Article

Supplementary Material: Modelling the start of season date of Hungarian grasslands using remote sensing data and 10 process-based models

Réka Ágnes Dávid [1,2], Zoltán Barcza [1,3,*], Roland Hollós [3,4] and Anikó Kern [5,6]

[1] Department of Meteorology, Institute of Geography and Earth Sciences, ELTE Eötvös Loránd University,
Pázmány P. st. 1/A, H-1117 Budapest, Hungary; rekaagnes@nimbus.elte.hu
[2] Doctoral School of Earth Sciences, ELTE Eötvös Loránd University, Pázmány P. st. 1/A, H-1117 Budapest, Hungary
[3] Global Change Research Institute, Czech Academy of Sciences, Bělidla 986/4a, 603 00 Brno, Czech Republic; hollos.roland@atk.hu
[4] Agricultural Institute, HUN-REN Centre for Agricultural Research, Brunszvik u. 2., H-2462 Martonvásár, Hungary
[5] Space Research Group, Department of Geophysics and Space Science, Institute of Geography and
Earth Sciences, ELTE Eötvös Loránd University, Pázmány P. st. 1/A, H-1117 Budapest, Hungary; aniko.kern@ttk.elte.hu
[6] Institute for Electrophysics/SpaceLab, Óbuda University, Szőlő út 4., H-1034 Budapest, Hungary
* Correspondence: zoltan.barcza@ttk.elte.hu

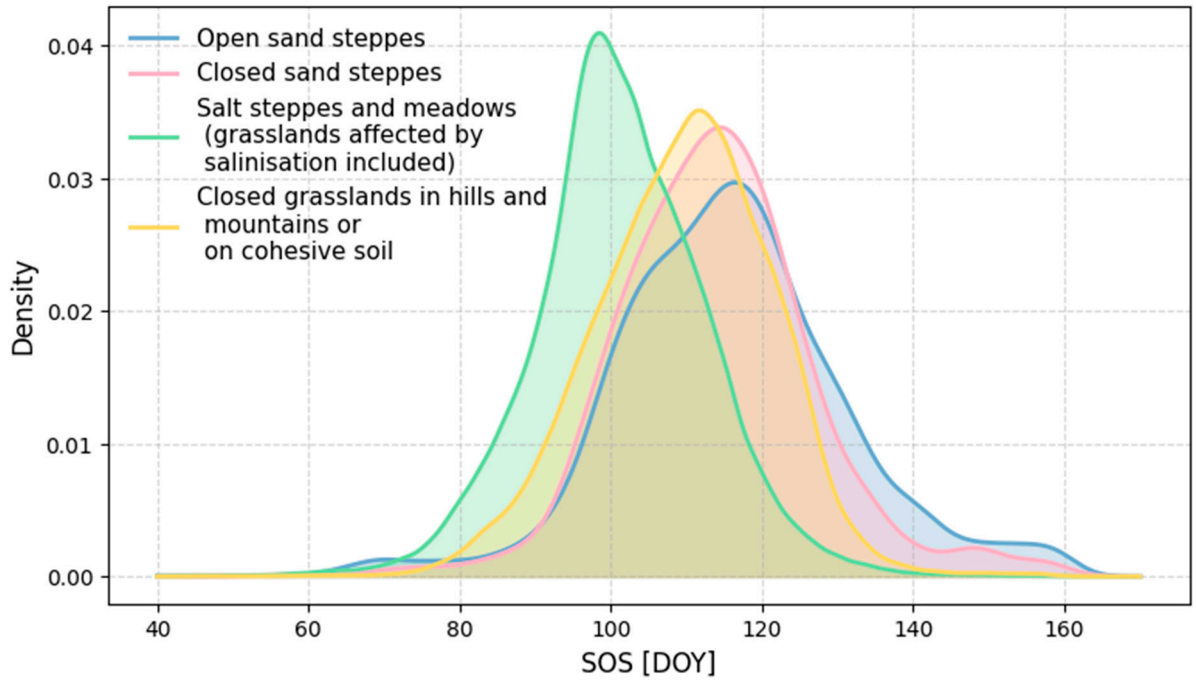


Figure S1. Kernel density estimates (KDE) of SOS date values (expressed as DOY) for the different grassland types. Salt steppes and meadows (grasslands affected by salinisation included) typically exhibit earlier green-up compared to the other types.

Table S1. Statistical indicators of the remote sensing-based SOS values for grasslands, where P1, P5, P95, and P99 denote the 1st, 5th, 95th, and 99th percentiles, respectively.

	Min	P1	P5	Mean	Median	P95	P99	Max	Std Dev
Open sand steppes	41	70	93	116	115	142	157	160	16
Closed sand steppes	54	79	95	114	114	135	152	160	13
Salt steppes and meadows (grasslands affected by salinisation included)	40	73	83	101	101	120	130	160	11
Closed grasslands in hills and mountains or on cohesive soil	44	81	90	110	110	127	135	160	12
All	40	74	84	103	103	123	133	160	12

1. GDD

The simplest and most commonly used approach is the so-called Warming Model, also known in the literature as the Growing Degree Days Model (abbreviated as GDD here), which is based on heat accumulation only. Heat accumulation is common in many models and is widely used in ecology and crop sciences. If the daily mean temperature exceeds the so-called base temperature (representing the plant's temperature threshold for functioning), then this daily mean temperature is included in the calculation of heat accumulation (by calculating the difference between the daily mean temperature and the base temperature). This procedure is done for many consecutive days. When the accumulated heat reaches a critical threshold value, the model indicates the occurrence of the SOS:

$$\sum_{t_0}^{t_{SOS}} \max\left(T_{avg} - T_b, 0\right) \geq GDDth, \quad \text{(S1)}$$

where $t_0 = 1\ DOY$ (i.e., the first day of the year in our case), t_{SOS} is the day of SOS, T_{avg} represents the daily mean temperature, T_b is the base temperature, and $GDDth$ is the critical threshold value that is to be reached to trigger SOS [1–3].

2. MGDD

In this model the heat requirement for the onset of growth is a function of the local climate. In warmer regions, greater heat accumulation is needed to reach the same phenological stage compared to cooler areas [4]. Consequently, the model logic suggests that it is advisable to adjust the predetermined threshold value for each region. The modified version of the GDD model (Modified Growing Degree Days Model; referred here as MGDD) defines the threshold value as a function of the long-term annual average heat accumulation:

$$\sum_{t_0}^{t_{SOS}} \max\left(T_{avg} - T_b, 0\right) \geq a\ \overline{GDD} + b, \quad \text{(S2)}$$

where $t_0 = 1\ DOY$, $\overline{GDD}$ is the long-term annual average heat accumulation, while a and b are adjustable parameters [2,3]. All other notations are consistent with definitions above.

3. MGDDwPP

The MGDDwPP (Modified Growing Degree Days with Photoperiod) model also relies on heat accumulation, but the calculation begins only after the photoperiod requirement has been fulfilled:

$$\sum_{t_0}^{t_{SOS}} \max\left(T_{avg} - T_b, 0\right) \geq a\ \overline{GDD} + b, \quad \text{(S3)}$$

where t_0 is the day of the year when the day length reaches p_0, that is a predefined photoperiod expressed by length of the daylight period (from sunrise to sunset). All other notations are consistent with those used above [3].

4. SEQ

The Sequential (hereinafter referred as SEQ) model assumes that heat accumulation (Eq. 1) only begins after the winter chilling requirement (which means cold weather with sufficient length) has been met. This method is common in

agro-ecological models for winter crops that need vernalization. Here heat accumulation should be calculated according to Eq. 1, but in this case the accumulation does not start at the first day of the year, but at DOY t_0, which is the day of the year when the chilling requirement (or the state of chilling (S_c); accumulated from the begginning of the calendar year) has been satisfied [2,3] :

$$\sum_{t_0}^{t_{SOS}} \max\left(T_{avg} < T_c, 0\right) \geq C, \tag{S4}$$

where T_c is the critical temperature, and C [°C] represents the chilling requirement of the vegetation.

5. PAR

The Parallel (PAR) model posits that heat and chilling requirements occur simultaneously, and the SOS occurs when the weighted heat accumulation reaches the threshold value [2,3]. In this model, heat accumulation is calculated similarly to the GDD model, but modulated by a weighting factor W:

$$\sum^{t_{SOS}} W \max\left(T_{avg} - T_b, 0\right) \geq F. \tag{S5}$$

where F is the forcing requirement, W reflects the degree to which the chilling requirement has been fulfilled, and is defined as:

$$W = \min\left(\frac{S_c}{C}, 1\right). \tag{S6}$$

All other notations are consistent with definitions used above. In this formulation, heat accumulation starts from the beginning of the year, but it is downregulated by W until the chilling requirement is fully met. As long as $S_c < C$, W remains below 1, reducing the effective contribution of daily heat. Once the chilling requirement is satisfied ($S_c = C$), W reaches 1, and daily heat accumulation proceeds without downregulation. Thus, the model captures the parallel nature of chilling and heat accumulation processes, with SOS occurring when the weighted heat sum reaches the critical threshold F.

6. GSI

We also used the Growing Season Index (GSI) model proposed by Jolly et al. (2005) [5]. It is based on a bioclimatic index, which requires minimum temperature, vapor pressure deficit (VPD), and photoperiod as input data. GSI can be interpreted as the quantification of conditions when plant growth is possible due to the optimal combination of the key environmental parameters. The GSI is calculated as the product of three sub-indices, each ranging between 0 and 1:

$$GSI = i_T \, i_{VPD} \, i_{PHTP}, \tag{S7}$$

where

$$i_T = \begin{cases} 0, \text{if } T_{min} \leq limit^1_{T_{min}} \\ \dfrac{T_{min} - limit^1_{T_{min}}}{limit^2_{T_{min}} - limit^1_{T_{min}}}, \\ 1, \text{if } T_{min} \geq limit^2_{T_{min}} \end{cases} \tag{S8}$$

$$i_{VPD} = \begin{cases} 0, \text{if } VPD \geq limit^1_{VPD} \\ 1 - \dfrac{VPD - limit^1_{VPD}}{limit^2_{VPD} - limit^1_{VPD}}, \\ 1, \text{if } VPD \leq limit^2_{VPD} \end{cases} \tag{S9}$$

$$i_{PHTP} = \begin{cases} 0, \text{if } PHTP \leq limit^1_{PHTP} \\ \dfrac{PHTP - limit^1_{PHTP}}{limit^2_{PHTP} - limit^1_{PHTP}} \\ 1, \text{if } PHTP \geq limit^2_{PHTP} \end{cases}, \quad \text{(S10)}$$

where $limit^1_{T_{min}}$, $limit^2_{T_{min}}$, $limit^1_{VPD}$, $limit^2_{VPD}$, $limit^1_{PHTP}$, and $limit^2_{PHTP}$ are adjustable parameters. SOS is identified as the first day on which the GSI exceeds a predefined threshold (SGS) that is also an adjustable parameter. According to Jolly et al. (2005) [5] the temporal change of GSI within a year is similar to that of NDVI which makes it applicable for SOS estimations. See Jolly et al. (2005) [5] for more details.

7. HSGSI

The Heat Sum Growing Season Index model (HSGSI; [6] is a combination of the GSI model described above [5] and the GDD model [1–3]. It calculates the GSI with an additional heat sum index, thereby considering not only the daily minimum temperature as a limiting factor, but also the cumulative heat sum over the past 10 days, thus integrating the properties of both models. In this sense, the cumulative GDD represents the memory of the system, while the other variables represent instantaneous values, theoretically suppressing or promoting plant growth. The index is defined as follows [6,7]:

$$HSGSI(d) = i_{HTSM_{10}} \cdot \frac{1}{10} \sum_{d-9}^{d} (i_T(d) \cdot i_{VPD}(d) \cdot i_{PHTP}(d)), \quad \text{(S11)}$$

where $i_T(d)$, $i_{VPD}(d)$, and $i_{PHTP}(d)$ are defined above (Eqs. 8–10) and represent daily quantities of the sub-indices, while $i_{HTSM_{10}}$ pertains to a 10-day period:

$$i_{HTSM_{10}} = \begin{cases} 0, \text{ha } HTSM_{10} \leq limit^1_{HTSM} \\ \dfrac{HTSM_{10} - limit^1_{HTSM}}{limit^2_{HTSM} - limit^1_{HTSM}} \\ 1, \text{ha } HTSM_{10} \geq limit^2_{HTSM} \end{cases}, \quad \text{(S12)}$$

$$HTSM_{10}(d) = \sum_{d-9}^{d} (T_{avg}(d) - (T_b)), \quad \text{(S13)}$$

$$T_{avg}(d) \leq T_{base}: HTSM_{10}(d) = 0$$

The parameters $limit^1_{HTSM}$ and $limit^2_{HTSM}$ are additional adjustable parameters.

8. AGSI

The Accumulated Heat Sum Growing Season Index model (AGSI; [3] also combines the GSI [5] and GDD [1–3] models but adopts a different approach. The model calculates the GSI-index using Eq. 7; however, the SOS is determined when the GSI accumulation reaches a critical threshold:

$$\sum^{t_{SOS}} GSI = G^*, \quad \text{(S14)}$$

where t_{SOS} represents the start of the growing season.

9. AGSIwSW

The Accumulated Heat Sum Growing Season Index Model with Soil Water Content Model (AGSIMwSW; [3] is an extended version of the AGSI model that also relies on the accumulated HSGSI-index (Eq. 14). However, instead of using

VPD as an indicator of atmospheric drought, this model incorporates soil water content as a key determinant of water availability:

$$i_{SW} = \begin{cases} 0, \text{if } SWC \le limit_{SW}^{1} \\ \dfrac{SWC - limit_{SW}^{1}}{limit_{SW}^{2} - limit_{SW}^{1}} \\ 1, \text{if } SWC \ge limit_{SW}^{2} \end{cases}, \quad \text{(S15)}$$

where $limit_{SW}^{1}$, and $limit_{SW}^{2}$ are adjustable parameters. The GSI-index is then computed as follows:

$$GSI = i_T \, i_{SW} \, i_{PHTP} \quad \text{(S16)}$$

SOS occurs when the index reaches a predefined value, as in case of Eq. 14.

10. AHSGSI

Accumulated Heat Sum Growing Season Index model calculates the HSGSI index [6] and defines the SOS as the accumulated HSGSI reaches a predefined threshold as in case of AGSI [3].

Model ensemble

It is a general approach in numerical weather prediction and other modelling types (like crop models or biogeochemical models) to construct model ensembles (ENS) and to calculate the multimodel mean (or median), which in many cases outperforms the individual models and provides a robust result [8,9]. Following this logic we created an additional model by averaging all model results at the pixel and individual year level to construct the ENS model.

Likelihood formulation

Although the phenology models are deterministic, they produce a single predicted SOS value for each observation. Therefore, the likelihood is defined at the level if the residuals between observed and modelled SOS values, rather than being associated with the internal model structure. The log-likelihood is given by

$$logL = -\frac{1}{2}\sum_{i=1}^{N}\left[+\frac{\varepsilon_i^2}{\sigma^2}\right], \quad \text{(S17)}$$

where N is the number of SOS observations, ε_i are the residuals, and σ^2 denotes the variance of the observed SOS values.

Table S2. Basic statistics for GEN calibrated models, where P1, P5, P95, and P99 denote the 1st, 5th, 95th, and 99th percentiles, respectively.

	Min	P1	P5	Mean	Median	P95	P99	Max	Std Dev
GEN GDD	81	84	93	102.9	103	113	116	133	6.4
GEN MGDD	85	93	98	105.9	105	113	117	130	5.0
GEN MGDDwPP	88	93	96	103.2	104	107	111	124	3.8
GEN SEQ	81	84	92	102.1	103	111	116	132	6.1
GEN PAR	81	84	93	102.1	103	111	116	132	6.0
GEN GSI	77	80	83	107.3	108	124	124	145	13.4
GEN HSGSI	79	79	83	107.8	108	123	124	146	12.2
GEN AGSI	82	88	91	102.8	98	118	122	156	10.3
GEN AGSIwSW	87	90	92	105.4	103	123	130	147	11.1
GEN AHSGSI	89	91	92	103.0	102	116	118	143	8.3

GEN ENS	85.3	88	92.3	104.3	106	115	118	137	7.3

Table S3. Basic statistics for GEN GRASS calibrated models, where P1, P5, P95, and P99 denote the 1st, 5th, 95th, and 99th percentiles, respectively.

	Min	P1	P5	Mean	Median	P95	P99	Max	Std Dev
GEN GRASS GDD	80	83	92	103	102	115	122	138	7
GEN GRASS MGDD	96	100	109	115	114	122	128	141	5
GEN GRASS MGDDwPP	85	91	96	105	103	120	126	139	7
GEN GRASS SEQ	79	84	92	103	103	117	123	139	7
GEN GRASS PAR	82	85	93	104	104	116	122	139	7
GEN GRASS GSI	77	78	79	107	108	124	128	159	15
GEN GRASS HSGSI	78	82	83	113	114	124	127	159	10
GEN GRASS AGSI	87	88	90	103	99	120	128	157	10
GEN GRASS AGSIwSW	86	88	90	105	103	124	130	150	11
GEN GRASS AHSGSI	95	97	99	109	107	122	124	149	8
GEN GRASS ENS	87	89	95	107	107.6	118.1	124.6	145.2	8

Table S4. Basic statistics for PIX calibrated models, where P1, P5, P95, and P99 denote the 1st, 5th, 95th, and 99th percentiles, respectively.

	Min	P1	P5	Mean	Median	P95	P99	Max	Std Dev
PIX GDD	70	83	89	105	104	121	125	137	9.6
PIX MGDD	71	83	90	105	104	120	125	137	9.2
PIX MGDDwPP	73	85	90	103	102	120	125	148	9.1
PIX SEQ	70	83	89	104	104	121	125	136	9.5
PIX PAR	71	83	89	104	104	121	125	136	9.4
PIX GSI	1	73	76	106	107	126	130	261	16.2
PIX HSGSI	70	73	76	106	106	126	131	262	16.1
PIX AGSI	79	85	88	105	105	124	129	146	11.3
PIX AGSIwSW	78	84	87	105	105	126	132	170	12.8
PIX AHSGSI	76	84	88	104	104	124	129	148	11.3
PIX ENS	73.9	82.8	87.2	105	104.8	121.9	126.4	141.6	10.4

Table S5. Pairwise comparisons of grassland types using Dunn's test with Bonferroni correction.

	Open sand steppes	**Closed sand steppes**	**Salt steppes and meadows (grasslands affected by salinisation included)**	**Closed grasslands in hills and mountains or on cohesive soil**
Open sand steppes	1	1	1	0.00021
Closed sand steppes	1	1	1	0.22857
Salt steppes and meadows (grasslands affected by salinisation included)	1	1	1	0
Closed grasslands in hills and mountains or on cohesive soil	0.00021	0.22857	0	1

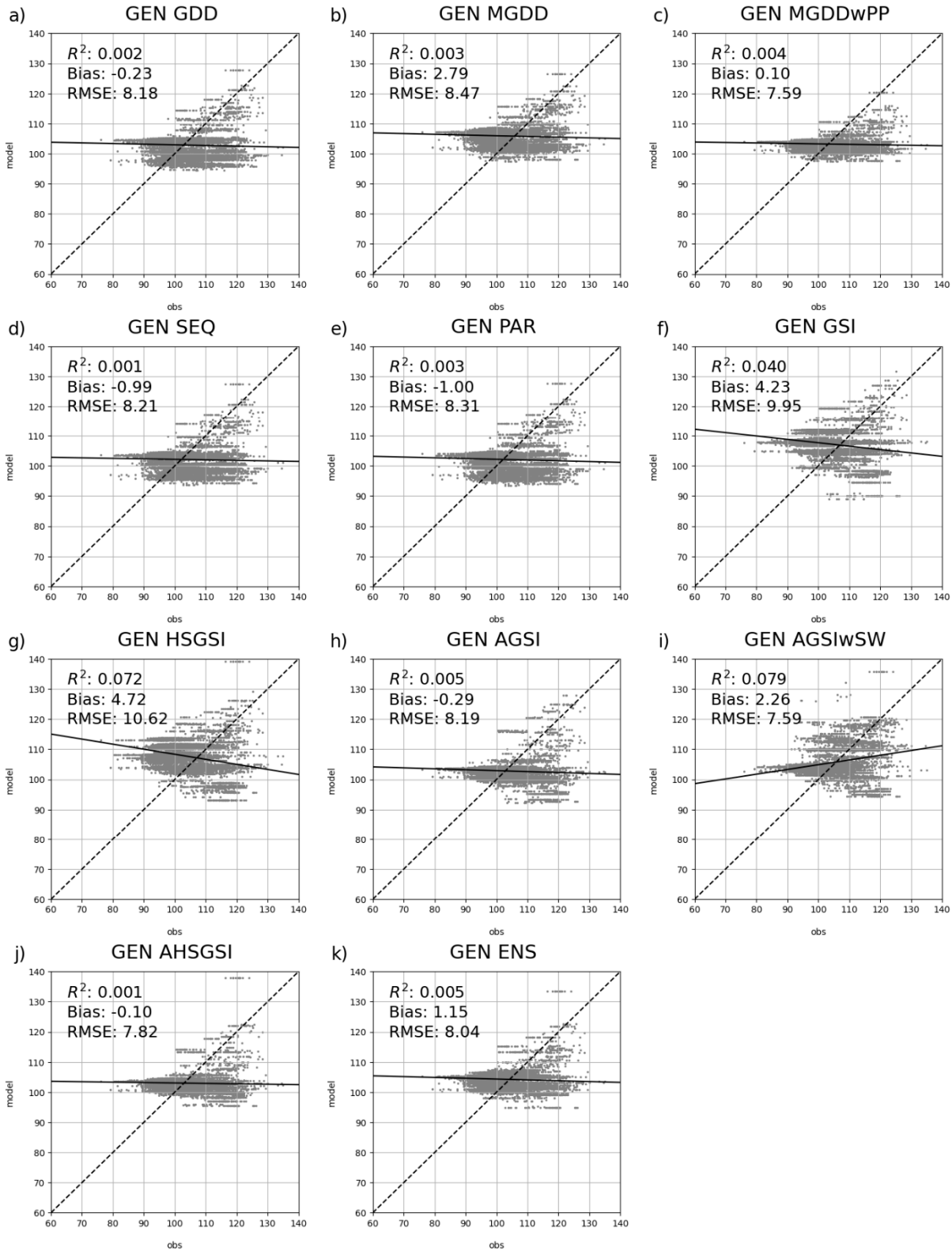


Figure S2. Comparison of the modelled and observed temporal mean SOS values from the GEN calibrations, where each point represents a 5-year average for a single pixel. Pearson r², bias and RMSE values are shown.

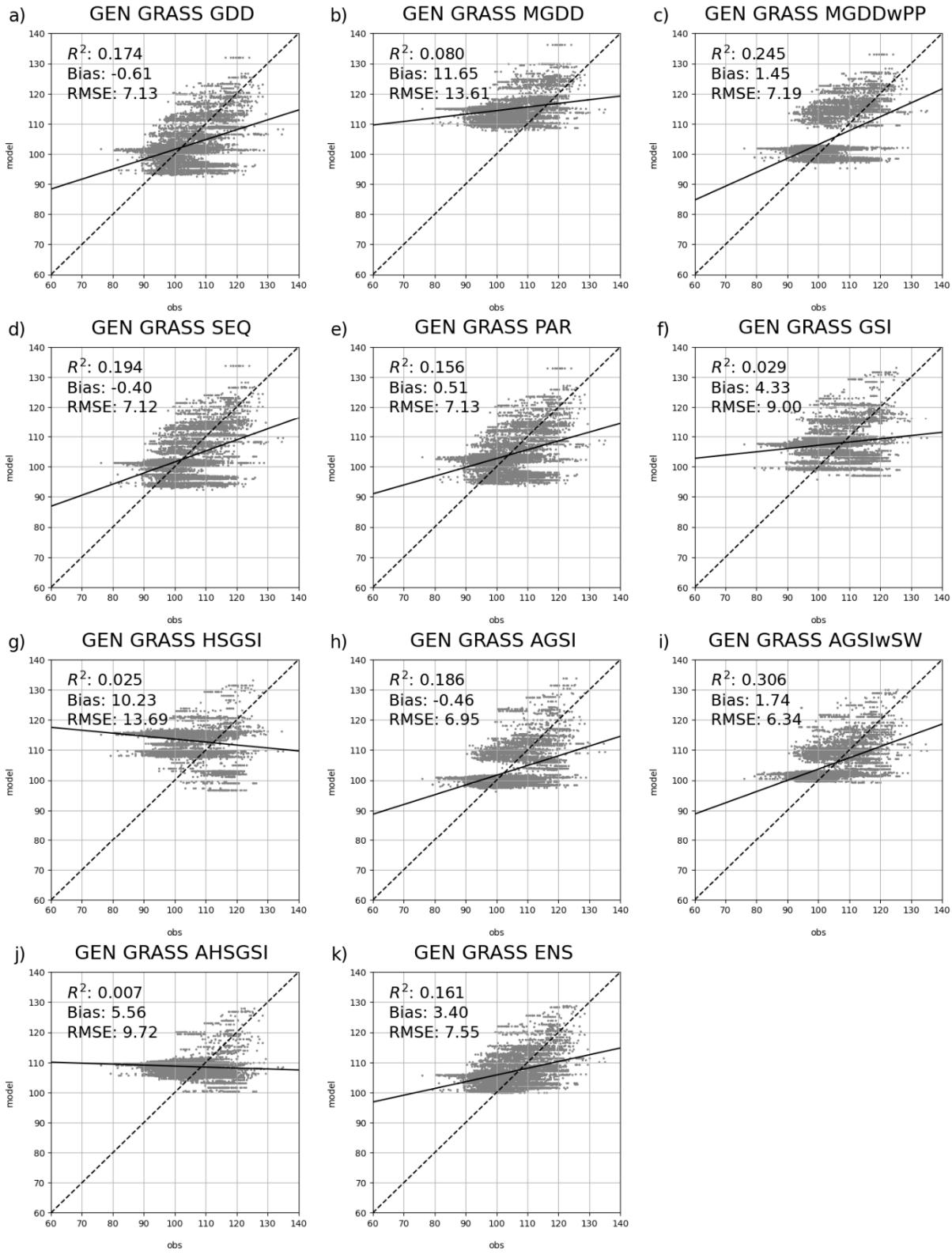


Figure S3. Comparison of the modelled and observed temporal mean SOS values from the GEN GRASS calibrations, where each point represents a 5-year average for a single pixel. Pearson r², bias and RMSE values are shown.

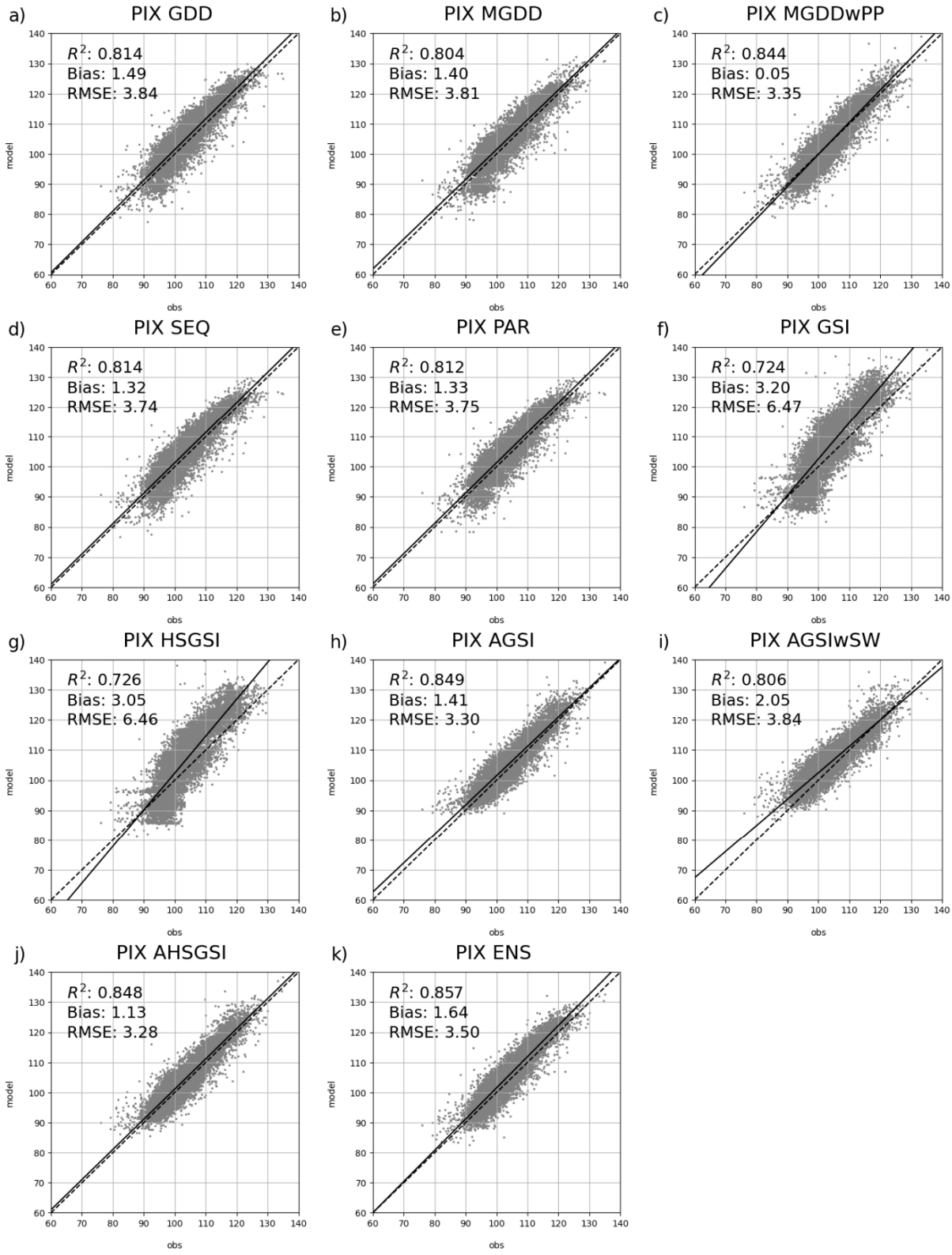


Figure S4. Comparison of the modelled and observed temporal mean SOS values from the PIX calibrations, where each point represents a 5-year average for a single pixel. Pearson r², bias and RMSE values are shown.

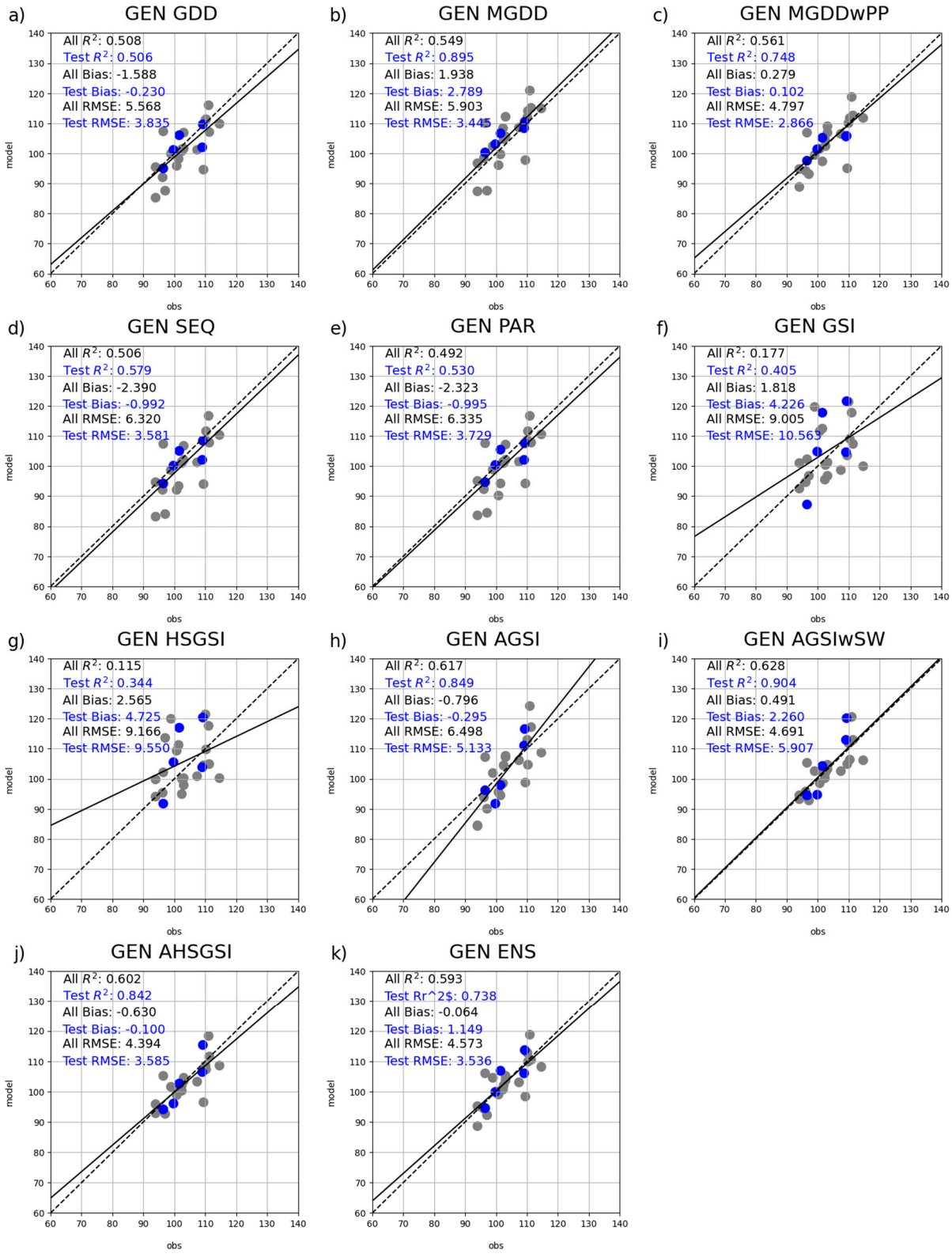


Figure S5. Spatial comparison of the observed and modelled SOS based on the GEN calibration. Each point represents the annual average of approximately 12,000 pixels. Gray points represent training data, and blue points indicating validation years. The dashed line denotes the 1:1 relationship, while the solid line is the result of the linear regression. Statistical metrics (Pearson r², bias, and RMSE) are reported for both the full dataset (including training and validation years), with test-year values highlighted in blue.

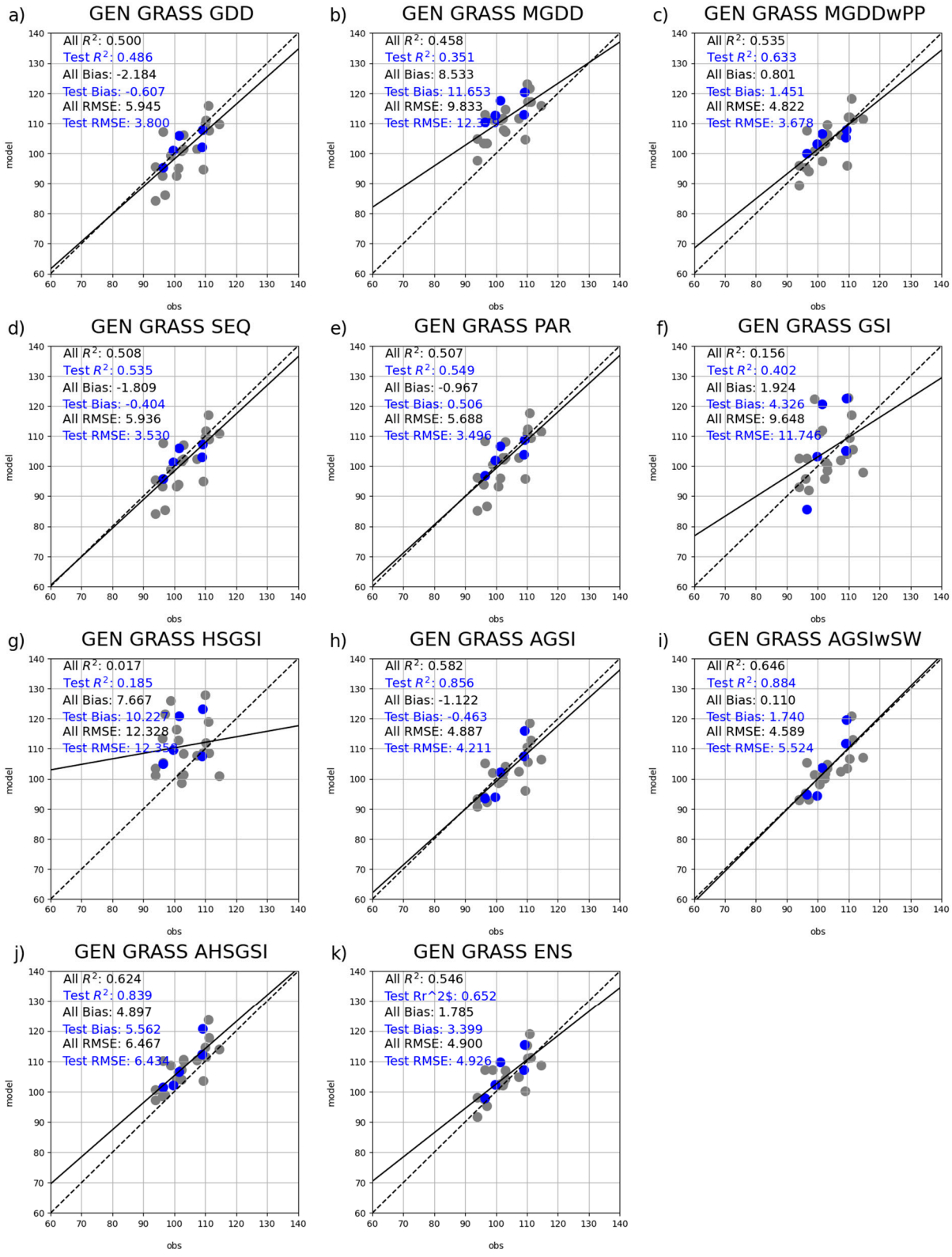


Figure S6. Spatial scatterplots of grassland-type-specific general calibrations. Each point represents the annual average of ~12,000 pixels. Gray points represent training data, and blue points indicating validation years. The dashed line denotes the 1:1 relationship, while the solid line is the result of the linear regression. Statistical metrics (Pearson r², bias, and RMSE) are reported for both the full dataset (including training and validation years), with test-year values highlighted in blue.

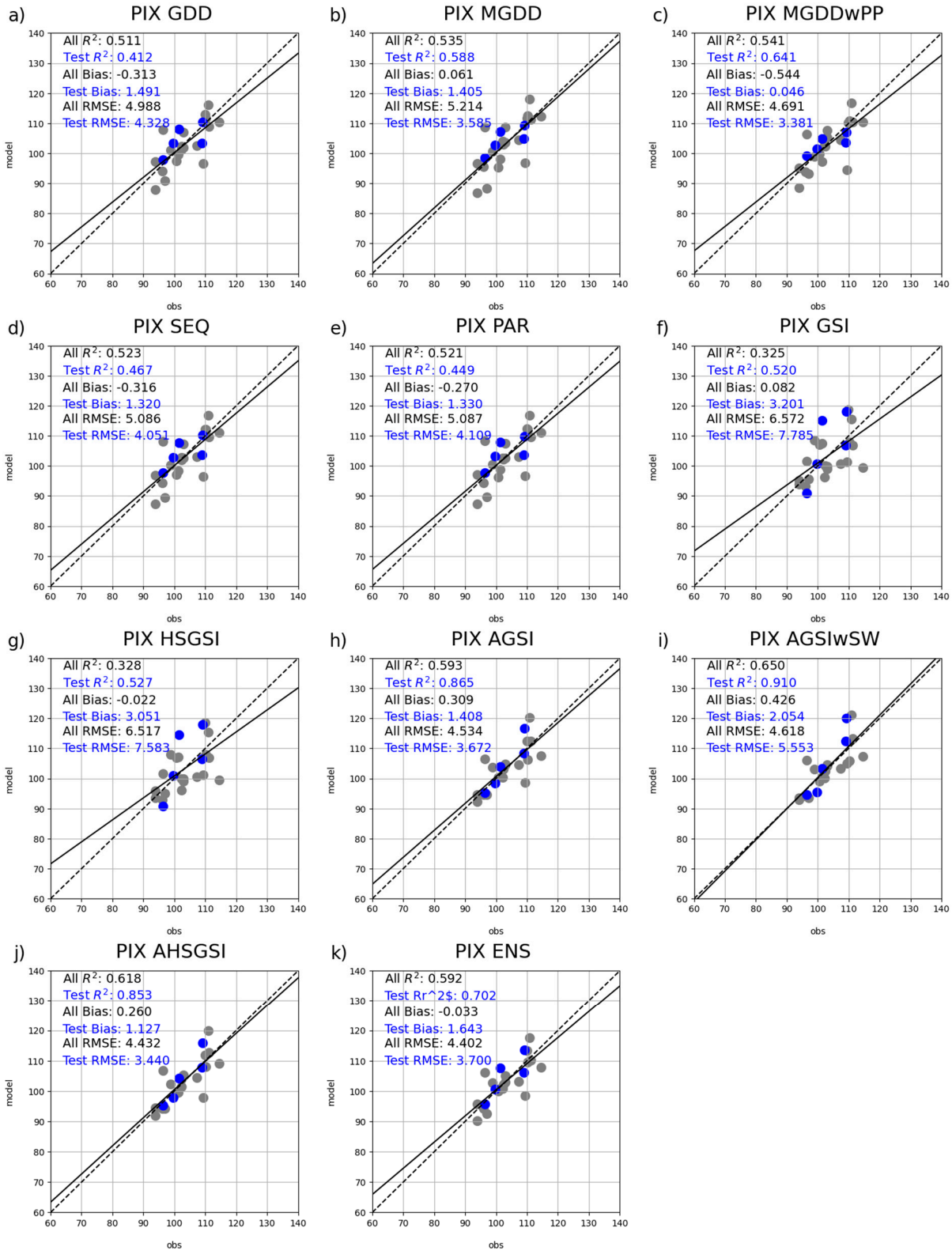


Figure S7. Spatial scatterplots of pixel-level calibrations. Each point corresponds to the annual average of approximately 12,000 pixels. Gray points represent training data, and blue points indicating validation years. The dashed line denotes the 1:1 relationship, while the solid line is the result of the linear regression. Statistical metrics (Pearson r², bias, and RMSE) are reported for both the full dataset (including training and validation years), with test-year values highlighted in blue.

Table S6. Pearson r values for each phenology model under three calibration schemes (GEN, GEN GRASS, and PIX) computed using data from the validation years. The statistics represent pixel-level performance, quantifying the spatial correspondence of SOS patterns across models computed from independent validation years.

Model Name	GEN	GEN GRASS	PIX
GDD	-0.045	0.417	0.902
MGDD	-0.058	0.283	0.897
MGDDwPP	-0.06	0.495	0.919
SEQ	-0.036	0.44	0.902
PAR	-0.051	0.395	0.901
GSI	-0.199	0.17	0.851
HSGSI	-0.269	-0.159	0.852
AGSI	-0.068	0.431	0.921
AGSIwSW	0.28	0.553	0.898
AHSGSI	-0.073	0.422	0.924
ENS	-0.07	0.401	0.926

Table S7. Statistics for each phenology model and the corrected models under three calibration schemes (GEN, GEN GRASS, and PIX) computed using data from all years. Note, that the statistical values presented here differ from Table 2 and Table S6, because those are calculated only for the validation years. The statistics represent pixel-level performance, quantifying the spatial correspondence of SOS patterns across models based on independent validation data.

Model Name	r	R^2	Bias	RMSE	r of the corrected models	R^2 of the corrected models	Bias of the corrected models	RMSE of the corrected models
GEN GDD	0.000	0.000	-1.6	8.1	0.000	0.000	0.0	7.9
GEN MGDD	-0.052	0.003	1.9	8.1	-0.052	0.003	0.0	7.9
GEN MGDDwPP	-0.061	0.004	0.3	7.6	-0.061	0.004	0.0	7.6
GEN SEQ	0.004	0.000	-2.4	8.3	0.004	0.000	0.0	7.9
GEN PAR	-0.010	0.000	-2.3	8.3	-0.010	0.000	0.0	8.0
GEN GSI	-0.069	0.005	1.8	8.4	-0.069	0.005	0.0	8.2
GEN HSGSI	-0.126	0.016	2.6	8.7	-0.126	0.016	0.0	8.3
GEN AGSI	-0.025	0.001	-0.8	8.1	-0.025	0.001	0.0	8.1
GEN AGSIwSW	0.373	0.139	0.5	7.0	0.373	0.139	0.0	6.9
GEN AHSGSI	-0.004	0.000	-0.6	7.8	-0.004	0.000	0.0	7.8
GEN GRASS GDD	0.429	0.184	-2.2	7.3	0.429	0.184	0.0	6.9
GEN GRASS MGDD	0.349	0.121	8.5	11.0	0.349	0.121	0.0	6.9
GEN GRASS MGDDwPP	0.494	0.245	0.8	6.8	0.494	0.245	0.0	6.7
GEN GRASS SEQ	0.433	0.187	-1.8	7.2	0.433	0.187	0.0	7.0
GEN GRASS PAR	0.413	0.171	-1.0	7.0	0.413	0.171	0.0	6.9
GEN GRASS GSI	0.328	0.108	1.9	7.4	0.328	0.108	0.0	7.2
GEN GRASS HSGSI	-0.198	0.039	7.7	11.6	-0.198	0.039	0.0	8.7
GEN GRASS AGSI	0.439	0.192	-1.1	7.1	0.439	0.192	0.0	7.0
GEN GRASS AGSIwSW	0.578	0.334	0.1	6.1	0.578	0.334	0.0	6.1
GEN GRASS AHSGSI	-0.017	0.000	4.9	9.2	-0.017	0.000	0.0	7.8
PIX GDD	0.958	0.918	-0.3	2.1	0.958	0.918	0.0	2.1
PIX MGDD	0.955	0.912	0.1	2.2	0.955	0.912	0.0	2.2
PIX MGDDwPP	0.972	0.944	-0.5	1.8	0.972	0.944	0.0	1.8
PIX SEQ	0.957	0.916	-0.3	2.2	0.957	0.916	0.0	2.2
PIX PAR	0.956	0.915	-0.3	2.2	0.956	0.915	0.0	2.2
PIX GSI	0.939	0.881	0.1	3.2	0.939	0.881	0.0	3.2
PIX HSGSI	0.938	0.880	0.0	3.2	0.938	0.880	0.0	3.2
PIX AGSI	0.978	0.957	0.3	1.6	0.978	0.957	0.0	1.5
PIX AGSIwSW	0.972	0.944	0.4	1.8	0.972	0.944	0.0	1.8
PIX AHSGSI	0.978	0.957	0.3	1.6	0.978	0.957	0.0	1.5

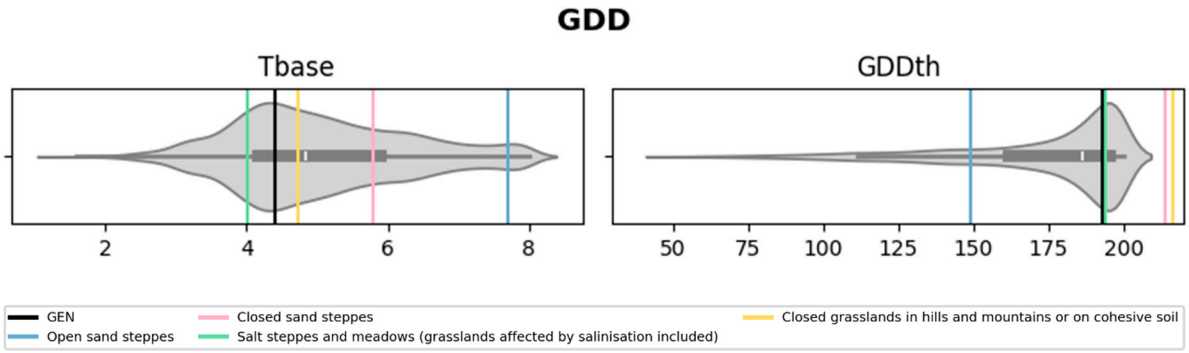


Figure S8. Violin plots illustrating the distribution of PIX-calibrated parameters for the GDD model. The black vertical line represents the GEN calibration value, while the blue, pink, green, and yellow vertical lines indicate the GEN GRASS calibrated parameters for Open sand steppes, Closed sand steppes, Salt steppes and meadows (grasslands affected by salinisation included), and Closed grasslands in hills and mountains or on cohesive soil, respectively.

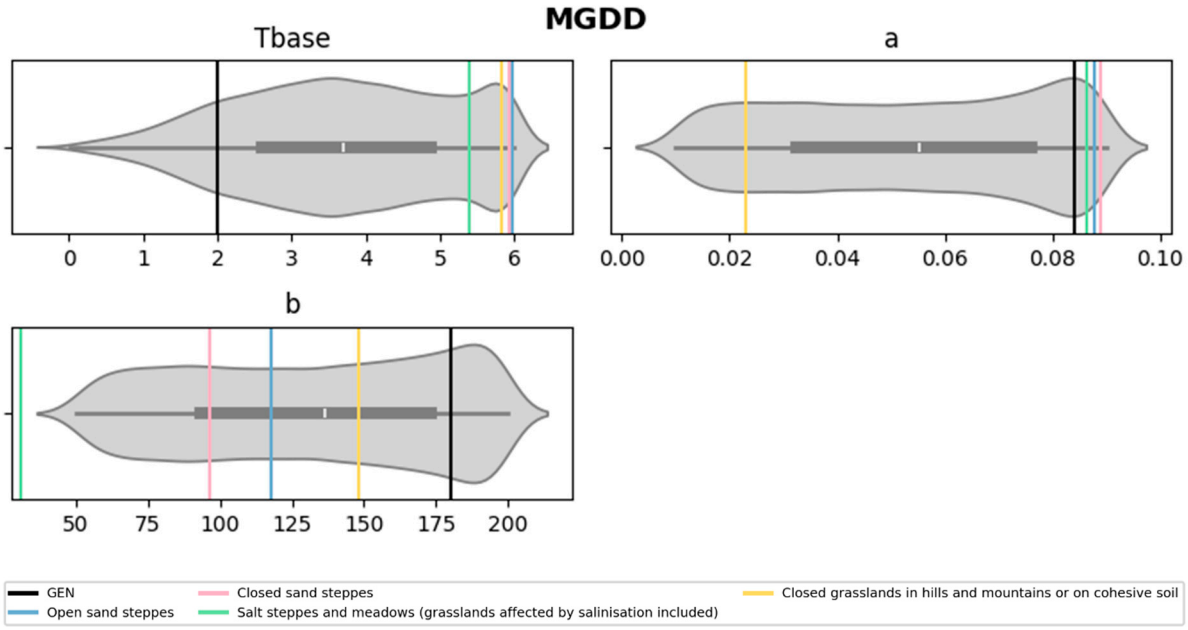


Figure S9. Violin plots illustrating the distribution of PIX-calibrated parameters for the MGDD model. The black vertical line represents the GEN calibration value, while the blue, pink, green, and yellow vertical lines indicate the GEN GRASS calibrated parameters for Open sand steppes, Closed sand steppes, Salt steppes and meadows (grasslands affected by salinisation included), and Closed grasslands in hills and mountains or on cohesive soil, respectively.

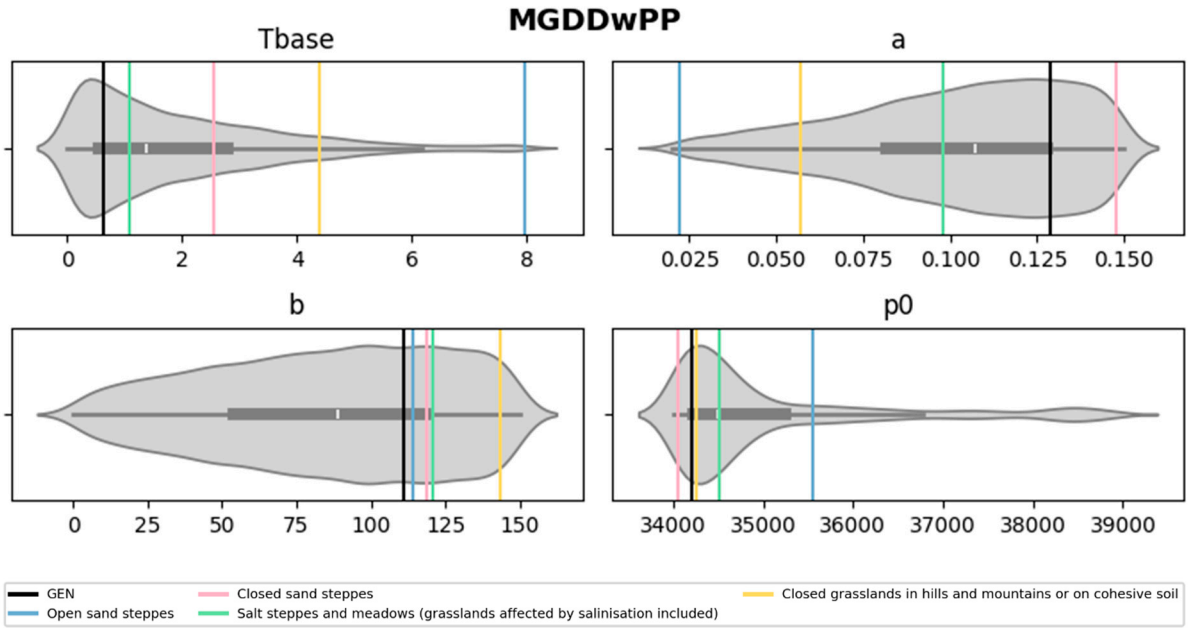


Figure S10. Violin plots illustrating the distribution of PIX-calibrated parameters for the MGDDwPP model. The black vertical line represents the GEN calibration value, while the blue, pink, green, and yellow vertical lines indicate the GEN GRASS calibrated parameters for Open sand steppes, Closed sand steppes, Salt steppes and meadows (grasslands affected by salinisation included), and Closed grasslands in hills and mountains or on cohesive soil, respectively.

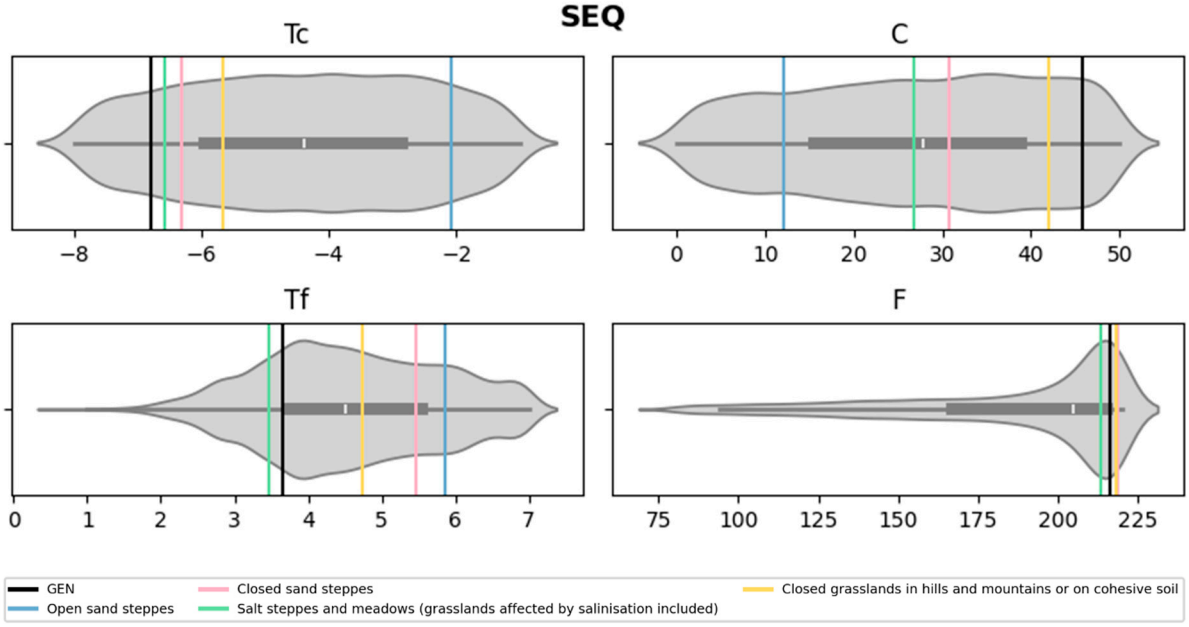


Figure S11. Violin plots illustrating the distribution of PIX-calibrated parameters for the SEQ model. The black vertical line represents the GEN calibration value, while the blue, pink, green, and yellow vertical lines indicate the GEN GRASS calibrated parameters for Open sand steppes, Closed sand steppes, Salt steppes and meadows (grasslands affected by salinisation included), and Closed grasslands in hills and mountains or on cohesive soil, respectively.

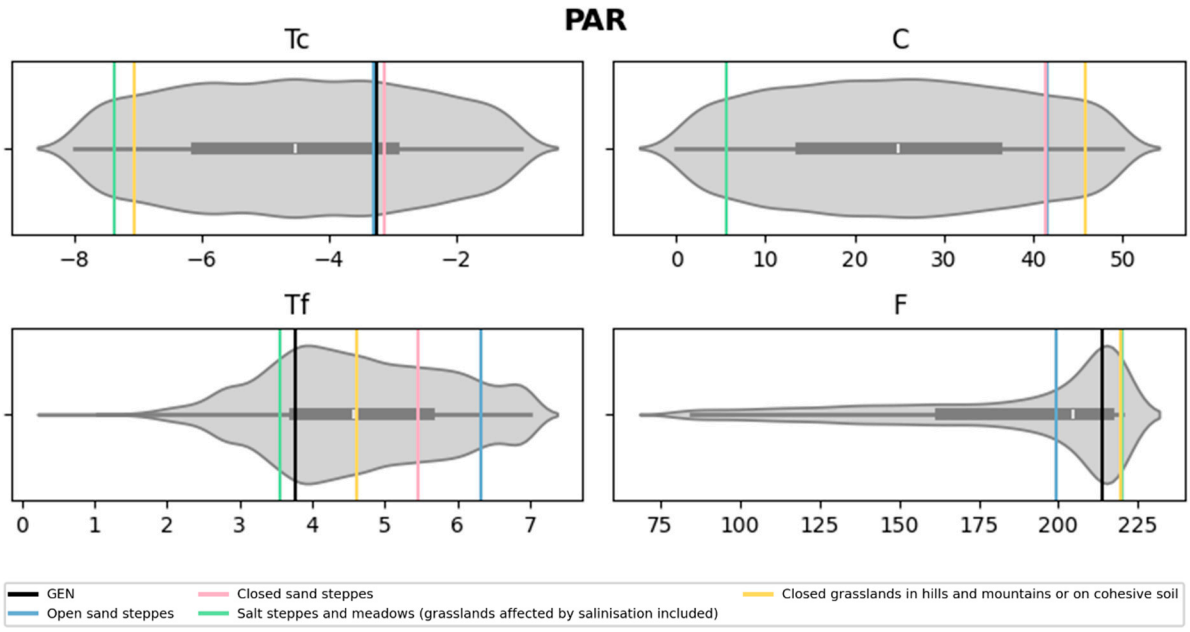


Figure S12. Violin plots illustrating the distribution of PIX-calibrated parameters for the PAR model. The black vertical line represents the GEN calibration value, while the blue, pink, green, and yellow vertical lines indicate the GEN GRASS calibrated parameters for Open sand steppes, Closed sand steppes, Salt steppes and meadows (grasslands affected by salinisation included), and Closed grasslands in hills and mountains or on cohesive soil, respectively.

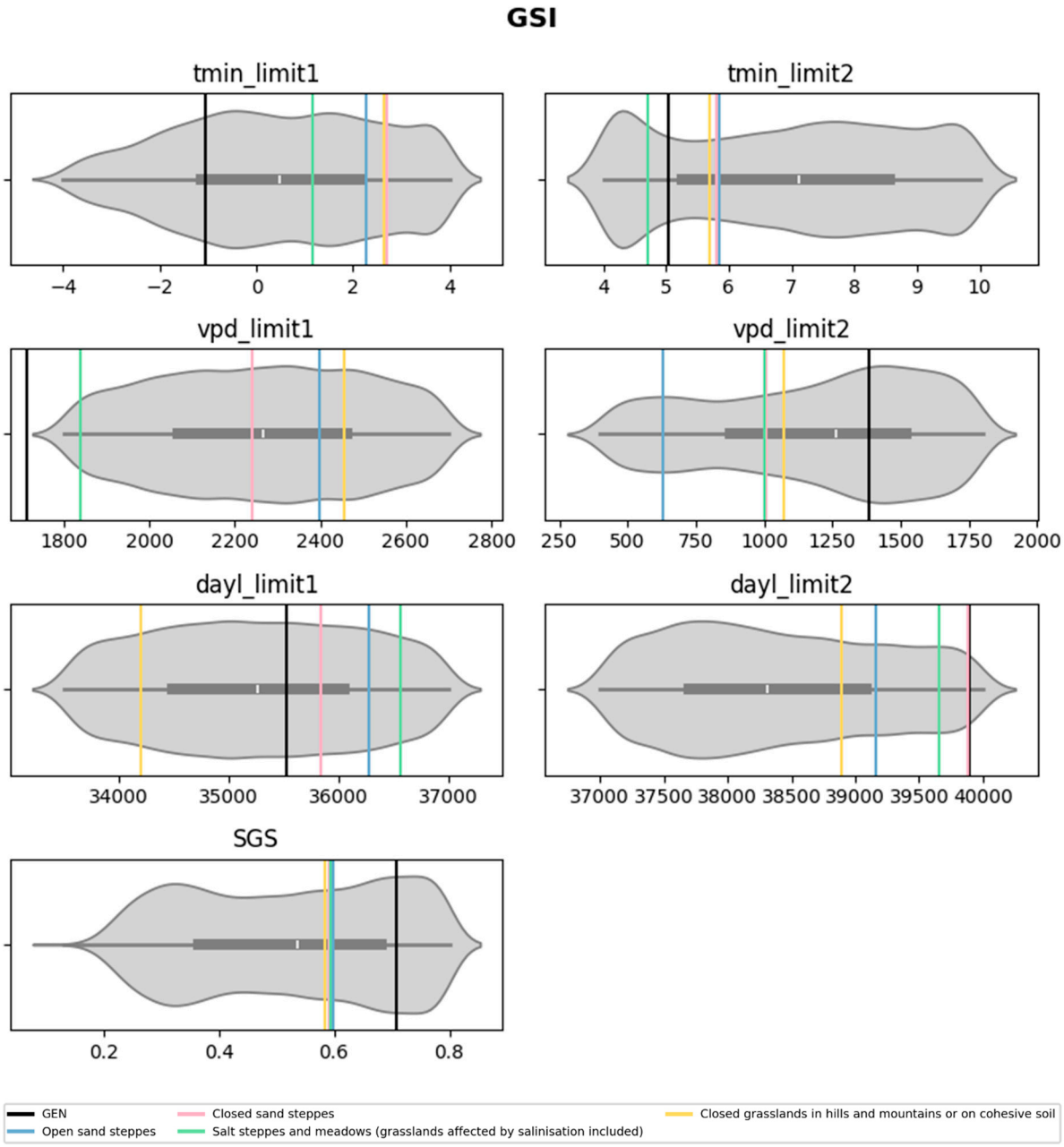


Figure S13. Violin plots illustrating the distribution of PIX-calibrated parameters for the GSI model. The black vertical line represents the GEN calibration value, while the blue, pink, green, and yellow vertical lines indicate the GEN GRASS calibrated parameters for Open sand steppes, Closed sand steppes, Salt steppes and meadows (grasslands affected by salinisation included), and Closed grasslands in hills and mountains or on cohesive soil, respectively.

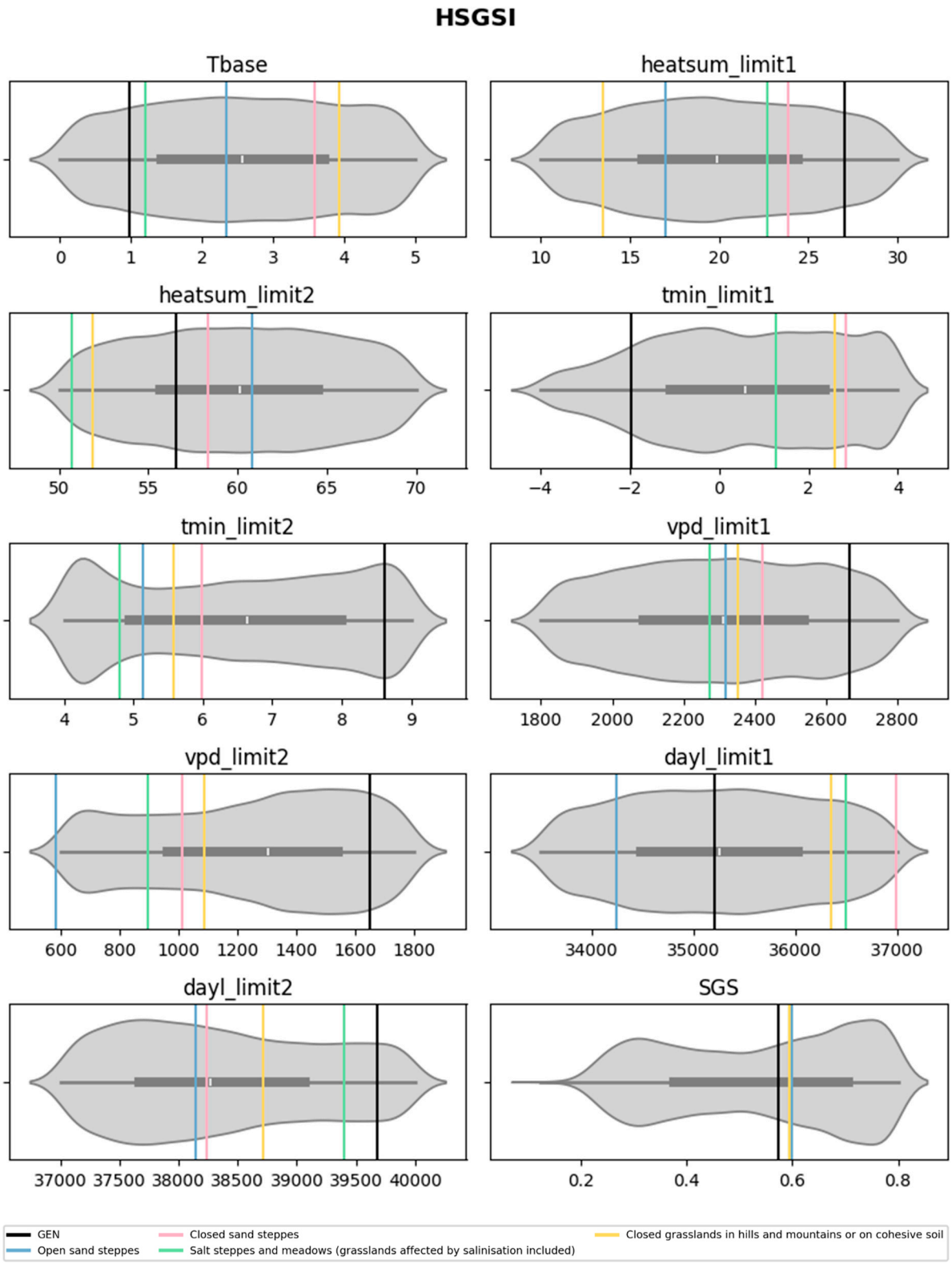


Figure S14. Violin plots illustrating the distribution of PIX-calibrated parameters for the HSGSI model. The black vertical line represents the GEN calibration value, while the blue, pink, green, and yellow vertical lines indicate the GEN GRASS calibrated parameters for Open sand steppes, Closed sand steppes, Salt steppes and meadows (grasslands affected by salinisation included), and Closed grasslands in hills and mountains or on cohesive soil, respectively.

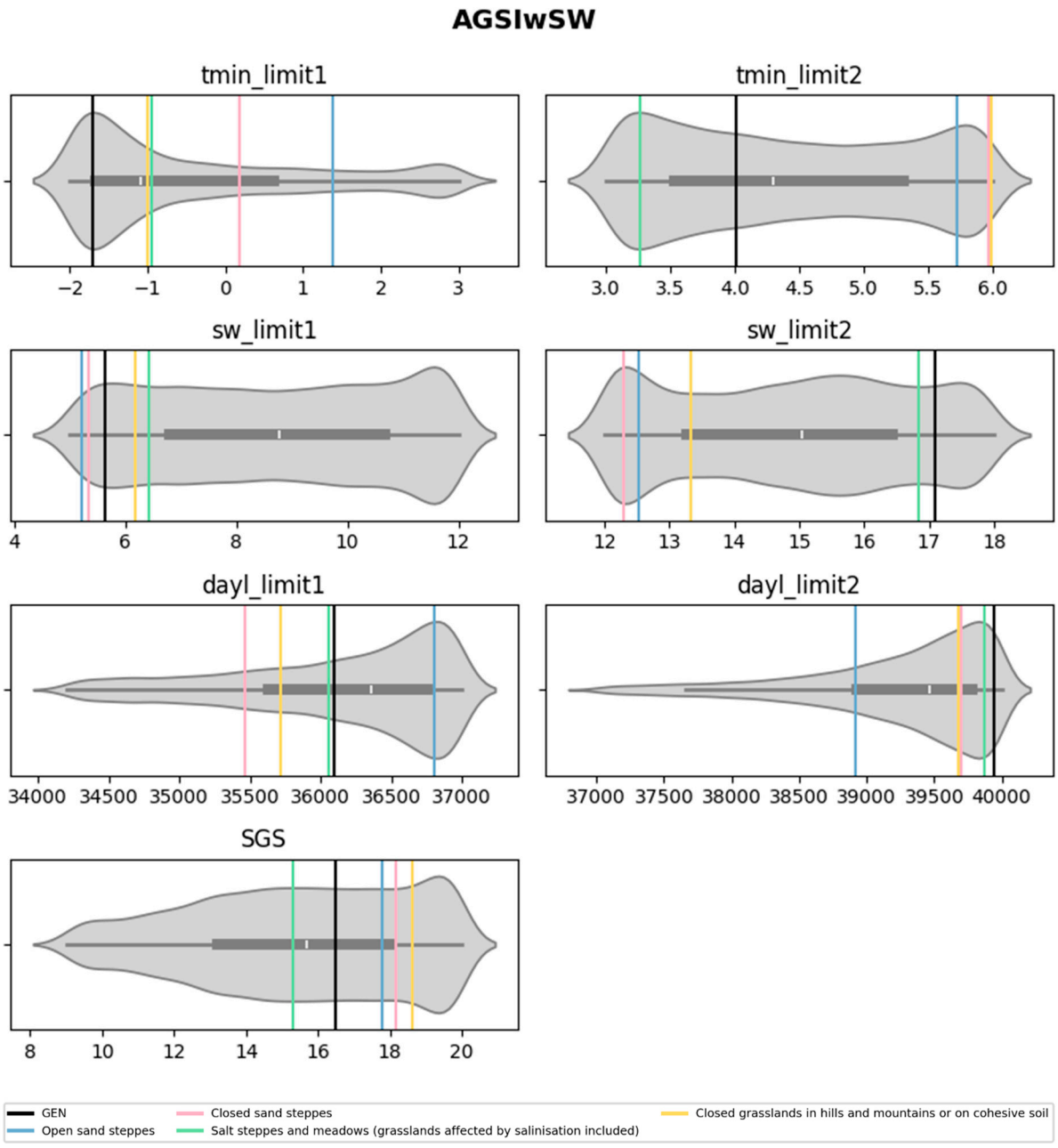


Figure S15. Violin plots illustrating the distribution of PIX-calibrated parameters for the AGSIwSW model. The black vertical line represents the GEN calibration value, while the blue, pink, green, and yellow vertical lines indicate the GEN GRASS calibrated parameters for Open sand steppes, Closed sand steppes, Salt steppes and meadows (grasslands affected by salinisation included), and Closed grasslands in hills and mountains or on cohesive soil, respectively.

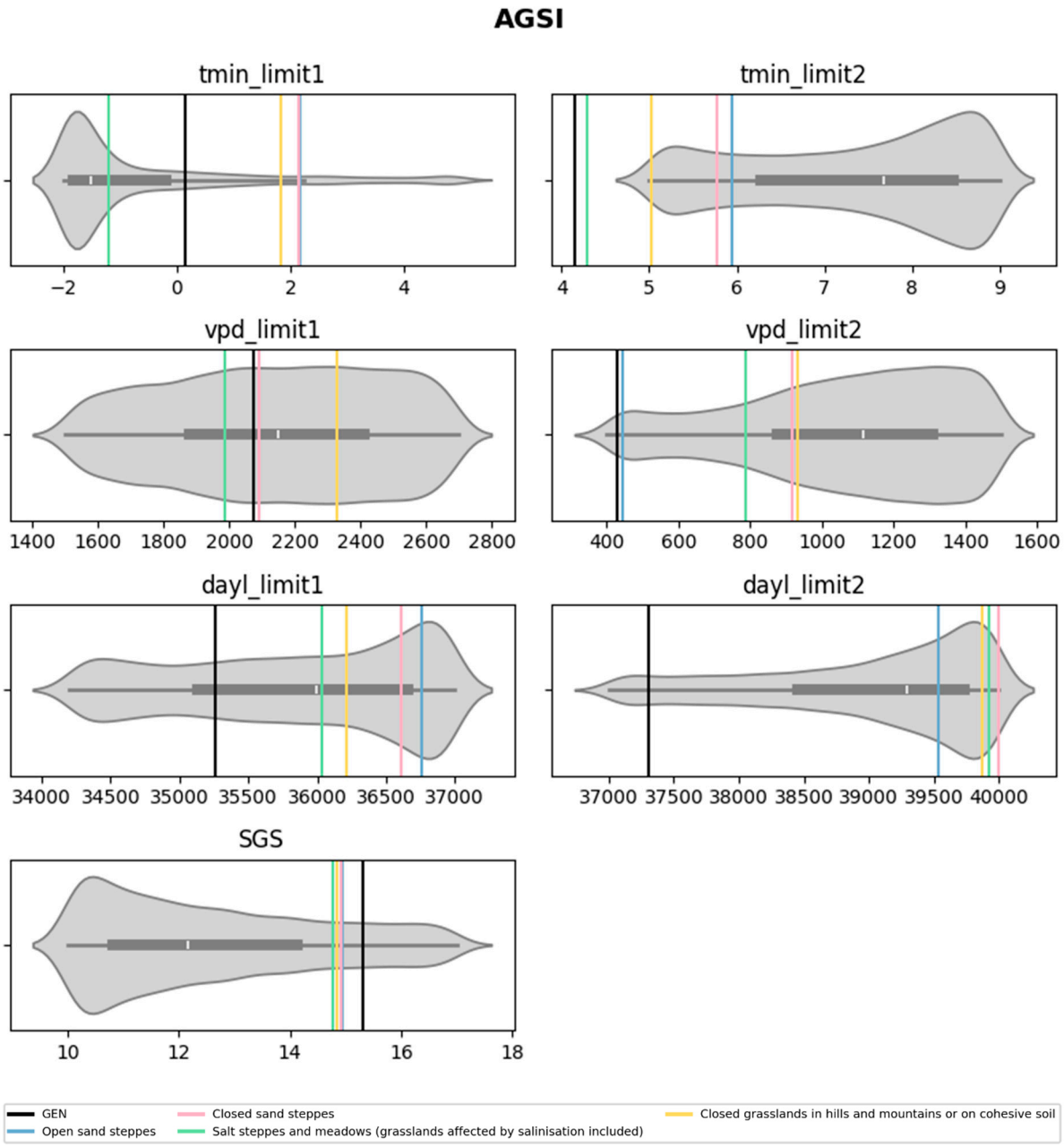


Figure S16. Violin plots illustrating the distribution of PIX-calibrated parameters for the AGSI model. The black vertical line represents the GEN calibration value, while the blue, pink, green, and yellow vertical lines indicate the GEN GRASS calibrated parameters for Open sand steppes, Closed sand steppes, Salt steppes and meadows (grasslands affected by salinisation included), and Closed grasslands in hills and mountains or on cohesive soil, respectively.

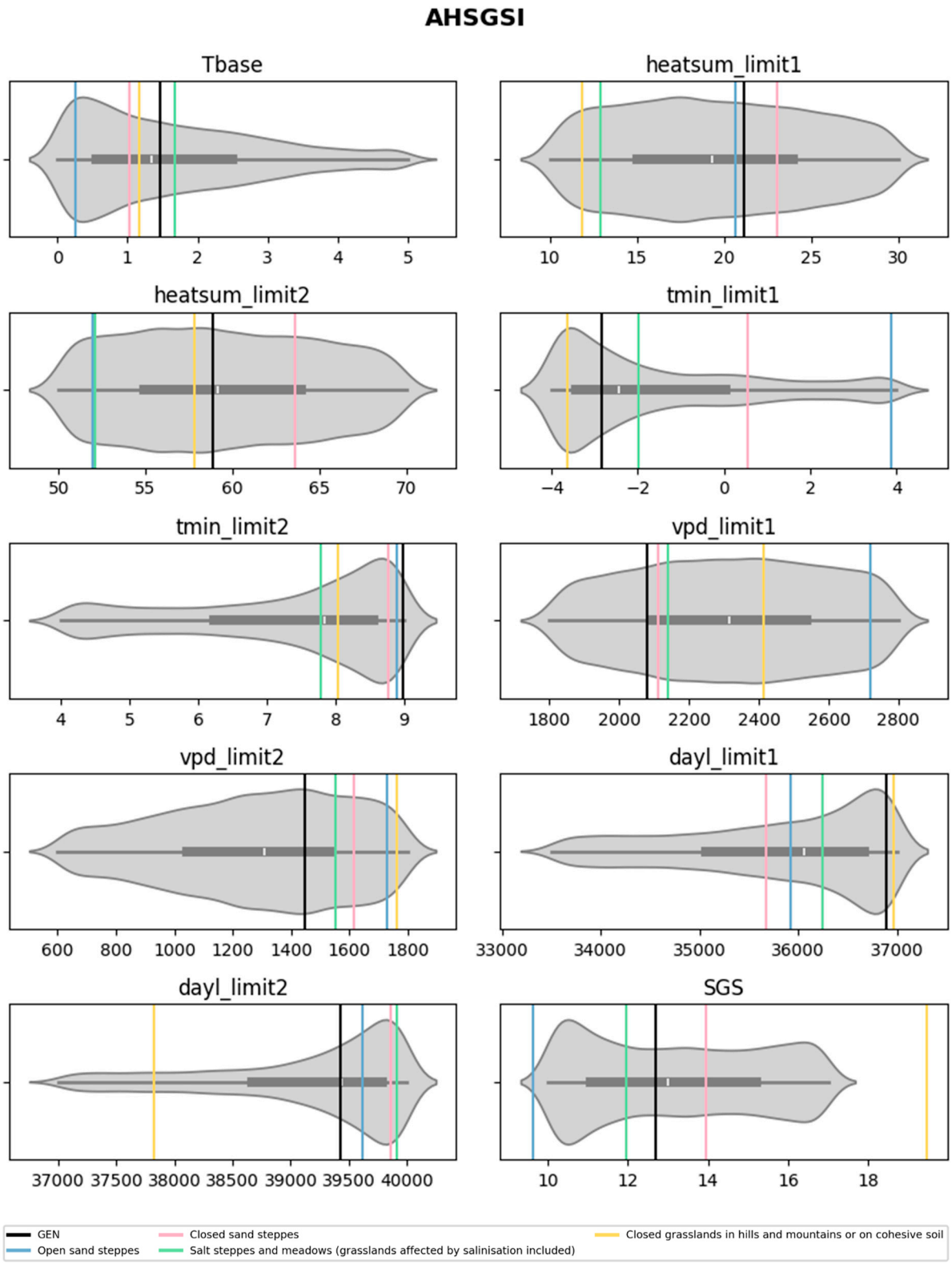


Figure S17. Violin plots illustrating the distribution of PIX-calibrated parameters for the AHSGSI model. The black vertical line represents the GEN calibration value, while the blue, pink, green, and yellow vertical lines indicate the GEN GRASS calibrated parameters for Open sand steppes, Closed sand steppes, Salt steppes and meadows (grasslands affected by salinisation included), and Closed grasslands in hills and mountains or on cohesive soil, respectively.

Table S8. Parameter intervals (minimum-maximum) for the heat accumulation-based models. The numbers displayed beneath each parameter name represent the parameter intervals: the left-hand value indicates the minimum (lower bound), and the right-hand value indicates the maximum (upper bound). The ranges were determined from expert knowledge, published literature, and empirical experience. The first rows within each model type (labeled *GEN parameters*) show the calibrated parameter values obtained from the general calibration. Within each model type (GDD, MGDD, MGDDwPP, SEQ, and PAR), the GEN, GEN GRASS, and PIX rows correspond to the parameter sets used for different calibration categories.

GDD	**Tbase**		**GDDth**	
GEN parameters	4.4		193	
GEN	0	8	100	220
GEN GRASS	0	8	80	220
PIX	0	8	50	200

MGDD	**Tbase**		**a**		**b**	
GEN parameters	2		0.1		180	
GEN	0	8	0.01	0.16	0	200
GEN GRASS	0	6	0.001	0.09	0	150
PIX	0	6	0.001	0.09	50	200

MGDDwPP	**Tbase**		**a**		**b**		**p0**	
GEN parameters	0.6		0.1		111		34212	
GEN	0	8	0.01	0.16	0	200	33500	39000
GEN GRASS	0	8	0.02	0.15	0	150	34000	39000
PIX	0	8	0.02	0.15	0	150	34000	39000

SEQ	**Tc**		**C**		**Tf**		**F**	
GEN parameters	-6.8		45.8		3.7		216	
GEN	-8	-1	0	50	0	7	80	220
GEN GRASS	-8	-1	0	50	0	7	80	220
PIX	-8	-1	0	50	0	7	80	220

PAR	**Tc**		**C**		**Tf**		**F**	
GEN parameters	-3.3		41.5		3.8		214	
GEN	-8	-1	0	50	0	7	80	220
GEN GRASS	-8	-1	0	50	0	7	80	220
PIX	-8	-1	0	50	0	7	80	220

Table S9. Parameter intervals (minimum-maximum) for the GSI-type models. The numbers displayed beneath each parameter name represent the parameter intervals: the left-hand value indicates the minimum (lower bound), and the right-hand value indicates the maximum (upper bound). The ranges were determined from expert knowledge, published literature, and empirical experience. The first rows within each model type (labeled *GEN parameters*) show the calibrated parameter values obtained from the general calibration. Within each model type (GSI, AGSI, AGSIwSW, HSGSI, and AHGSI), the GEN, GEN GRASS, and PIX rows correspond to the parameter sets used for different calibration categories.

GSI	**tmin_limit1**		**tmin_limit2**		**vpd_limit1**		**vpd_limit2**		**dayl_limit1**		**dayl_limit2**		**SGS**	
GEN parameters	-1		5		1715		1386		35524		39893		0.7	
GEN	-2	5	5	9	1600	2700	400	1600	34200	37000	37000	40000	0.1	0.8
GEN GRASS	-2	3	3	6	1500	2500	400	1100	34200	37000	37000	40000	0.3	0.6
PIX	-4	4	4	10	1800	2700	400	1800	33500	37000	37000	40000	0.1	0.8
AGSI	**tmin_limit1**		**tmin_limit2**		**vpd_limit1**		**vpd_limit2**		**dayl_limit1**		**dayl_limit2**		**SGS**	
GEN parameters	0.2		4.1		2072		428		35260		37310		15.3	
GEN	-2	3	3	6	1500	2500	400	1100	34200	37000	37000	40000	9	20
GEN GRASS	-2	3	3	6	1500	2500	400	1100	34200	37000	37000	40000	10	17
PIX	-2	5	5	9	1500	2700	400	1500	34200	37000	37000	40000	10	17
AGSIwSW	**tmin_limit1**		**tmin_limit2**		**sw_limit1**		**sw_limit2**		**dayl_limit1**		**dayl_limit2**		**SGS**	
GEN parameters	-1.7		4		6		17		36095		39943		16.5	
GEN	-2	3	3	6	5	12	12	18	34200	37000	37000	40000	9	17
GEN GRASS	-2	3	3	6	5	12	12	18	34200	37000	37000	40000	9	20
PIX	-2	3	3	6	5	12	12	18	34200	37000	37000	40000	9	20

HSGSI	**Tbase**	**heatsum_limit1**	**heatsum_limit2**	**tmin_limit1**	**tmin_limit2**	**vpd_limit1**	**vpd_limit2**	**dayl_limit1**	**dayl_limit2**	**SGS**
GEN parameters	1	27	57	-2	8.6	2666	1648	35208	39683	0.6

GEN	0	5	10	30	50	70	-3	6	6	9	1800	2700	400	1800	34200	37000	37000	40000	0.1	0.8
GEN GRASS	0	5	10	30	50	70	-2	3	3	6	1500	2500	400	1100	34200	37000	37000	40000	0.3	0.6
PIX	0	5	10	30	50	70	-4	4	4	9	1800	2800	600	1800	33500	37000	37000	40000	0.1	0.8

AHGSI	**Tbase**	**heatsum_limit1**	**heatsum_limit2**	**tmin_limit1**	**tmin_limit2**	**vpd_limit1**	**vpd_limit2**	**dayl_limit1**	**dayl_limit2**	**SGS**
GEN parameters	1.5	21	59	-2.8	9	2079	1447	36894	39425	12.7

GEN	0	5	10	30	50	70	-3	6	6	9	1800	2700	400	1800	34200	37000	37000	40000	9	17
GEN GRASS	0	5	10	30	50	70	-4	4	4	9	1800	2800	600	1800	33500	37000	37000	40000	1	20
PIX	0	5	10	30	50	70	-4	4	4	9	1800	2800	600	1800	33500	37000	37000	40000	10	17

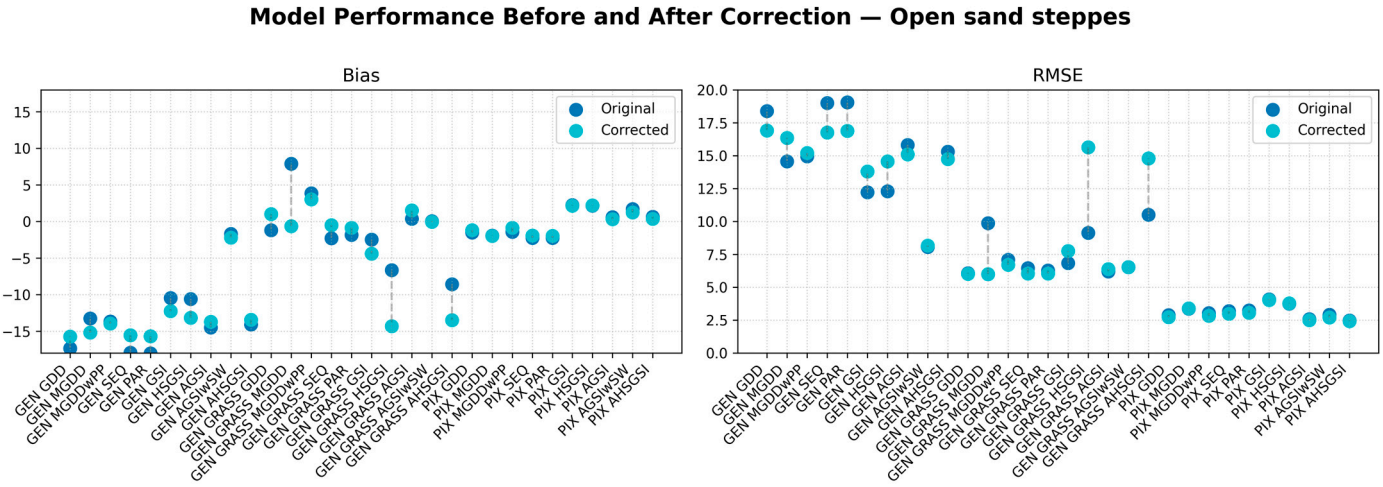


Figure S18. Performance of the evaluated models for Open sand steppes, shown in terms of bias and RMSE before (Original) and after bias correction (Corrected). Each point represents a model, while dashed vertical lines connect original and corrected values for the same model to enhance visual clarity.

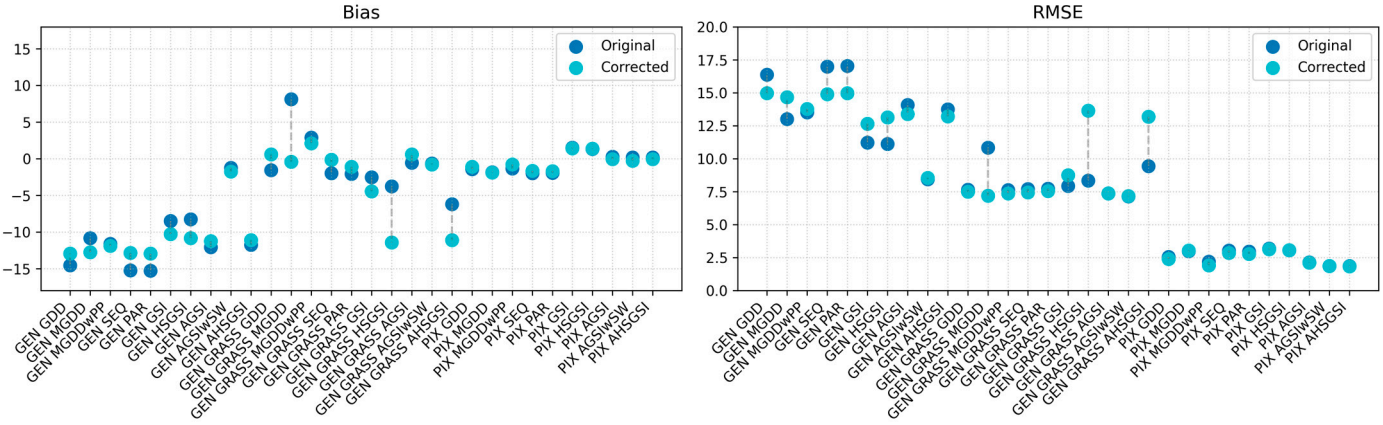


Figure S19. Performance of the evaluated models for Closed sand steppes, shown in terms of bias and RMSE before (Original) and after bias correction (Corrected). Each point represents a model, while dashed vertical lines connect original and corrected values for the same model to enhance visual clarity.

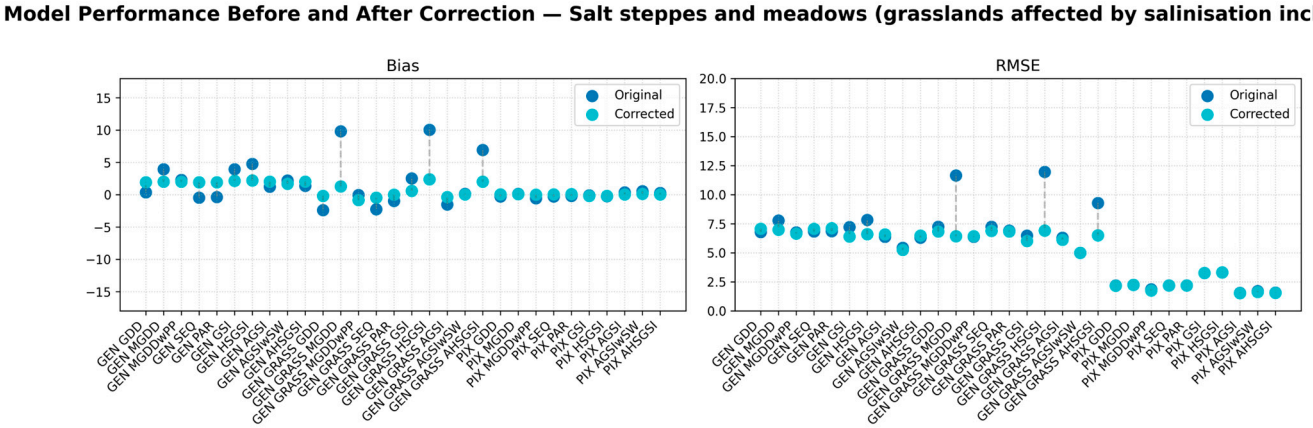


Figure S20. Performance of the evaluated models for Salt steppes and meadows (grasslands affected by salinisation included), shown in terms of bias and RMSE before (Original) and after bias correction (Corrected). Each point represents a model, while dashed vertical lines connect original and corrected values for the same model to enhance visual clarity.

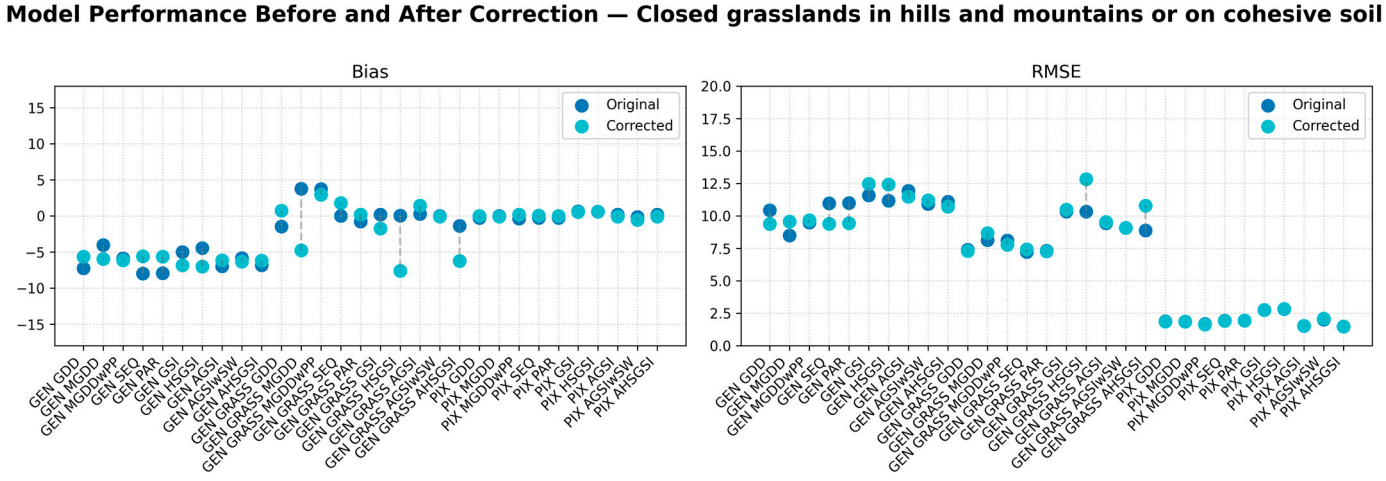


Figure S21. Performance of the evaluated models for Closed grasslands in hills and mountains or on cohesive soil, shown in terms of bias and RMSE before (Original) and after bias correction (Corrected). Each point represents a model, while dashed vertical lines connect original and corrected values for the same model to enhance visual clarity.

References

1. Jeong, S.J.; Medvigy, D.; Shevliakova, E.; Malyshev, S. Uncertainties in Terrestrial Carbon Budgets Related to Spring Phenology. *J Geophys Res Biogeosci* 2012, *117*, 1–17, doi:10.1029/2011JG001868.

2. Melaas, E.K.; Richardson, A.D.; Friedl, M.A.; Dragoni, D.; Gough, C.M.; Herbst, M.; Montagnani, L.; Moors, E. Using FLUXNET Data to Improve Models of Springtime Vegetation Activity Onset in Forest Ecosystems. *Agric For Meteorol* 2013, *171–172*, 46–56, doi:10.1016/j.agrformet.2012.11.018.

3. Xin, Q.; Broich, M.; Zhu, P.; Gong, P. Modeling Grassland Spring Onset across the Western United States Using Climate Variables and MODIS-Derived Phenology Metrics. *Remote Sens Environ* 2015, *161*, 63–77, doi:10.1016/j.rse.2015.02.003.

4. White, M.A.; Thornton, P.E.; Running, S.W. A Continental Phenology Model for Monitoring Vegetation Responses to Interannual Climatic Variability. *Global Biogeochem Cycles* 1997, *11*, 217–234, doi:10.1029/97GB00330.

5. Jolly, W.M.; Nemani, R.; Running, S.W. A Generalized, Bioclimatic Index to Predict Foliar Phenology in Response to Climate. *Glob Chang Biol* 2005, *11*, 619–632, doi:10.1111/j.1365-2486.2005.00930.x.

6. Hidy, D.; Barcza, Z.; Haszpra, L.; Churkina, G.; Pintér, K.; Nagy, Z. Development of the Biome-BGC Model for Simulation of Managed Herbaceous Ecosystems. *Ecol Modell* 2012, *226*, 99–119, doi:10.1016/j.ecolmodel.2011.11.008.

7. Hidy, D.; Barcza, Z.; Marjanovic, H.; Sever, M.Z.O.; Dobor, L.; Gelybó, G.; Fodor, N.; Pintér, K.; Churkina, G.; Running, S.; et al. Terrestrial Ecosystem Process Model Biome-BGCMuSo v4.0: Summary of Improvements and New Modeling Possibilities. *Geosci Model Dev* 2016, *9*, 4405–4437, doi:10.5194/gmd-9-4405-2016.

8. Asseng, S.; Ewert, F.; Rosenzweig, C.; Jones, J.W.; Hatfield, J.L.; Ruane, A.C.; Boote, K.J.; Thorburn, P.J.; Rötter, R.P.; Cammarano, D.; et al. Uncertainty in Simulating Wheat Yields under Climate Change. *Nat Clim Chang* 2013, *3*, 827–832, doi:10.1038/nclimate1916.

9. Sándor, R.; Barcza, Z.; Hidy, D.; Lellei-Kovács, E.; Ma, S.; Bellocchi, G. Modelling of Grassland Fluxes in Europe: Evaluation of Two Biogeochemical Models. *Agric Ecosyst Environ* 2016, *215*, 1–19, doi:10.1016/j.agee.2015.09.001.