*Supporting material of*

# Groundwater head responses to droughts across Germany

Pia Ebeling[1], Andreas Musolff[1], Rohini Kumar[2], Andreas Hartmann[3], Jan H. Fleckenstein[1,4]

[1]Department of Hydrogeology, Helmholtz Centre for Environmental Research - UFZ, Leipzig, Germany.
[2]Department of Computational Hydrosystems, Helmholtz Centre for Environmental Research - UFZ, Leipzig, Germany.
[3]Institute of Groundwater Management, Technical University Dresden, Dresden, Germany
[4]Bayreuth Center of Ecology and Environmental Research (BayCEER), University of Bayreuth, Bayreuth, Germany.

*Correspondence to*: Pia Ebeling (pia.ebeling@ufz.de)

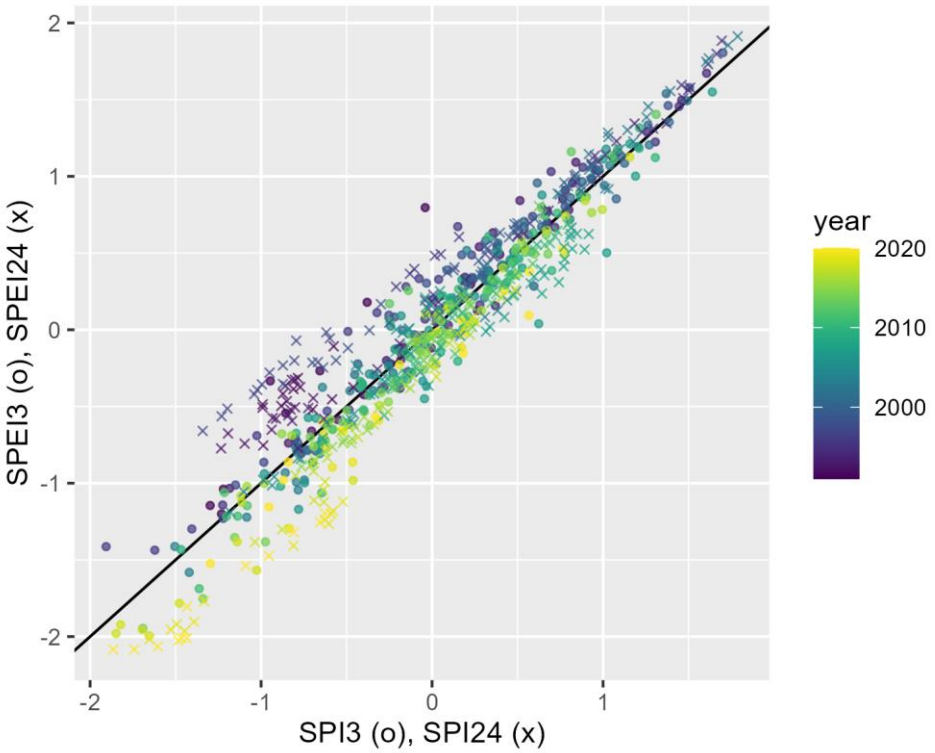


**Figure S1: Differences between average SPI and SPEI for short-term accumulation period of 3 months (circles) and long accumulation of 24 months (crosses) with colors according to the year. It can be seen that SPEI is generally higher at the beginning of the 30-year period and lower at the end, corresponding to negative overall trends.**

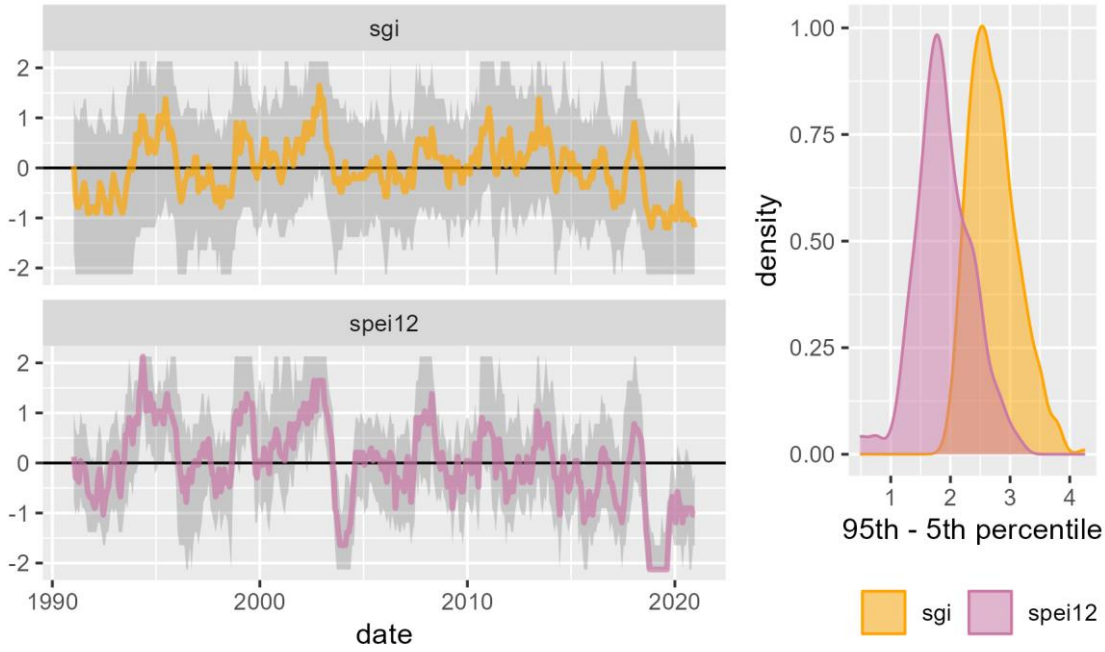


**Figure S2: Median SGI and $SPEI_{12}$ time series across all wells with the band between the $95^{th}$ and $5^{th}$ percentile in gray (left) and density distribution of the bandwidths between the $95^{th}$ and $5^{th}$ percentile (right).**

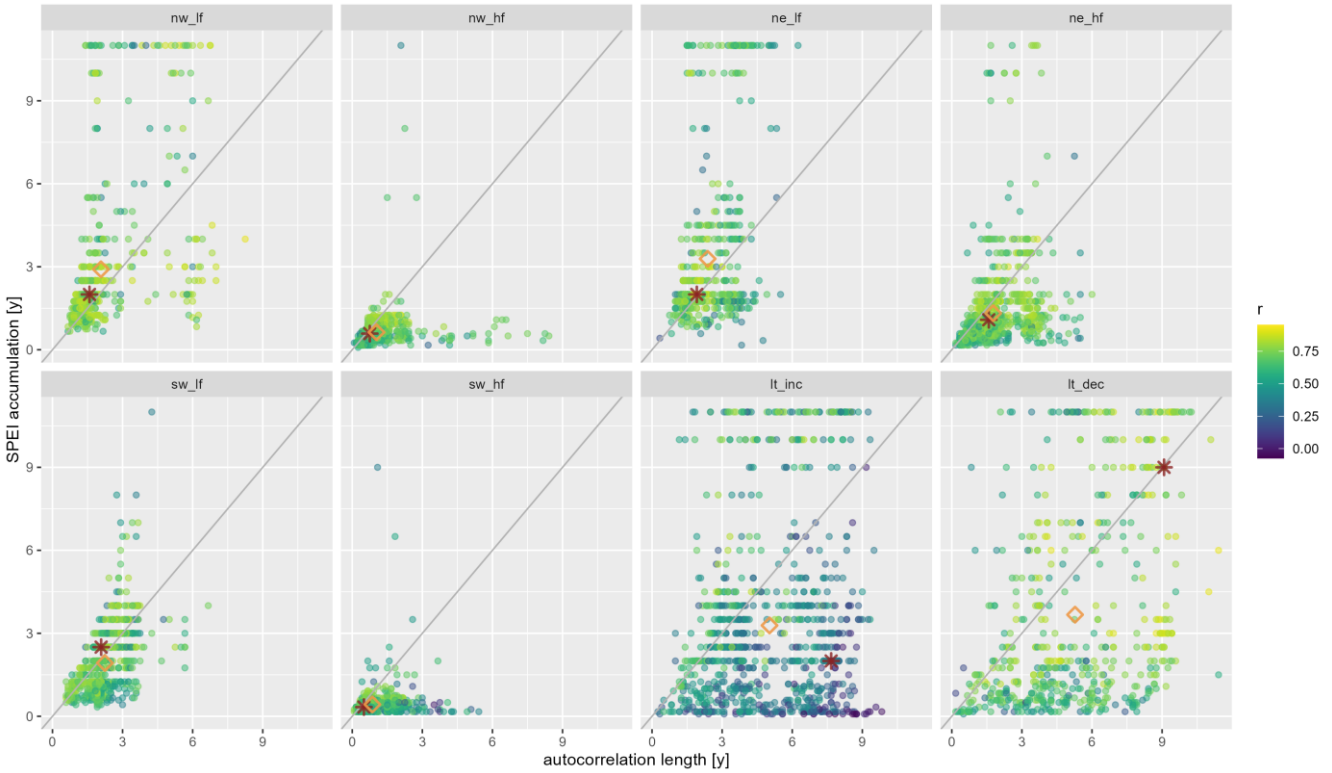


**Figure S3: Groundwater response time scales (SPEI accumulation $acc_{SPEI}$) against autocorrelation length (acf_lag) of individual wells per cluster, of cluster centers (brown star) and mean across cluster members (orange diamond). Colors indicate the cross-correlation coefficients between SGI and $SPEI_{max}$.**

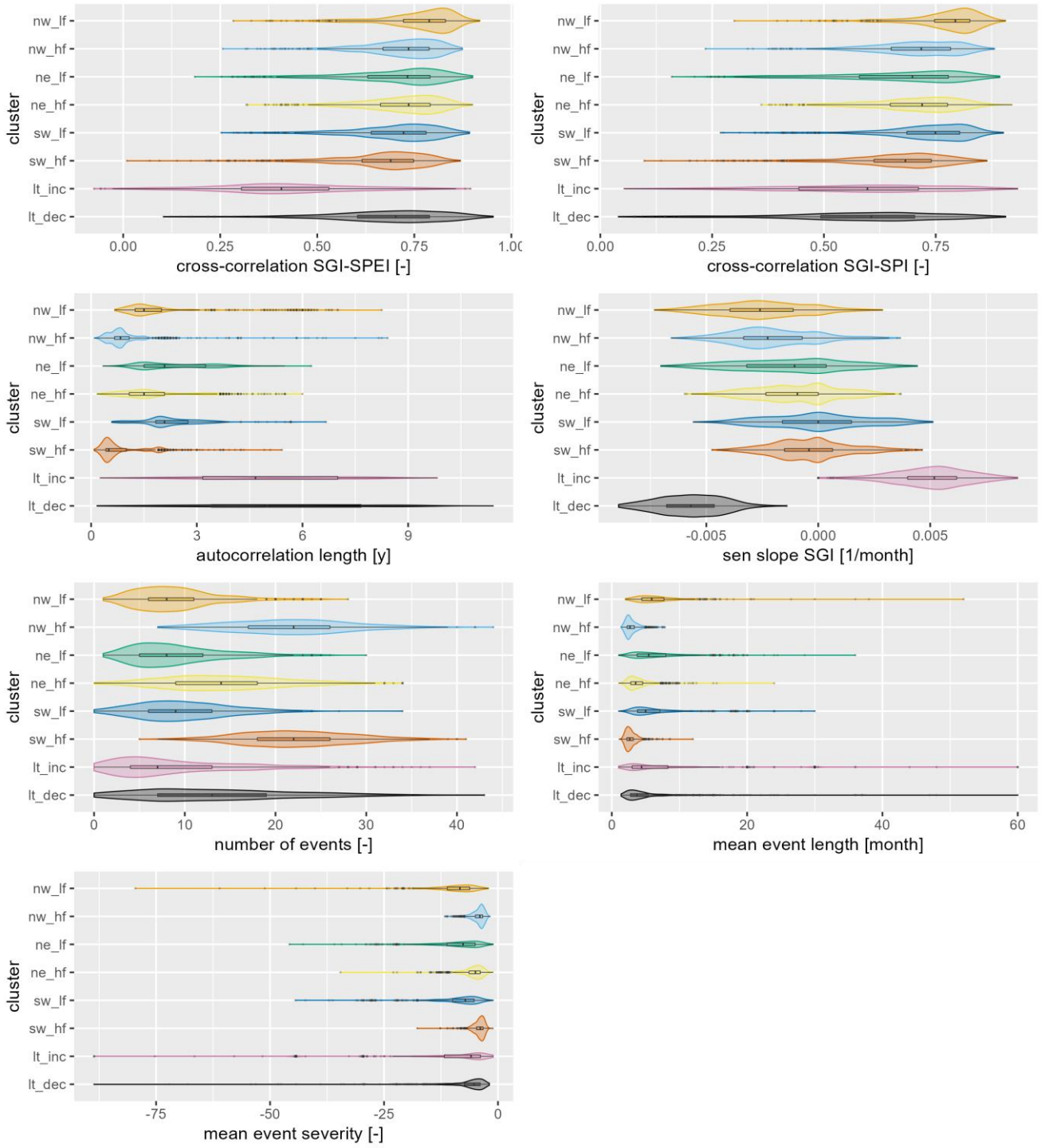


**Figure S4: Distribution of groundwater response characteristics per cluster as violin plots with additional boxplots visualizing summary statistics (median, the 25$^{th}$ and 75$^{th}$ percentiles).**

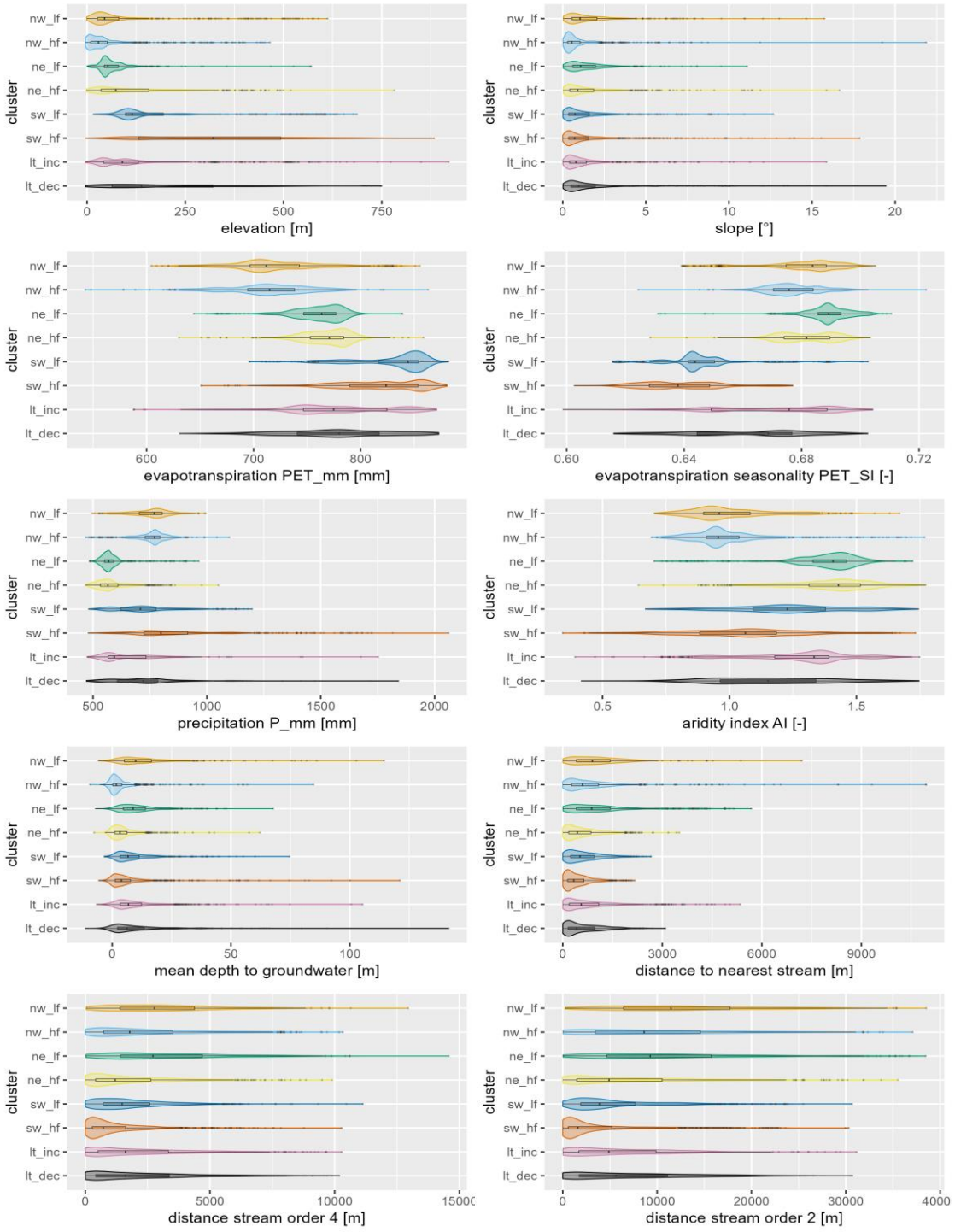


**Figure S5: Distribution of spatial controls per cluster as violin plots with additional boxplots visualizing summary statistics (median, the 25$^{th}$ and 75$^{th}$ percentiles).**

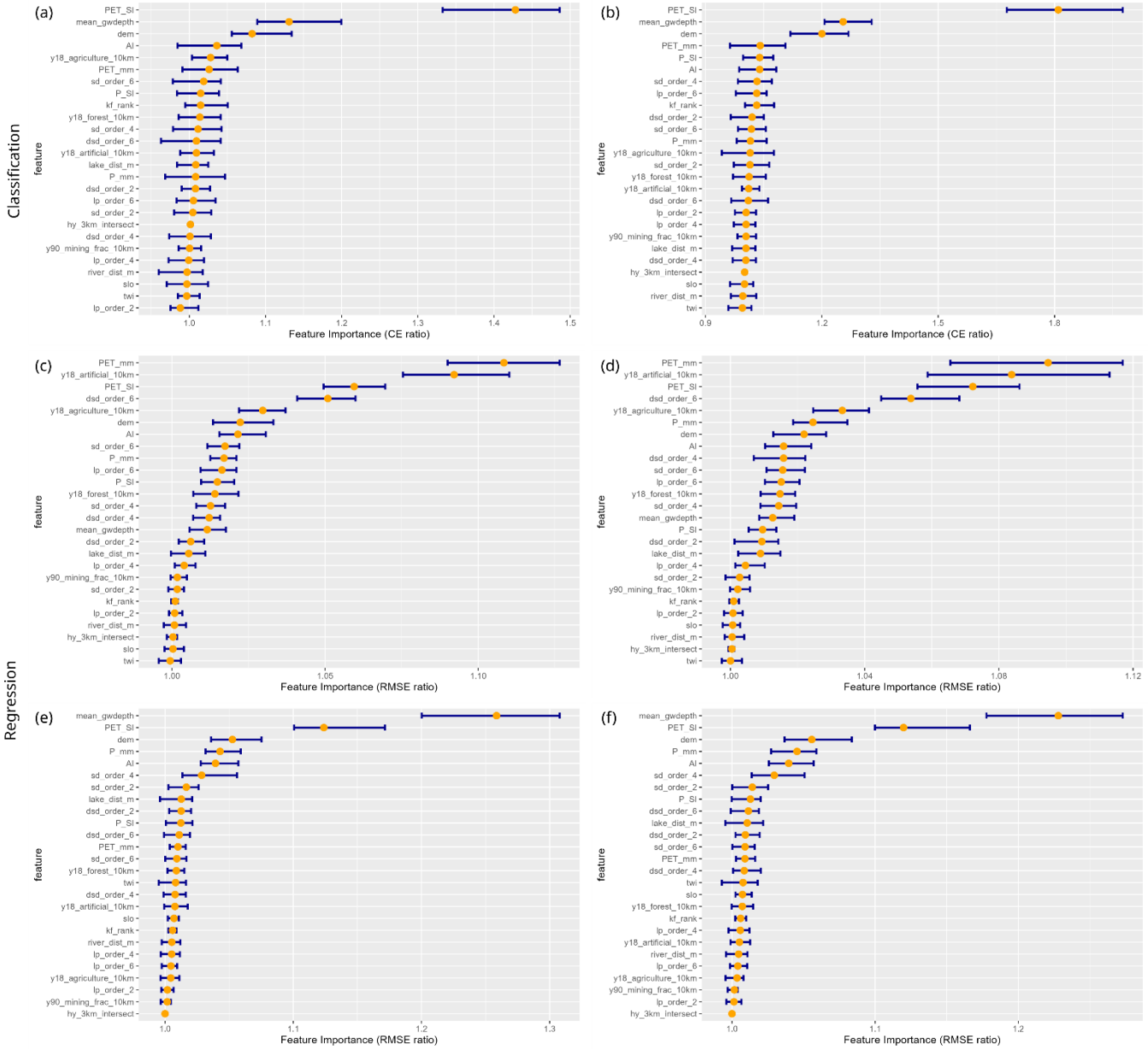


**Figure S6: Feature importance of classification and regression models presented in Table 3 for (a) all 8 clusters, (b) the six regional clusters, (c) the trend in residuals between SGI and $SPEI_{acc}$, (d) the trend in residuals between SGI and $SPI_{acc}$, (e) the response time $resp_{SPEI}$, and (f) the optimal accumulation time $acc_{SPEI}$. Note: The importance for classification is measured by the classification error (CE) with permutation over the CE without permutation of the respective feature, while for regression the root mean square error (RMSE) was used. The orange dot marks the average feature importance calculated as the mean across the five cross-validation resamplings, i.e. the medians from the 10 permutations applied for each resampling (iteration). The blue bar marks the range between the minimum of the 5$^{th}$ percentile and the maximum of the 95$^{th}$ percentile of importance values (percentiles are from the 10 permutations of each iteration).**

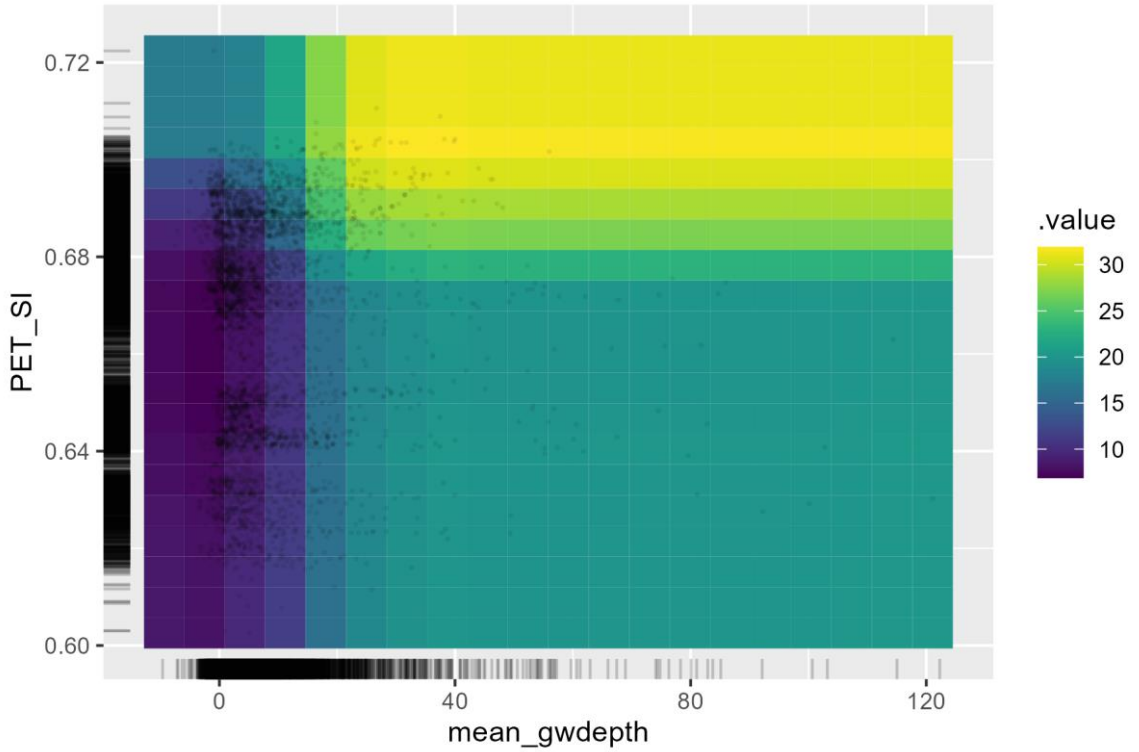


**Figure S7: 2D partial dependence plot (PDP) of the effects of mean_gwdepth and PET_SI on the predicted $resp_{SPEI}$ for the 6-cluster data subset RF regression model. Note: Marginal plots and black dots show the data distribution across the variable space.**

**Table S1: Random forest (RF) results including the three most important features from permutation. Performance is given as mean accuracy for classification and coefficient of variation ($R^2$) for regression models across the cross-validation iterations. Note: All models are shown to complement models (here bold) presented in the main text in Table 3, here bold. Feature importance is given as the mean (across cross-validation iterations) of median importance of the permutation repetitions.**

| RF model | Variable | Number of samples | Performance (accuracy/$R^2$) | Feature | Importance (mean of medians) |
|---|---|---|---|---|---|
| classification | **cluster (all)** | 6620 | **0.68** | PET_SI | 1.43 |
| | | | | mean_gwdepth | 1.13 |
| | | | | dem | 1.08 |
| | **cluster (regional)** | 5120 | **0.79** | PET_SI | 1.81 |
| | | | | mean_gwdepth | 1.25 |
| | | | | dem | 1.20 |
| regression | acf_lag | 6620 | 0.34 | mean_gwdepth | 1.08 |
| | | | | AI | 1.06 |
| | | | | y18_agriculture_10km | 1.05 |
| | $respt_{SPEI}$ | 6620 | 0.35 | mean_gwdepth | 1.20 |
| | | | | AI | 1.07 |
| | | | | PET_SI | 1.05 |
| | $respt_{SPI}$ | 6620 | 0.35 | mean_gwdepth | 1.17 |
| | | | | AI | 1.05 |
| | | | | y18_artificial_10km | 1.05 |
| | $acc_{SPEI}$ | 6620 | 0.35 | mean_gwdepth | 1.18 |
| | | | | AI | 1.07 |
| | | | | P_mm | 1.07 |
| | $acc_{SPI}$ | 6620 | 0.39 | mean_gwdepth | 1.15 |
| | | | | AI | 1.06 |
| | | | | y18_artificial_10km | 1.05 |
| | acf_lag | 5120 | 0.34 | mean_gwdepth | 1.08 |
| | | | | AI | 1.08 |
| | | | | dem | 1.05 |
| | **$respt_{SPEI}$** | 5120 | **0.42** | mean_gwdepth | 1.26 |
| | | | | PET_SI | 1.12 |
| | | | | dem | 1.05 |
| | $respt_{SPI}$ | 5120 | 0.32 | mean_gwdepth | 1.17 |
| | | | | PET_SI | 1.06 |
| | | | | AI | 1.04 |
| | **$acc_{SPEI}$** | 5120 | **0.41** | mean_gwdepth | 1.23 |
| | | | | PET_SI | 1.12 |
| | | | | dem | 1.06 |
| | $acc_{SPI}$ | 5120 | 0.35 | mean_gwdepth | 1.16 |
| | | | | PET_SI | 1.07 |
| | | | | AI | 10.5 |
| | **$resid\_sen_{SPEI}$** | 6620 | **0.42** | PET_mm | 1.11 |
| | | | | y18_artificial_10km | 1.09 |
| | | | | PET_SI | 1.06 |
| | **$resid\_sen_{SPI}$** | 6620 | **0.41** | PET_mm | 1.09 |
| | | | | y18_artificial_10km | 1.08 |
| | | | | PET_SI | 1.07 |