Choapa, Colorado, and Cautín River basins during the calibration period (2005–2018). Each line corresponds to a behavioral parameter set selected using one of two criteria: (1) the best 1% of parameter sets ranked by objective function value, or (2) all parameter sets with an objective function value exceeding max[0.5*KGE(Q) + 0.5*KGE(1/Q)] - 0.02. Streamflow observations and remotely sensed variables are referred to as "reference" and symbolized with white dots. The normalization of ET and SM_1 is computed as $X' = (X - X_{min})/(X_{max} - X_{min})$. Notice that winter corresponds to JJA, while summer to DJF.

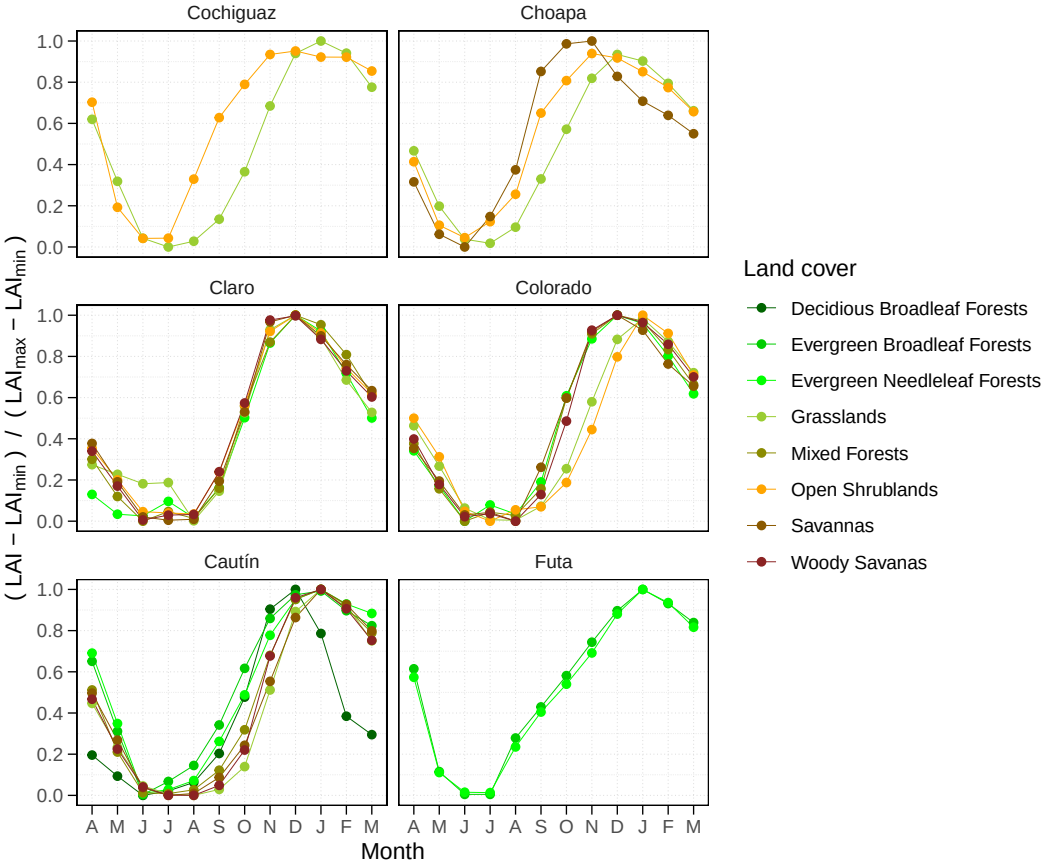

Figure S.35. Mean (dimensionless) LAI values for all the basins and all the land cover considered in each basin. Land cover categories correspond to the IGBP library, from the MODIS product.

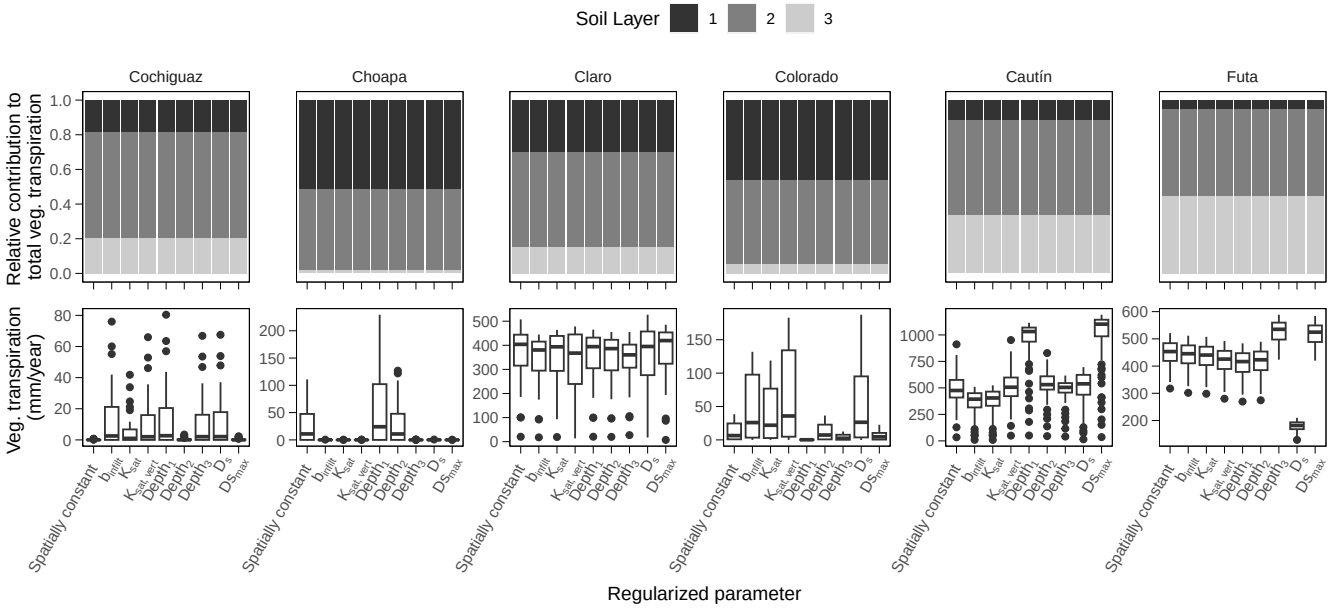

Figure S.36. Contribution of each soil layer to the total transpiration at the basin scale. Note that the root fraction assigned to each soil layer reflects model setup choices.