Supplement of Earth Syst. Dynam., 17, 811–827, 2026
https://doi.org/10.5194/esd-17-811-2026-supplement

Earth System Dynamics
Open Access
EGU

Supplement of

Large impact of extreme precipitation on projected blue–green water partitioning

Simon P. Heselschwerdt et al.

Correspondence to: Simon P. Heselschwerdt (simon.heselschwerdt@hereon.de)

S1 CMIP6 ensemble overview

Table S1. Overview of the 12 CMIP6 ESMs used in this study, including the land model, vegetation representation, member ID, and reference.

No.	Institution	Model	Land model	Member ID	Veg. repr.	Reference
1	BCC	BCC-CSM2-MR†	BCC_AVIM2.0	r1i1p1f1	Prog. LAI	Wu et al. (2019)
2	CCCMa	CanESM5	CLASS3.6 + CTEM1.2	r1i1p1f1	Prog. LAI	Swart et al. (2019)
3	NCAR	CESM2	CLM5	r11i1p1f1	Prog. LAI	Danabasoglu et al. (2020)
4	CMCC	CMCC-ESM2	CLM4.5 (BGC)	r1i1p1f1	Prog. LAI	Lovato et al. (2022)
5	CNRM/CERFACS	CNRM-ESM2-1	ISBA + CTRIP	r1i1p1f2	Prog. LAI	Séférian et al. (2019)
6	EC-Earth Consortium	EC-Earth3-Veg	HTESSEL + LPJ-GUESS	r1i1p1f1	Dyn. veg.	Döscher et al. (2022)
7	NOAA-GFDL	GFDL-ESM4	GFDL-LM4.1	r1i1p1f1	Dyn. veg.	Dunne et al. (2020)
8	IPSL	IPSL-CM6A-LR	ORCHIDEE	r1i1p1f1	Prog. LAI	Boucher et al. (2020)
9	MIROC	MIROC-ES2L†	MATSIRO6 + VISIT-e	r1i1p1f2	Prog. LAI	Hajima et al. (2020)
10	MPI-M	MPI-ESM1-2-LR	JSBACH3.2	r11i1p1f1	Dyn. veg.	Mauritsen et al. (2019)
11	NCC	NorESM2-MM†	CLM5	r1i1p1f1	Prog. LAI	Seland et al. (2020)
12	MOHC	UKESM1-0-LL	JULES-ES-1.0	r1i1p1f2	Dyn. veg.	Sellar et al. (2019)

Veg. repr.: Prog. LAI = prognostic/time-varying LAI and phenology; Dyn. veg. = prognostic LAI with interactive vegetation cover, competition, or demography. †Model excluded from the 9-model land-cover sensitivity subset; the subset includes only models for which tree and crop fraction changes were available.

S2 Predictor screening and sensitivity tests

To support the selection of the final common predictor set, we first performed a screening analysis using a larger candidate predictor pool (Table S2). For each historical BGWS regime, we fitted repeated spatially blocked Elastic Net models with nested grouped cross-validation and a one-standard-error rule, using the same train–test design as in the main analysis. From these runs, we summarised each predictor by its selection frequency across repeated models and by the consistency of the sign of its non-zero coefficients. This screening was used to assess the robustness and interpretability of candidate predictors before defining the final common predictor set. The approach follows earlier work showing that predictor selection in spatial environmental models should be evaluated under spatially structured cross-validation, and that repeated cross-validation and variable preselection can help identify parsimonious and robust predictor sets (Meyer et al., 2019; Laimighofer et al., 2022). Selection frequency and sign consistency were used as screening diagnostics rather than as strict selection rules. The final set was defined by combining these diagnostics with physical interpretability and predictor availability across the full 12-model ensemble.

We assessed both total-column and near-surface soil moisture during predictor screening. Near-surface soil moisture was retained in the final predictor set because it was selected more consistently than total-column soil moisture (Table S2). For total-column soil moisture, relative changes were computed separately for each ESM as future-minus-historical change normalised by the model's historical baseline, before averaging across models to form the ensemble mean, because the effective soil-column depth represented by total-column soil moisture differs across models. Total cloud cover was retained because it was selected consistently in the blue water regime.

We further performed targeted sensitivity analyses (Table S3). First, we tested ΔRX1day as an alternative extreme-precipitation metric. With the final analysis mask, replacing ΔRX5day with ΔRX1day yields similar out-of-sample performance and leaves extreme precipitation as the leading predictor in both regimes, with ΔRX1day ranking first (Table S3). This confirms that the dominant role of extreme precipitation is robust to the choice of extreme-precipitation metric. We retain ΔRX5day in the main analysis because multi-day precipitation accumulation is a more robust large-scale indicator at coarse model resolution as it captures better saturation-excess runoff generation, the predominant mechanism in CMIP6 land schemes (Hou et al., 2023), whereas single-day extremes are more strongly affected by drizzle bias and coarse-resolution underestimation in CMIP6 ESMs (Brunner et al., 2025).

Second, we assessed land-cover predictors in a reduced 9-model subset for which tree and crop fraction changes were available. This test should be interpreted relative to the reduced-ensemble baseline, because the 9-model subset already differs from the full 12-model ensemble in its predictor ranking (Table S3). Within this reduced ensemble, adding ΔTreeFrac and ΔCropFrac changes some secondary ranking positions but does not overturn the leading regime-specific controls: ΔRX5day remains the leading predictor in the blue regime, and $\Delta\mathrm{VPD}_{\mathrm{seas}}$ remains the leading predictor in the green regime. ΔTreeFrac enters the top three in the blue regime and ranks 6th in the green regime, while ΔCropFrac contributes little (10th blue, 8th green). The remaining ranking shifts likely reflect a combination of the smaller 9-model ensemble and the correlated structure among vegetation-related predictors. In particular, ΔLAI and ΔTreeFrac are strongly correlated in the reduced ensemble (Pearson's $r = 0.70$), indicating that part of the vegetation-related signal may be shared between them and redistributed in the Elastic Net regression. Because tree and crop fraction changes are unavailable for the full 12-model ensemble and do not materially alter the broader process interpretation, we retain the 12-model fixed predictor set in the main analysis.

Third, we tested the spatial attribution restricted to grid cells with high inter-model sign agreement in ΔBGWS ($\geq$8/12 ESMs). This subset comprises $\sim$25 % of the analysed domain (Fig. S1), with predictive skill of $R^2 = 0.57$ in the blue and $R^2 = 0.85$ in the green regime. The blue-regime ranking is essentially unchanged, with ΔRX5day strongly dominant. In the green regime, $\Delta\mathrm{VPD}_{\mathrm{seas}}$ and ΔVPD become the leading predictors, with ΔLAI third and ΔRX5day dropping to rank 5. The high-confidence cells cluster in specific climate zones: boreal and monsoonal/tropical wet regions in the blue regime, and warm-temperate and water-limited zones in the green regime. The greater importance of atmospheric dryness in the high-confidence green subset is consistent with established findings that vegetation responses are particularly sensitive to VPD in water-limited and transitional regimes (Seneviratne et al., 2010; Denissen et al., 2022). The high-confidence subset attribution should therefore be read as a regime-conditional finding for these specific climate zones rather than as a revised regime-wide attribution.

Table S2. Predictor screening results from repeated spatially blocked Elastic Net models using the full candidate predictor pool. Selection frequency gives the fraction of repeated models in which a predictor was retained with a non-zero coefficient. Sign consistency gives the mean sign of non-zero coefficients across repeats, where values close to $+1$ or -1 indicate a stable coefficient direction. The final-set column indicates whether the predictor was retained in the common 12-model fixed predictor set used in the main analysis. Abbreviations are as follows: P, mean precipitation; RX1day, annual maximum 1-day precipitation; RX5day, annual maximum consecutive 5-day precipitation; P_{seas}, precipitation seasonality; VPD, vapour pressure deficit; VPD$_{seas}$, VPD seasonality; T, near-surface air temperature; RSDS, surface downwelling shortwave radiation; CLT, total cloud cover; SM, total-column soil moisture; SM$_{surf}$, near-surface soil moisture; LAI, leaf area index; GPP, gross primary productivity; WUE, water-use efficiency.

Predictor	Blue water regime		Green water regime		Final
	Sel. freq.	Sign cons.	Sel. freq.	Sign cons.	set
ΔP	0.78	1.00	1.00	1.00	yes
ΔRX1day	1.00	1.00	1.00	1.00	no
ΔRX5day	0.53	1.00	1.00	1.00	yes
ΔP_{seas}	0.98	1.00	0.45	1.00	yes
ΔVPD	0.75	1.00	1.00	1.00	yes
ΔVPD$_{\mathrm{seas}}$	0.63	-1.00	1.00	-1.00	yes
ΔT	0.55	-1.00	1.00	-1.00	no
ΔRSDS	0.50	-1.00	0.65	1.00	no
ΔCLT	0.98	1.00	0.15	1.00	yes
ΔSM	0.93	1.00	0.40	1.00	no
ΔSM$_{\mathrm{surf}}$	1.00	1.00	0.98	1.00	yes
ΔLAI	1.00	-1.00	1.00	-1.00	yes
ΔGPP	0.38	-1.00	0.10	-0.50	no
ΔWUE	1.00	1.00	1.00	1.00	yes

Table S3. Performance and predictor rankings for the fixed predictor-set attribution models and the sensitivity tests. $\overline{R^2}$ is the mean out-of-sample R^2 from repeated spatially blocked train–test splits. The three most important predictors (by ensemble-mean median permutation importance) are listed per regime in decreasing order.

Setup	ESMs	$\overline{R^2}_{\mathrm{blue}}$	$\overline{R^2}_{\mathrm{green}}$	Top-3 blue	Top-3 green
Main analysis	12	0.51	0.74	RX5day, LAI, $\mathrm{SM_{surf}}$	RX5day, LAI, $\mathrm{VPD_{seas}}$
Robust subset ($\geq$8/12 ESMs)	12	0.57	0.85	RX5day, $\mathrm{SM_{surf}}$, LAI	$\mathrm{VPD_{seas}}$, VPD, LAI
ΔRX1day instead of ΔRX5day	12	0.54	0.81	RX1day, LAI, $\mathrm{SM_{surf}}$	RX1day, LAI, $\mathrm{VPD_{seas}}$
Reduced ensemble	9	0.50	0.72	RX5day, LAI, CLT	$\mathrm{VPD_{seas}}$, LAI, RX5day
+ ΔTreeFrac	9	0.50	0.73	RX5day, CLT, TreeFrac	$\mathrm{VPD_{seas}}$, RX5day, P
+ ΔTreeFrac + ΔCropFrac	9	0.50	0.74	RX5day, CLT, TreeFrac	$\mathrm{VPD_{seas}}$, RX5day, P

ΔTreeFrac ranks 3rd (blue) and 6th (green); ΔCropFrac ranks 10th (blue) and 8th (green).

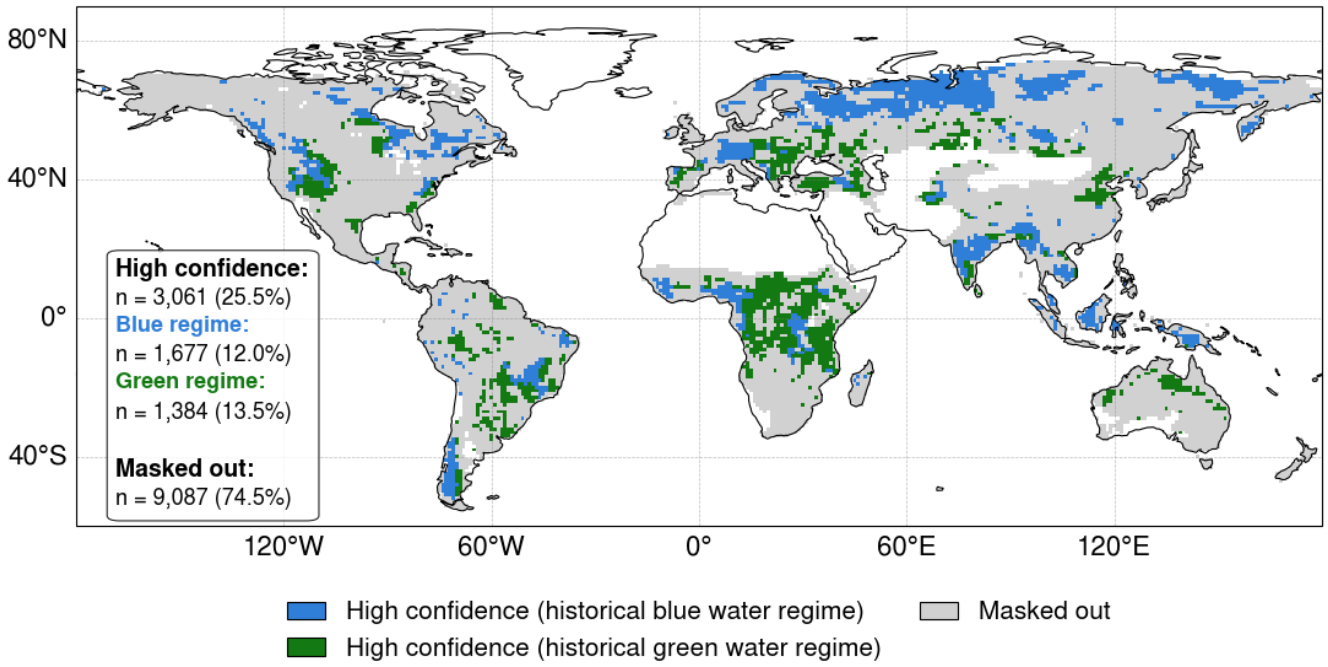

Figure S1. High-confidence subset used for the attribution sensitivity analysis under SSP3–7.0. Coloured grid cells show locations where at least 8 of 12 models agree on the sign of projected ΔBGWS. These high-confidence cells are separated by their historical BGWS regime: blue indicates the historical blue water regime (BGWS > 0), and green indicates the historical green water regime (BGWS < 0). Light grey areas show grid cells that are part of the valid analysis domain but do not meet the high-confidence criterion. Numbers indicate grid-cell counts and area-weighted fractions relative to the valid analysis domain.

S3 Benchmarking against ERA5-Land and observation-based data

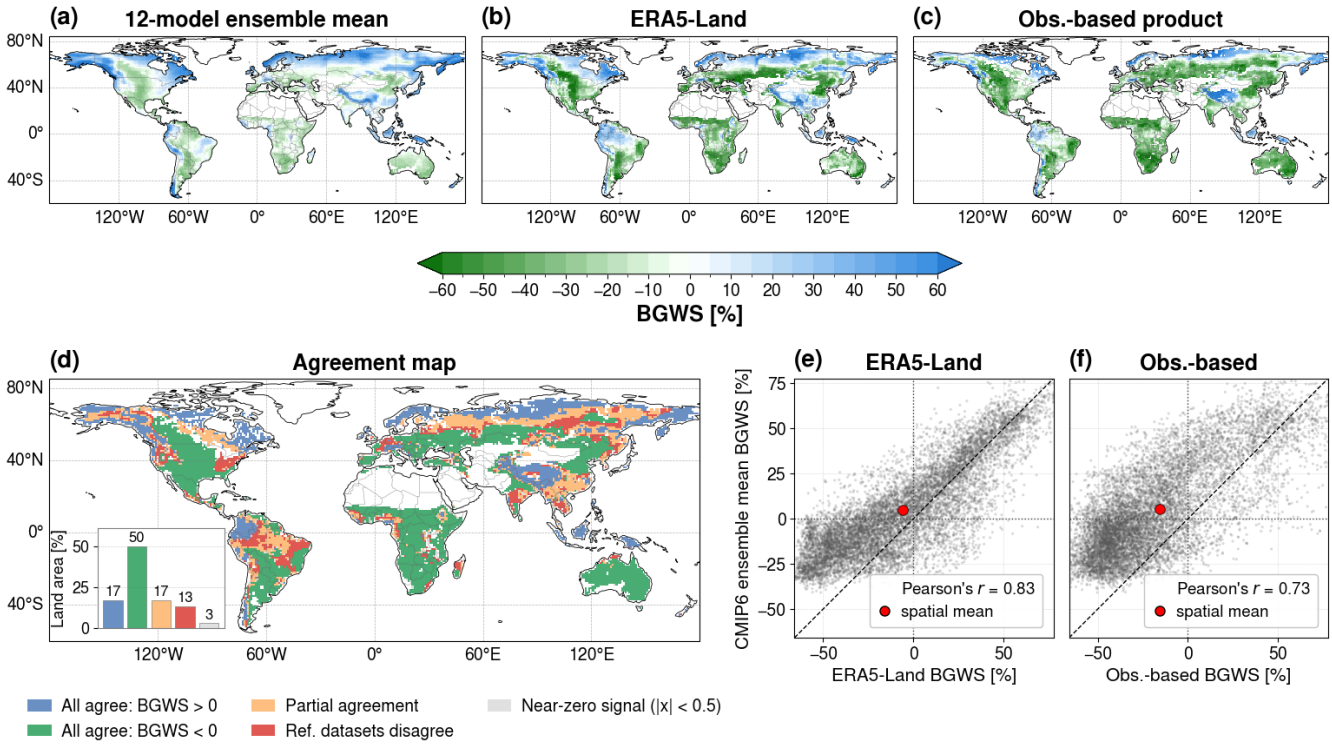

Figure S2. Comparison of historical mean BGWS (1985–2014) across the 12-model CMIP6 ensemble mean, ERA5-Land, and observation-based data. Panels (a)–(c) show global maps of BGWS for the ensemble mean, ERA5-Land, and the observation-based product, respectively. Panel (d) distinguishes grid cells where all three datasets agree on the BGWS sign, where the 12-model ensemble mean agrees with only one reference dataset (partial agreement), where both reference datasets disagree with the ensemble mean, and where at least one dataset has a near-zero signal ($|x| < 0.5\,\%$). The inset in panel (d) shows the fraction of land area in each agreement category. Panels (e) and (f) show grid-cell scatter plots comparing ERA5-Land and the observation-based product, respectively, against the CMIP6 ensemble mean. Only grid cells retained by the fixed historical analysis mask are shown.

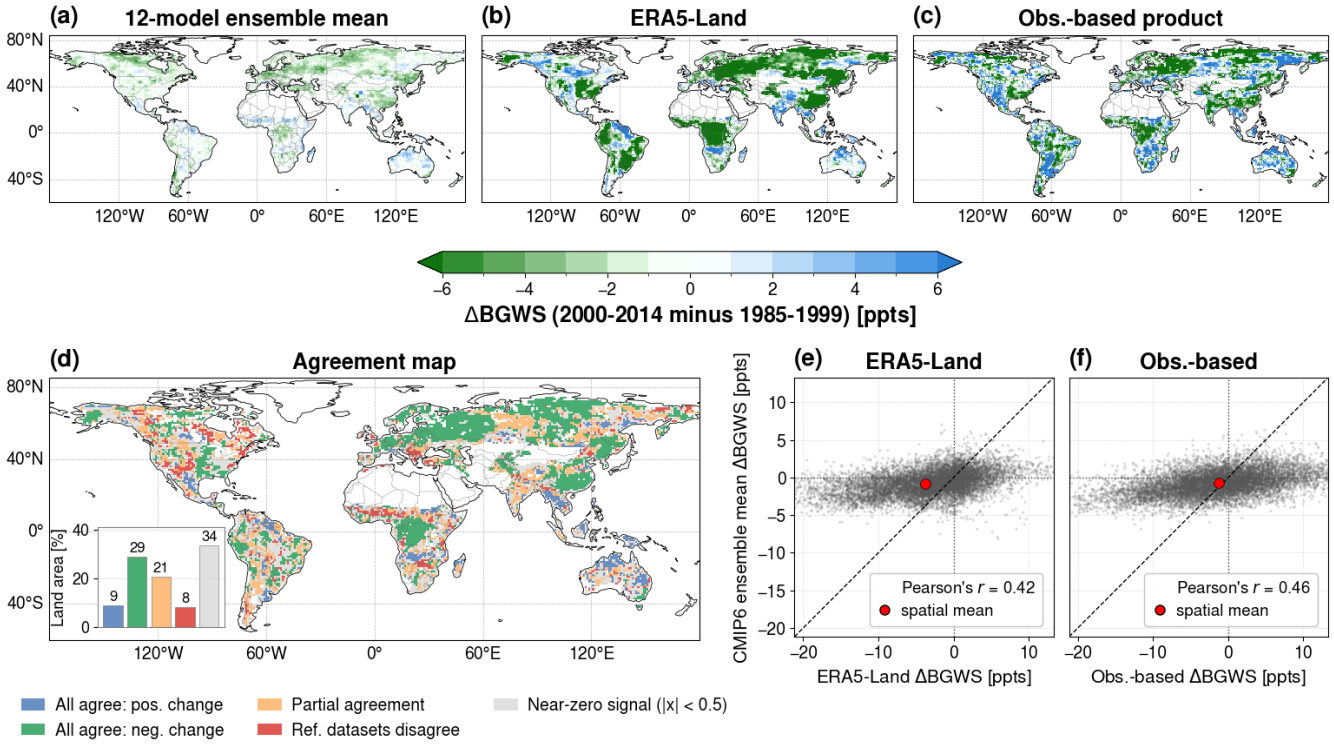

Figure S3. As in Fig. S2, but for historical changes in BGWS (2000–2014 minus 1985–1999). Categories defined as in Fig. S2

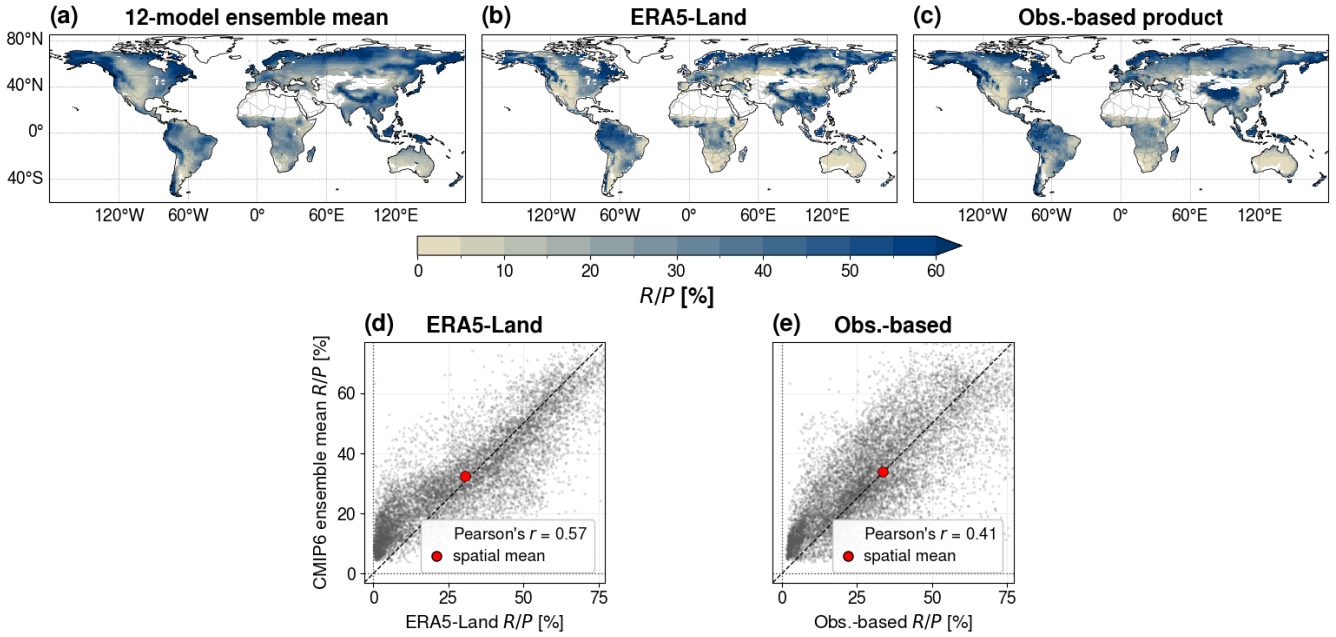

Figure S4. Comparison of the historical mean runoff-to-precipitation ratio (R/P; 1985–2014) across the 12-model CMIP6 ensemble mean, ERA5-Land, and observation-based data. Panels (a)–(c) show global maps of R/P for the ensemble mean, ERA5-Land, and the observation-based product, respectively. Panels (d) and (e) show grid-cell scatter plots comparing ERA5-Land and the observation-based product, respectively, against the CMIP6 ensemble mean. Point density is indicated by transparency, the dashed line marks the 1:1 relationship, and the red marker denotes the spatial mean. Only grid cells retained by the fixed historical analysis mask are shown.

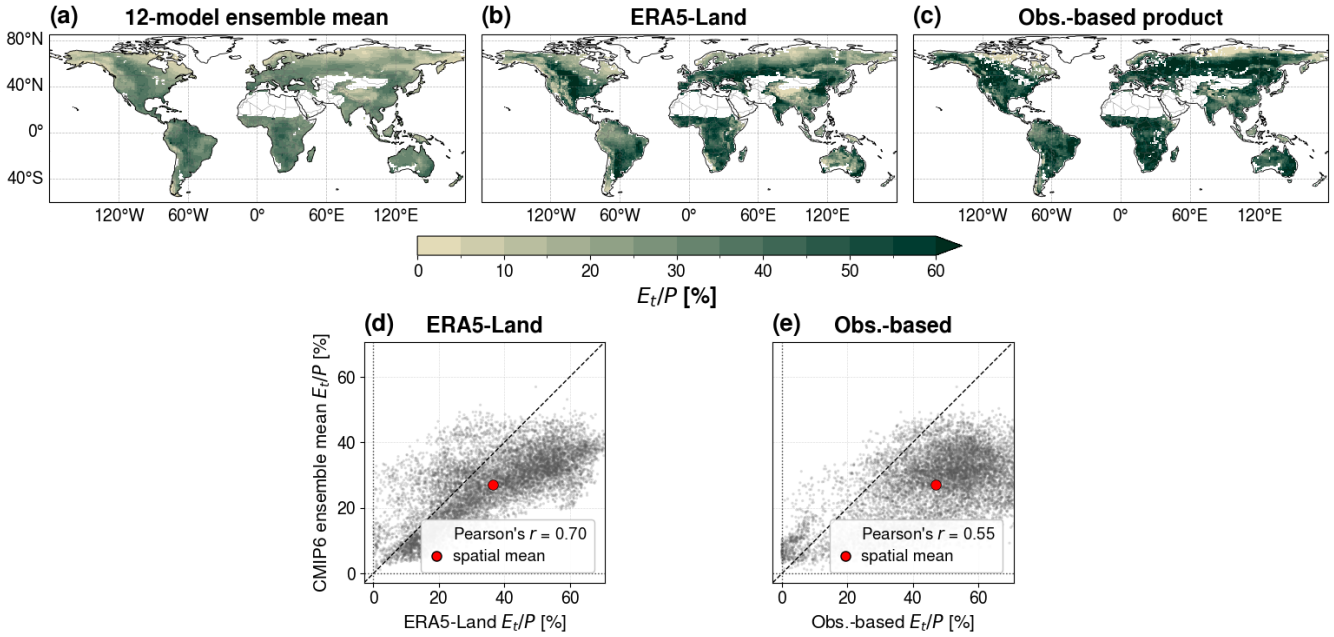

Figure S5. As in Fig. S4, but for the historical mean transpiration-to-precipitation ratio (E_t/P; 1985–2014).

S4 Historical climatic and hydroecological states

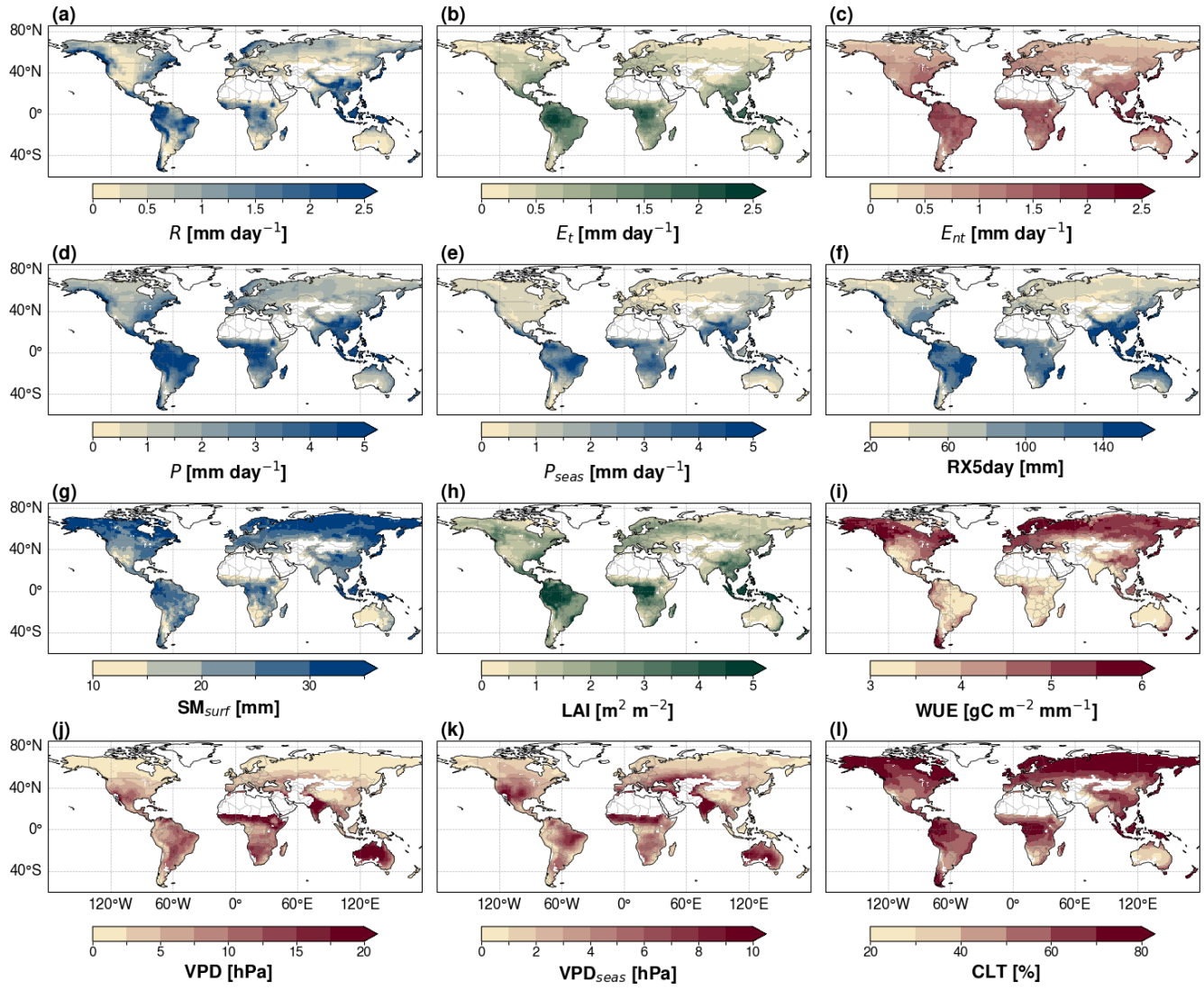

Figure S6. Historical mean states (1985–2014) of selected climatic and hydroecological variables in the 12-model CMIP6 ensemble mean. Panels show (a) R, (b) E_t, (c) E_{nt}, (d) P, (e) P_{seas}, (f) RX5day, (g) $\mathrm{SM}_{\mathrm{surf}}$, (h) LAI, (i) WUE, (j) VPD, (k) $\mathrm{VPD}_{\mathrm{seas}}$, and (l) CLT. Only grid cells retained by the fixed historical analysis mask are shown.

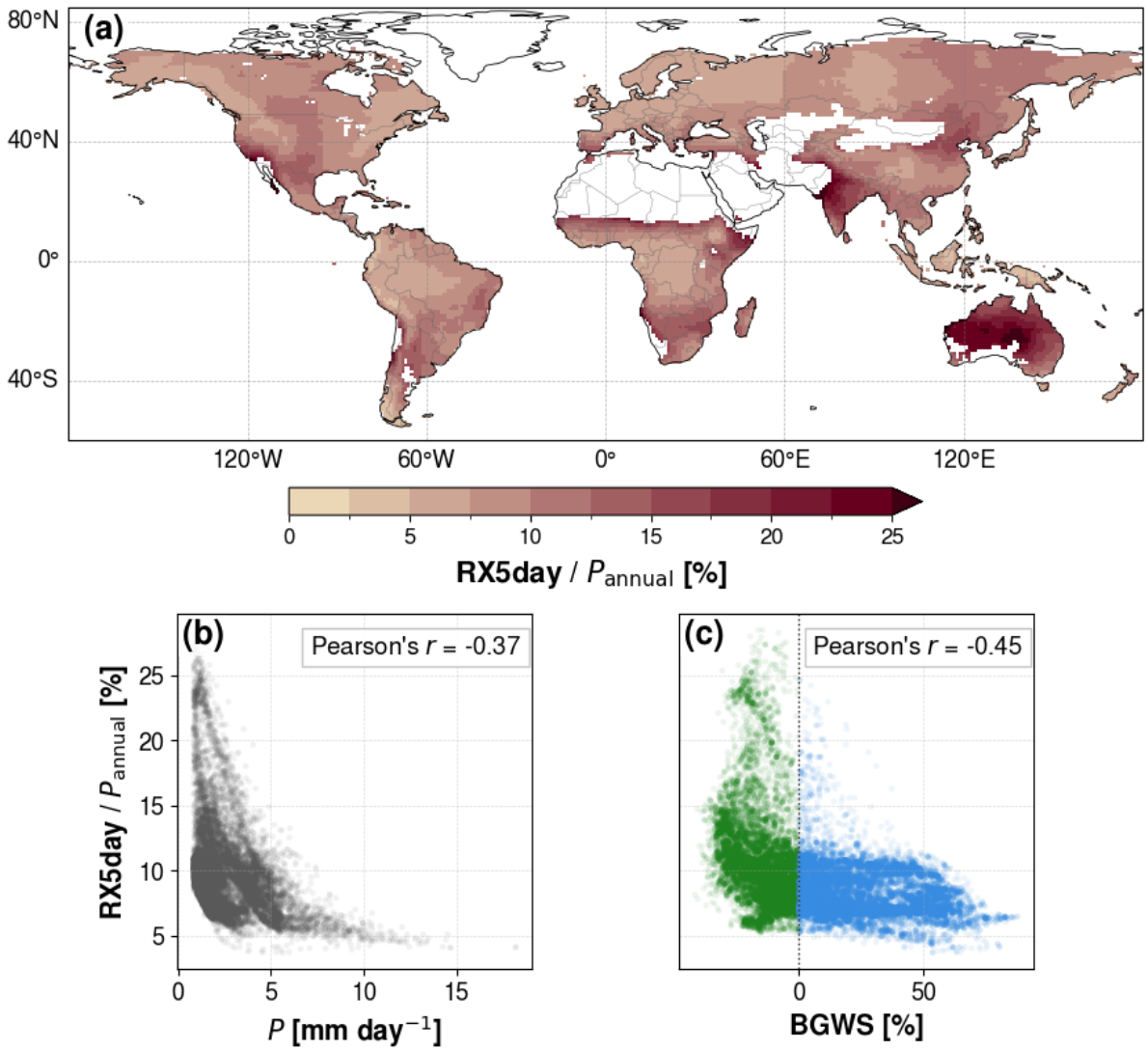

Figure S7. Historical mean RX5day ratio (RX5day/P_{annual}) in the 12-model CMIP6 ensemble mean and its relationships with P and BGWS. Panel (a) shows the ensemble mean RX5day/P_{annual} ratio [%]. Panels (b) and (c) show scatter plots of RX5day/P_{annual} against P and BGWS, respectively. In panel (c), blue points indicate grid cells with BGWS > 0 and green points indicate grid cells with BGWS < 0. Pearson's correlation coefficient (r) is shown in each scatter panel. Variable definitions and units are given in the Methods. Only grid cells retained by the fixed historical analysis mask are shown.

S5 Inter-model uncertainty in BGWS

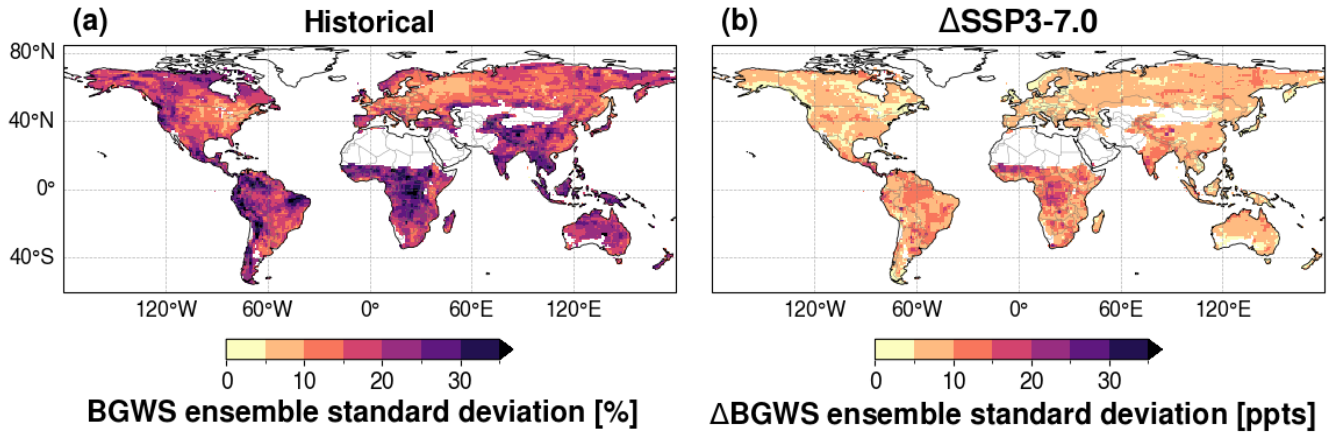

Figure S8. Ensemble standard deviation of BGWS across the 12-model CMIP6 ensemble. Panel (a) shows the historical ensemble standard deviation of BGWS for 1985–2014, and panel (b) shows the ensemble standard deviation of projected ΔBGWS under SSP3-7.0, calculated from changes between 2071–2100 and 1985–2014. Each grid cell represents the standard deviation across the model ensemble. Only grid cells retained by the fixed historical analysis mask are shown.

Table S4. Global mean BGWS [%], P [mm day^{-1}], R [mm day^{-1}], and E_t [mm day^{-1}] for the historical period, the SSP3-7.0 scenario, and the change (SSP3-7.0 minus historical) based on 12 CMIP6 ESMs. Only grid cells retained by the fixed historical analysis mask are shown.

	BGWS [%]			P [mm day^{-1}]			R [mm day^{-1}]			E_t [mm day^{-1}]		
Model	Historical	SSP3-7.0	Change	Historical	SSP3-7.0	Change	Historical	SSP3-7.0	Change	Historical	SSP3-7.0	Change
BCC-CSM2-MR	16.59	15.78	-0.82	2.98	3.07	0.10	1.30	1.37	0.07	0.59	0.57	-0.01
CanESM5	12.51	10.66	-1.86	3.01	3.27	0.26	1.10	1.24	0.13	0.50	0.55	0.05
CESM2	6.37	10.26	3.89	2.96	3.02	0.05	1.25	1.28	0.03	0.82	0.69	-0.13
CMCC-ESM2	12.83	13.13	0.31	3.08	3.25	0.16	1.81	1.97	0.16	0.95	1.02	0.07
CNRM-ESM2-1	-5.62	-1.62	3.99	2.77	2.88	0.10	1.17	1.27	0.10	0.95	0.87	-0.08
EC-Earth3-Veg	-13.46	-12.90	0.56	2.86	3.09	0.23	0.89	0.91	0.03	0.96	0.93	-0.02
GFDL-ESM4	-7.19	-7.19	0.00	3.00	2.96	-0.05	0.69	0.67	-0.02	0.82	0.71	-0.11
IPSL-CM6A-LR	6.26	8.35	2.09	3.36	3.58	0.22	1.41	1.65	0.24	0.85	0.86	0.01
MIROC-ES2L	1.63	0.38	-1.25	3.33	3.47	0.14	0.84	0.91	0.07	0.81	0.85	0.04
MPI-ESM1-2-LR	-5.66	-5.77	-0.11	2.70	2.83	0.13	0.62	0.70	0.07	0.76	0.78	0.02
NorESM2-MM	5.65	9.05	3.40	2.80	2.82	0.02	1.15	1.17	0.01	0.82	0.68	-0.13
UKESM1-0-LL	0.50	5.60	5.10	2.91	3.08	0.17	1.07	1.22	0.15	0.98	0.88	-0.10
Ensemble mean	2.95	4.29	1.35	2.98	3.11	0.13	1.13	1.23	0.09	0.82	0.78	-0.04

S6 Projected BGWS regime shifts

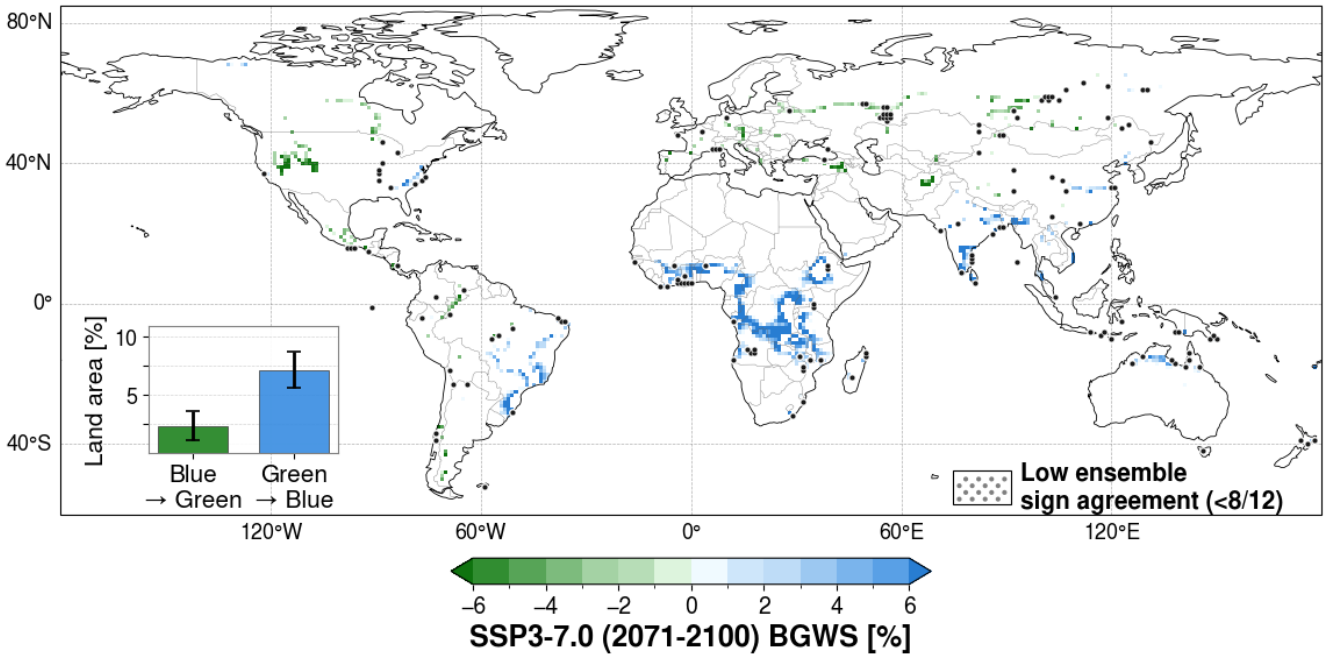

Figure S9. Ensemble mean BGWS under SSP3-7.0 for 2071–2100 across the 12-model CMIP6 ensemble. Only grid cells that undergo a regime shift between the historical period (1985–2014) and the far-future period are coloured, distinguishing shifts from blue to green water regimes and from green to blue water regimes. Stippling denotes regions of low ensemble sign agreement where fewer than 8 of 12 models agree on the sign of ΔBGWS. The inset shows the fraction of global land area undergoing each type of regime shift; error bars indicate the 95% confidence interval across the model ensemble. Only grid cells retained by the fixed historical analysis mask are shown.

S7 Projected changes in climatic and hydroecological variables

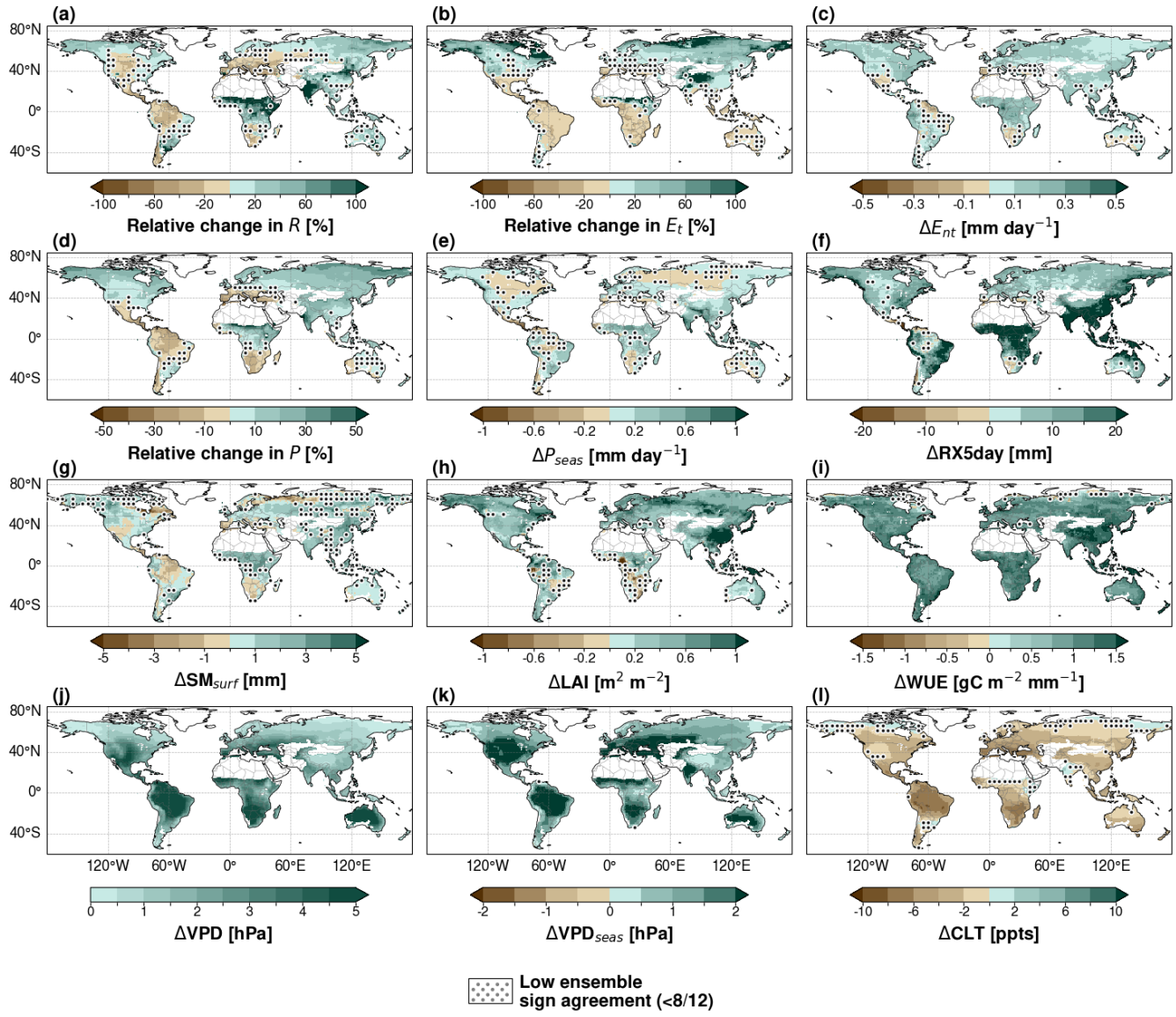

Figure S10. Ensemble mean changes in selected climatic and hydroecological variables under SSP3-7.0, shown as differences between 2071–2100 and 1985–2014. Panels show (a) relative change in R, (b) relative change in E_t, (c) ΔE_{nt}, (d) relative change in P, (e) ΔP_{seas}, (f) ΔRX5day, (g) $\Delta \mathrm{SM}_{\mathrm{surf}}$, (h) ΔLAI, (i) ΔWUE, (j) ΔVPD, (k) $\Delta \mathrm{VPD}_{\mathrm{seas}}$, and (l) ΔCLT. Stippling denotes regions of low ensemble sign agreement where fewer than 8 of 12 CMIP6 models agree on the sign of change and only when variable change is not solely positive. Only grid cells retained by the fixed historical analysis mask are shown.

S8 Random Forest regression and SHAP feature importance

As a nonlinear robustness check, we additionally applied Random Forest (RF) regression to the same final predictor set used in the blocked linear attribution analysis. RF models were fitted separately for the blue and green water regimes using repeated spatially blocked train–test splits, consistent with the main workflow. Hyperparameters were selected once for each regime using grouped cross-validation and then held fixed across repeated blocked splits. Model performance was evaluated on held-out test blocks using R^2.

To quantify predictor importance in the nonlinear model, we computed both held-out permutation importance and SHAP (Shapley additive explanations; Lundberg and Lee 2017). SHAP values assign each predictor an additive contribution to the model output and are used here as a diagnostic measure of variable importance and direction of influence within the RF model. Predictor importance was summarised across repeated blocked splits using median mean absolute SHAP values and compared with the held-out permutation importance from the blocked linear model (Table S5).

Figure S11 shows SHAP summary plots for representative blocked test splits with test performance closest to the median across repeats. Positive SHAP values indicate contributions towards more positive modelled ΔBGWS, whereas negative SHAP values indicate contributions towards more negative modelled ΔBGWS. Dot colours represent normalised predictor values. The RF analysis is intended as a complementary nonlinear sensitivity analysis rather than a replacement for the main blocked linear attribution framework.

Table S5. Comparison of predictor ranks from the blocked linear model, RF permutation importance, and RF SHAP importance for the blue and green water regimes. Rank 1 denotes the most important predictor within each method and regime.

Predictor	Blue linear	Blue RF perm.	Blue RF SHAP	Green linear	Green RF perm.	Green RF SHAP
ΔRX5day	1	1	1	1	1	1
ΔLAI	2	3	4	2	2	2
$\Delta\mathrm{SM}_{\mathrm{surf}}$	3	5	2	7	5	6
ΔWUE	5	8	8	6	9	8
ΔCLT	6	2	3	9	6	9
ΔP_{seas}	4	7	5	8	8	4
ΔVPD	8	4	6	4	7	7
ΔP	7	6	7	5	4	5
$\Delta\mathrm{VPD}_{\mathrm{seas}}$	9	9	9	3	3	3

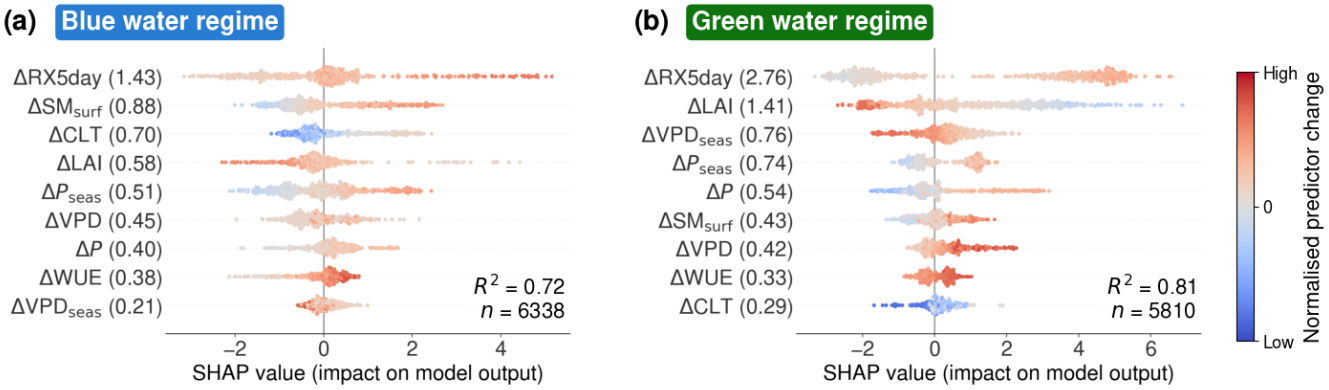

Figure S11. Random-Forest-based variable importance of the final predictor set for projected ΔBGWS in the two historical BGWS regimes based on the 12-model CMIP6 ensemble mean. (a) Blue water regime and (b) Green water regime. SHAP summary plots are shown for representative blocked test splits. Positive and negative SHAP values indicate contributions towards more positive or more negative modelled ΔBGWS, respectively. Dot colours represent normalised predictor values from low to high. Predictors are ordered by median mean absolute SHAP values across repeated blocked runs; values in brackets denote the corresponding median mean absolute SHAP values. Reported R^2 values give the mean held-out test performance across repeated blocked runs, and n is the number of grid cells analysed in each regime.

References

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., D'Andrea, F., Davini, P., de Lavergne, C., Denvil, S., Deshayes, J., Devilliers, M., Ducharne, A., Dufresne, J.-L., Dupont, E., Éthé, C., Fairhead, L., Falletti, L., Flavoni, S., Foujols, M.-A., Gardoll, S., Gastineau, G., Ghattas, J., Grandpeix, J.-Y., Guenet, B., Guez, Lionel, E., Guilyardi, E., Guimberteau, M., Hauglustaine, D., Hourdin, F., Idelkadi, A., Joussaume, S., Kageyama, M., Khodri, M., Krinner, G., Lebas, N., Levavasseur, G., Lévy, C., Li, L., Lott, F., Lurton, T., Luyssaert, S., Madec, G., Madeleine, J.-B., Maignan, F., Marchand, M., Marti, O., Mellul, L., Meurdesoif, Y., Mignot, J., Musat, I., Ottlé, C., Peylin, P., Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thiéblemont, R., Traore, A. K., Vancoppenolle, M., Vial, J., Vialard, J., Viovy, N., and Vuichard, N.: Presentation and Evaluation of the IPSL-CM6A-LR Climate Model, Journal of Advances in Modeling Earth Systems, 12, e2019MS002 010, https://doi.org/10.1029/2019MS002010, 2020.

Brunner, L., Poschlod, B., Dutra, E., Fischer, E. M., Martius, O., and Sillmann, J.: A global perspective on the spatial representation of climate extremes from km-scale models, Environmental Research Letters, 20, 074 054, https://doi.org/10.1088/1748-9326/ade1ef, 2025.

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L. K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G., Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T. M., Lindsay, K., Lipscomb, W. H., Mills, M. J., Neale, R., Oleson, K. W., Otto-Bliesner, B., Phillips, A. S., Sacks, W., Tilmes, S., van Kampenhout, L., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J. E., Kinnison, D., Kushner, P. J., Larson, V. E., Long, M. C., Mickelson, S., Moore, J. K., Nienhouse, E., Polvani, L., Rasch, P. J., and Strand, W. G.: The Community Earth System Model Version 2 (CESM2), Journal of Advances in Modeling Earth Systems, 12, e2019MS001 916, https://doi.org/10.1029/2019MS001916, 2020.

Denissen, J. M. C., Teuling, A. J., Pitman, A. J., Koirala, S., Migliavacca, M., Li, W., Reichstein, M., Winkler, A. J., Zhan, C., and Orth, R.: Widespread shift from ecosystem energy to water limitation with climate change, Nature Climate Change, 12, 677–684, https://doi.org/10.1038/s41558-022-01403-8, 2022.

Dunne, J. P., Horowitz, L. W., Adcroft, A. J., Ginoux, P., Held, I. M., John, J. G., Krasting, J. P., Malyshev, S., Naik, V., Paulot, F., Shevliakova, E., Stock, C. A., Zadeh, N., Balaji, V., Blanton, C., Dunne, K. A., Dupuis, C., Durachta, J., Dussin, R., Gauthier, P. P. G., Griffies, S. M., Guo, H., Hallberg, R. W., Harrison, M., He, J., Hurlin, W., McHugh, C., Menzel, R., Milly, P. C. D., Nikonov, S., Paynter, D. J., Ploshay, J., Radhakrishnan, A., Rand, K., Reichl, B. G., Robinson, T., Schwarzkopf, D. M., Sentman, L. T., Underwood, S., Vahlenkamp, H., Winton, M., Wittenberg, A. T., Wyman, B., Zeng, Y., and Zhao, M.: The GFDL Earth System Model Version 4.1 (GFDL-ESM 4.1): Overall Coupled Model Description and Simulation Characteristics, Journal of Advances in Modeling Earth Systems, 12, e2019MS002 015, https://doi.org/10.1029/2019MS002015, 2020.

Döscher, R., Acosta, M., Alessandri, A., Anthoni, P., Arsouze, T., Bergman, T., Bernardello, R., Boussetta, S., Caron, L.-P., Carver, G., Castrillo, M., Catalano, F., Cvijanovic, I., Davini, P., Dekker, E., Doblas-Reyes, F. J., Docquier, D., Echevarria, P., Fladrich, U., Fuentes-Franco, R., Gröger, M., v. Hardenberg, J., Hieronymus, J., Karami, M. P., Keskinen, J.-P., Koenigk, T., Makkonen, R., Massonnet, F., Ménégoz, M., Miller, P. A., Moreno-Chamarro, E., Nieradzik, L., van Noije, T., Nolan, P., O'Donnell, D., Ollinaho, P., van den Oord, G., Ortega, P., Prims, O. T., Ramos, A., Reerink, T., Rousset, C., Ruprich-Robert, Y., Le Sager, P., Schmith, T., Schrödner, R., Serva, F., Sicardi, V., Sloth Madsen, M., Smith, B., Tian, T., Tourigny, E., Uotila, P., Vancoppenolle, M., Wang, S., Wårlind, D., Willén, U., Wyser, K., Yang, S., Yepes-Arbós, X., and Zhang, Q.: The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6, Geoscientific Model Development, 15, 2973–3020, https://doi.org/10.5194/gmd-15-2973-2022, 2022.

Hajima, T., Watanabe, M., Yamamoto, A., Tatebe, H., Noguchi, M. A., Abe, M., Ohgaito, R., Ito, A., Yamazaki, D., Okajima, H., Ito, A., Takata, K., Ogochi, K., Watanabe, S., and Kawamiya, M.: Development of the MIROC-ES2L Earth system model and the evaluation of biogeochemical processes and feedbacks, Geoscientific Model Development, 13, 2197–2244, https://doi.org/10.5194/gmd-13-2197-2020, 2020.

Hou, Y., Guo, H., Yang, Y., and Liu, W.: Global Evaluation of Runoff Simulation From Climate, Hydrological and Land Surface Models, Water Resources Research, 59, e2021WR031 817, https://doi.org/10.1029/2021WR031817, 2023.

Laimighofer, J., Melcher, M., and Laaha, G.: Parsimonious statistical learning models for low-flow estimation, Hydrology and Earth System Sciences, 26, 129–148, https://doi.org/10.5194/hess-26-129-2022, 2022.

Lovato, T., Peano, D., Butenschön, M., Materia, S., Iovino, D., Scoccimarro, E., Fogli, P. G., Cherchi, A., Bellucci, A., Gualdi, S., Masina, S., and Navarra, A.: CMIP6 Simulations With the CMCC Earth System Model (CMCC-ESM2), Journal of Advances in Modeling Earth Systems, 14, e2021MS002 814, https://doi.org/10.1029/2021MS002814, 2022.

Lundberg, S. M. and Lee, S.-I.: A Unified Approach to Interpreting Model Predictions, Advances in Neural Information Processing Systems, 30, https://doi.org/10.48550/arXiv.1705.07874, 2017.

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., Fast, I., Fiedler, S., Fläschner, D., Gayler, V., Giorgetta, M., Goll, D. S., Haak, H., Hagemann, S., Hedemann, C., Hohenegger, C., Ilyina, T., Jahns, T., Jimenéz-de-la Cuesta, D., Jungclaus, J., Kleinen, T., Kloster, S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornblueh, L.,

Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Möbis, B., Müller, W. A., Nabel, J. E. M. S., Nam, C. C. W., Notz, D., Nyawira, S.-S., Paulsen, H., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., Popp, M., Raddatz, T. J., Rast, S., Redler, R., Reick, C. H., Rohrschneider, T., Schemann, V., Schmidt, H., Schnur, R., Schulzweida, U., Six, K. D., Stein, L., Stemmler, I., Stevens, B., von Storch, J.-S., Tian, F., Voigt, A., Vrese, P., Wieners, K.-H., Wilkenskjeld, S., Winkler, A., and Roeckner, E.: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO2, Journal of Advances in Modeling Earth Systems, 11, 998–1038, https://doi.org/10.1029/2018MS001400, 2019.

Meyer, H., Reudenbach, C., Wöllauer, S., and Nauss, T.: Importance of spatial predictor variable selection in machine learning applications – Moving from data reproduction to spatial prediction, Ecological Modelling, 411, 108 815, https://doi.org/10.1016/j.ecolmodel.2019.108815, 2019.

Seland, O., Bentsen, M., Olivié, D., Toniazzo, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y.-C., Kirkevåg, A., Schwinger, J., Tjiputra, J., Aas, K. S., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Karset, I. H. H., Landgren, O., Liakka, J., Moseid, K. O., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iversen, T., and Schulz, M.: Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations, Geoscientific Model Development, 13, 6165–6200, https://doi.org/10.5194/gmd-13-6165-2020, 2020.

Sellar, A. A., Jones, C. G., Mulcahy, J. P., Tang, Y., Yool, A., Wiltshire, A., O'Connor, F. M., Stringer, M., Hill, R., Palmieri, J., Woodward, S., de Mora, L., Kuhlbrodt, T., Rumbold, S. T., Kelley, D. I., Ellis, R., Johnson, C. E., Walton, J., Abraham, N. L., Andrews, M. B., Andrews, T., Archibald, A. T., Berthou, S., Burke, E., Blockley, E., Carslaw, K., Dalvi, M., Edwards, J., Folberth, G. A., Gedney, N., Griffiths, P. T., Harper, A. B., Hendry, M. A., Hewitt, A. J., Johnson, B., Jones, A., Jones, C. D., Keeble, J., Liddicoat, S., Morgenstern, O., Parker, R. J., Predoi, V., Robertson, E., Siahaan, A., Smith, R. S., Swaminathan, R., Woodhouse, M. T., Zeng, G., and Zerroukat, M.: UKESM1: Description and Evaluation of the U.K. Earth System Model, Journal of Advances in Modeling Earth Systems, 11, 4513–4558, https://doi.org/10.1029/2019MS001739, 2019.

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., and Teuling, A. J.: Investigating soil moisture–climate interactions in a changing climate: A review, Earth-Science Reviews, 99, 125–161, https://doi.org/10.1016/j.earscirev.2010.02.004, 2010.

Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Hanna, S., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Sigmond, M., Solheim, L., von Salzen, K., Yang, D., and Winter, B.: The Canadian Earth System Model version 5 (CanESM5.0.3), Geoscientific Model Development, 12, 4823–4873, https://doi.org/10.5194/gmd-12-4823-2019, 2019.

Séférian, R., Nabat, P., Michou, M., Saint-Martin, D., Voldoire, A., Colin, J., Decharme, B., Delire, C., Berthet, S., Chevallier, M., Sénési, S., Franchisteguy, L., Vial, J., Mallet, M., Joetzjer, E., Geoffroy, O., Guérémy, J.-F., Moine, M.-P., Msadek, R., Ribes, A., Rocher, M., Roehrig, R., Salas-y Mélia, D., Sanchez, E., Terray, L., Valcke, S., Waldman, R., Aumont, O., Bopp, L., Deshayes, J., Éthé, C., and Madec, G.: Evaluation of CNRM Earth System Model, CNRM-ESM2-1: Role of Earth System Processes in Present-Day and Future Climate, Journal of Advances in Modeling Earth Systems, 11, 4182–4227, https://doi.org/10.1029/2019MS001791, 2019.

Wu, T., Lu, Y., Fang, Y., Xin, X., Li, L., Li, W., Jie, W., Zhang, J., Liu, Y., Zhang, L., Zhang, F., Zhang, Y., Wu, F., Li, J., Chu, M., Wang, Z., Shi, X., Liu, X., Wei, M., Huang, A., Zhang, Y., and Liu, X.: The Beijing Climate Center Climate System Model (BCC-CSM): the main progress from CMIP5 to CMIP6, Geoscientific Model Development, 12, 1573–1600, https://doi.org/10.5194/gmd-12-1573-2019, 2019.