Supplementary material

Table of Contents

S1. Sensitivity analysis

The Morris or elementary effects screening method (Morris, 1991) was used to conduct a qualitative global sensitivity analysis on phenological development of maize. Sensitivity analysis was only performed for site-year 6_2010 under the assumption that ranks of the most sensitive parameters would not change significantly due do the different weather and initial conditions in Kraichgau and the Swabian Alb. The *sensitivity* package in R (Bertrand Iooss et al., 2020) was used. The one-at-a-time (OAT) design in the *morris* function was used to define the parameter vectors. A total of 11 parameters that influence phenological development in the SPASS model were pre-selected based on expert knowledge. Uniform parameter distributions with a range equal to three standard deviations from the expected value were used. It is noted that different distributions have been used for Bayesian calibration (platykurtic prior distribution) and sensitivity analysis (uniform distribution). However, this is assumed to have a limited influence in identifying the most sensitive parameters. Settings to the *morris* function were provided: 1000 samples, 10 levels and a grid jump-size of 2 units. Phenology was simulated using the SPASS model in XN5 software for all the proposed parameter vectors. The *morris* function was then used to estimate elementary effects (Cuntz et al., 2015; Morris, 1991) of phenological development at an interval of every 5 days within the growing season. The sensitivity measures, namely, the mean (μ^*) of the absolute value of the elementary effects of a parameter and the standard deviation (σ) were calculated on these days to evaluate parameter sensitivity over the growing season.

$$\mu^*_{\theta_i} = \frac{1}{N}\sum_{n=1}^{N} |ee_{n,\theta_i}| \quad \text{S1.}$$

$$\sigma_{\theta_i} = \sqrt{\frac{\sum_{n=1}^{N}(ee_{n,\theta_i} - \mu_{\theta_i})^2}{N}} \quad \text{S2.}$$

where $\mu^*_{\theta_i}$and σ_{θ_i} are the μ^* and σ sensitivity measures for the i^{th} parameter in the parameter vector θ, ee_n is the elementary effects for the n^{th} parameter vector, N are the total parameter vectors and $\mu_{\theta_i}$is given by:

$$\mu_{\theta_i} = \frac{1}{N}\sum_{n=1}^{N} ee_{n,\theta_i} \qquad \text{S3.}$$

Based on μ^*, the effective sowing depth (SOWDEPTH) was the most and only sensitive parameter during emergence, which is intuitive as the other parameters influence development after emergence (Fig. S1). Then the relative importance of parameters that define the cardinal temperatures (DELTMAX1, DELTOPT1 and TMINDEV1) and the physiological development days (PDD1) of the vegetative phase increased. These parameters continued to be the most influential parameters even through the generative phase of development. Even though DELTOPT2 and PDD2 are important parameters for the generative phase of development, their influence was small and over-shadowed by the influence of the vegetative phase parameters.

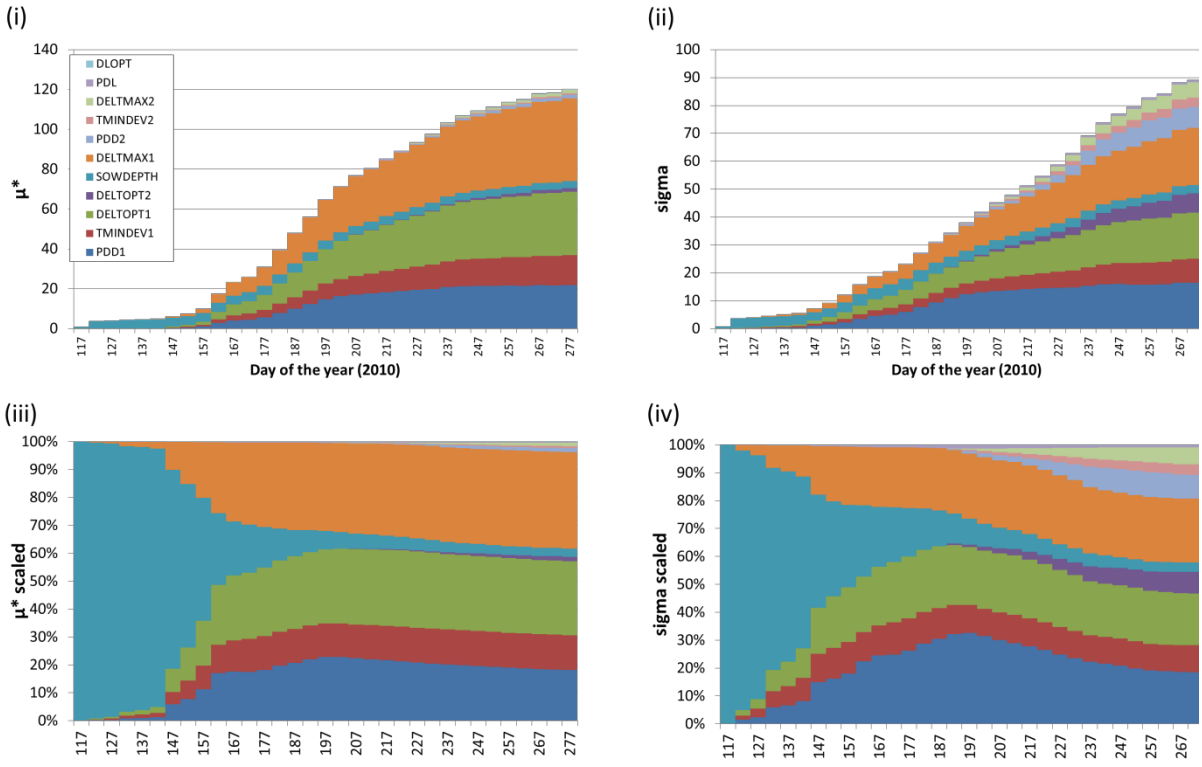

Figure S1: Plots of (i) $\boldsymbol{\mu}^*$ and (ii) sigma of elementary effects calculated for simulated phenological development at an interval of 5 days over the growing season of silage maize (between sowing day 112 and harvest day 278 of the year) at site 6 in the year 2010. The parameters that influence phenological development in the SPASS model are listed in the legend. Plots (iii) and (iv) are the normalized $\boldsymbol{\mu}^*$ and sigma values per day, respectively, expressed as a percentages.

S2. Estimation of information entropy

Information entropy (H) for a continuous distribution is given by:

$$H = -\int f(\theta)\ln\big(\mathrm{f}(\theta)\big)\mathrm{d}\theta \quad \text{S4.}$$

where $f(\theta)$ is the probability density function of θ.

Information entropy estimates of the posterior parameter distributions were obtained using the redistribution estimate equation (Beirlant et al., 1997):

$$H_n = -\frac{1}{n}\sum_{i=1}^{n} \ln f_n(\theta_i) \quad \text{S5.}$$

where H_n is the estimate of information entropy, f_n is the Kernel Density Estimate (KDE) and $\theta_1,\ldots\theta_n$ are independent and identically distributed (i.i.d.) parameter vector samples from the posterior distribution. The KDE was obtained by using the kde function from the *ks* package in R (Duong, 2020). Least Squares Cross-Validation (LSCV) was used for bandwidth selection.

S3. Residual analysis

Residuals were analysed for the synthetic and true sequences for simulated phenology at the maximum a posteriori probability (MAP) estimate of the model parameters. The residual plots provided in the following sections have been separated into the synthetic sequences (section S3.1), Swabian Alb true sequence (section S3.2), and Kraichgau true sequence (section S3.3).

Homoscedasticity was checked by plotting the residuals against days-after-sowing and simulated phenology (Fig. S2, Fig. S3, Fig. S8 – Fig. S13, Fig. S18 – Fig. S20). In general, heteroscedasticity was not observed. Normal assumption of the error model was verified by plotting histograms of the residuals and quantile-quantile plots (Fig. S4, Fig. S5, Fig. S14 – Fig. S16, Fig. S21). For the first few sequential updates, the number of observations were limited making a thorough analysis difficult. For the latter few sequential updates, the residuals were found to be nearly normal.

In the synthetic sequences, the residual error distribution was nearly normal (Fig. S4, Fig. S5). The slight skewness is attributed to model limitations (controlled cultivar-environment sequence) and specific site-years that had a different phenological development as compared to the remaining site-years in the calibration sequence (both synthetic sequences).

The slight skewness observed in the true sequence is attributed to model limitations where the model is unable to capture the slow development during the vegetative phase that was observed at a few site-years like 6_2013 (Fig. S14, Fig. S15, Fig. S16) and 5_2016 (Fig. S16). Autocorrelation was estimated after padding the dataset as the observations are not at regular time-intervals. Therefore, there is no ACF estimated at some lags. Figure S17: contains the autocorrelation (ACF) plot of the residuals after the model is calibrated to data from site-years 6_2010, 5_2011, 5_2012, 6_2013, 5_2015, and 5_2016. Based on the limited data with unequal lags, no autocorrelation was detected. However, it is suspected that with state variables like phenology, which are based on cumulative sums, autocorrelation of errors could

theoretically exist. However, due to data limitations, error modelling would be limited in its scope for improving the results.

S3.1 Synthetic sequences

In the ideal sequence where there is no model structural error, the skewness in the residual distribution (Fig. S4) is caused by site-year 2. This site-year exhibits a different development-behaviour as compared to other site-years in the calibration sequence (Fig. S6). In the controlled cultivar-environment sequence the slight skewness (Fig. S5) in the distribution of the residuals are caused due to two reasons. The site-year 9 exhibits a different phenological development-behaviour as compared to other site-years in the calibration sequence (Fig. S7). Additionally, the model is unable to capture the rapid growth seen in site-years 3, 4, 5, 8 and 9 between 82 and 110 days after sowing.

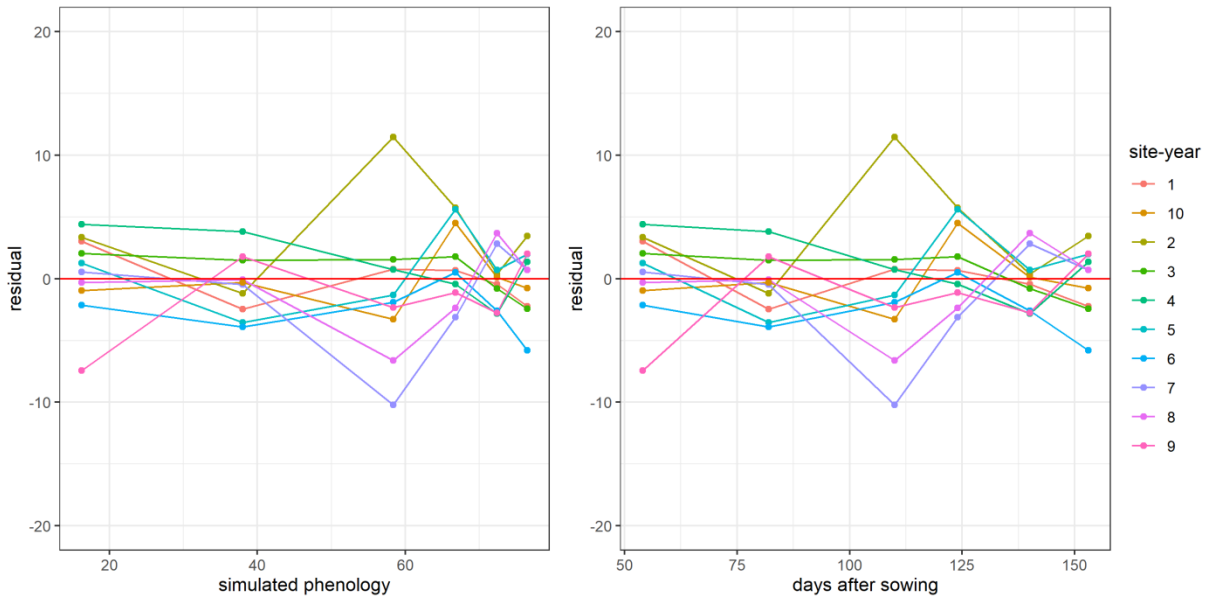

Figure S2: Residuals vs simulated phenology and days after sowing after calibration to 10 site-years in the ideal synthetic sequence

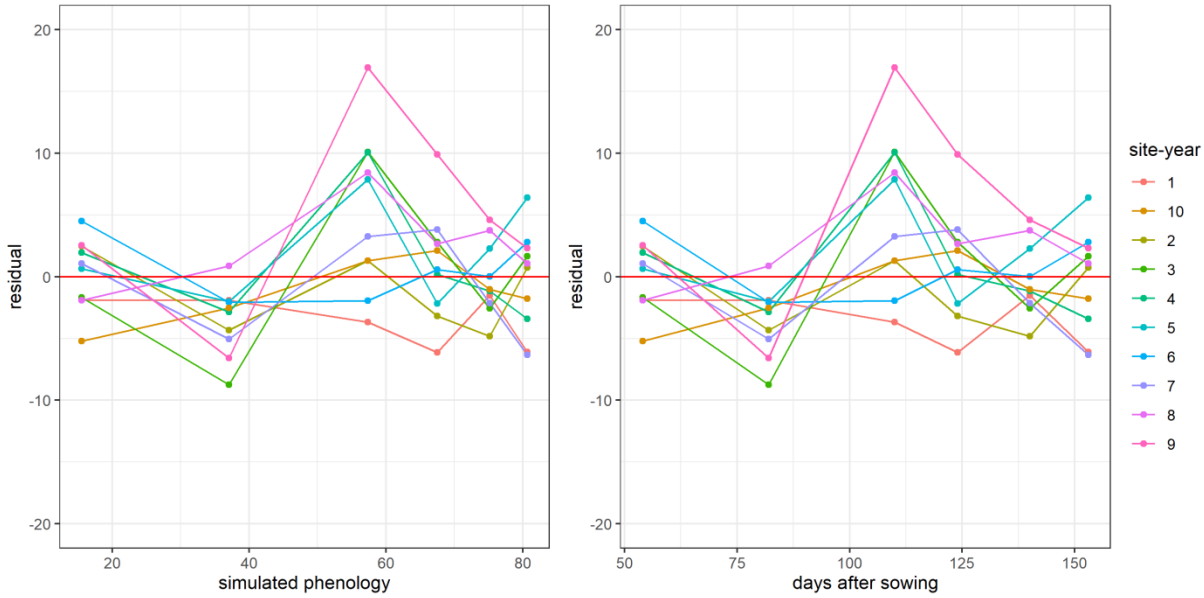

Figure S3: Residuals vs simulated phenology and days after sowing after calibration to 10 site-years in the controlled cultivar-environment synthetic sequence

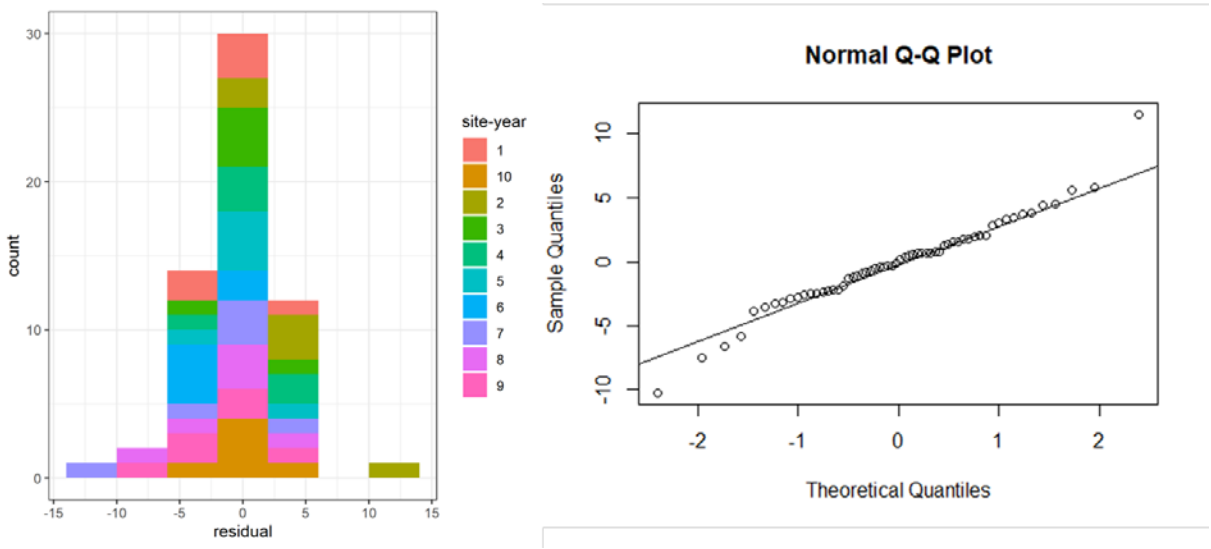

Figure S4: Histogram and quantile-quantile plots of the residuals after calibration to 10 site-years of the ideal synthetic sequence

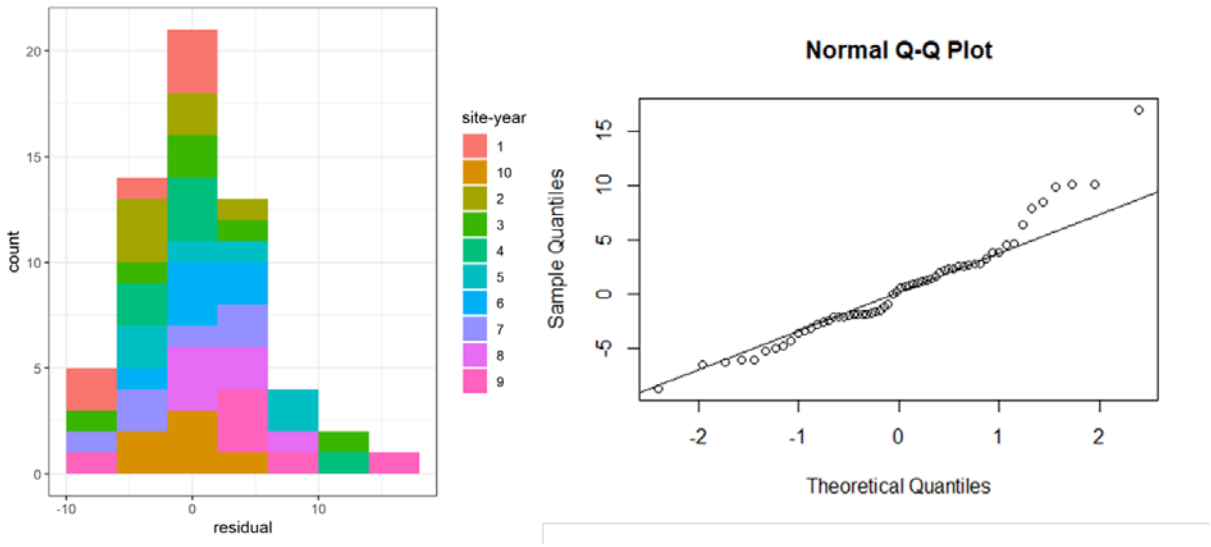

Figure S5: Histogram and quantile-quantile plots of the residuals after calibration to 10 site-years of the controlled cultivar-environment synthetic sequence

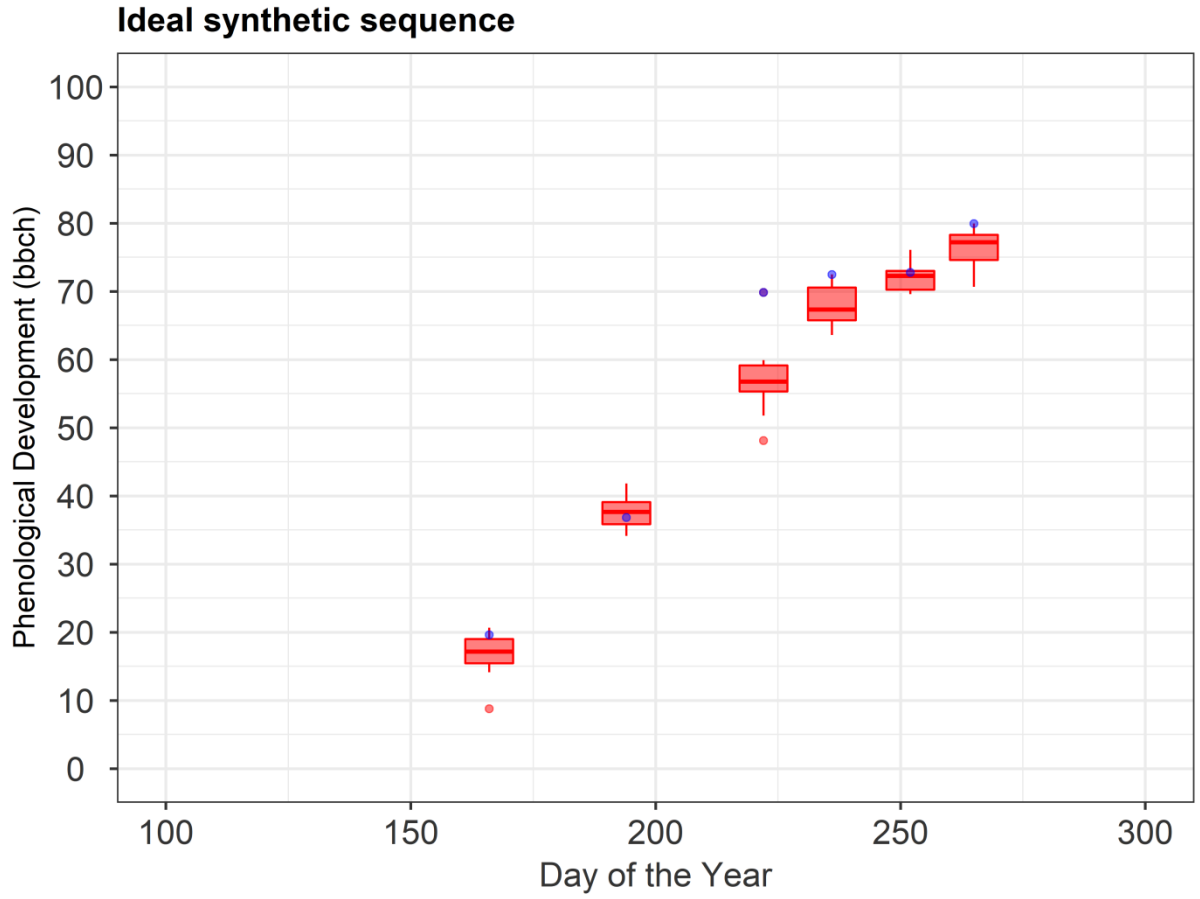

Figure S6: The boxplots show the phenological development (BBCH) of all the site-years used in calibration in the ideal synthetic sequence. The blue point corresponds to the phenological development (BBCH) for site-year 2.

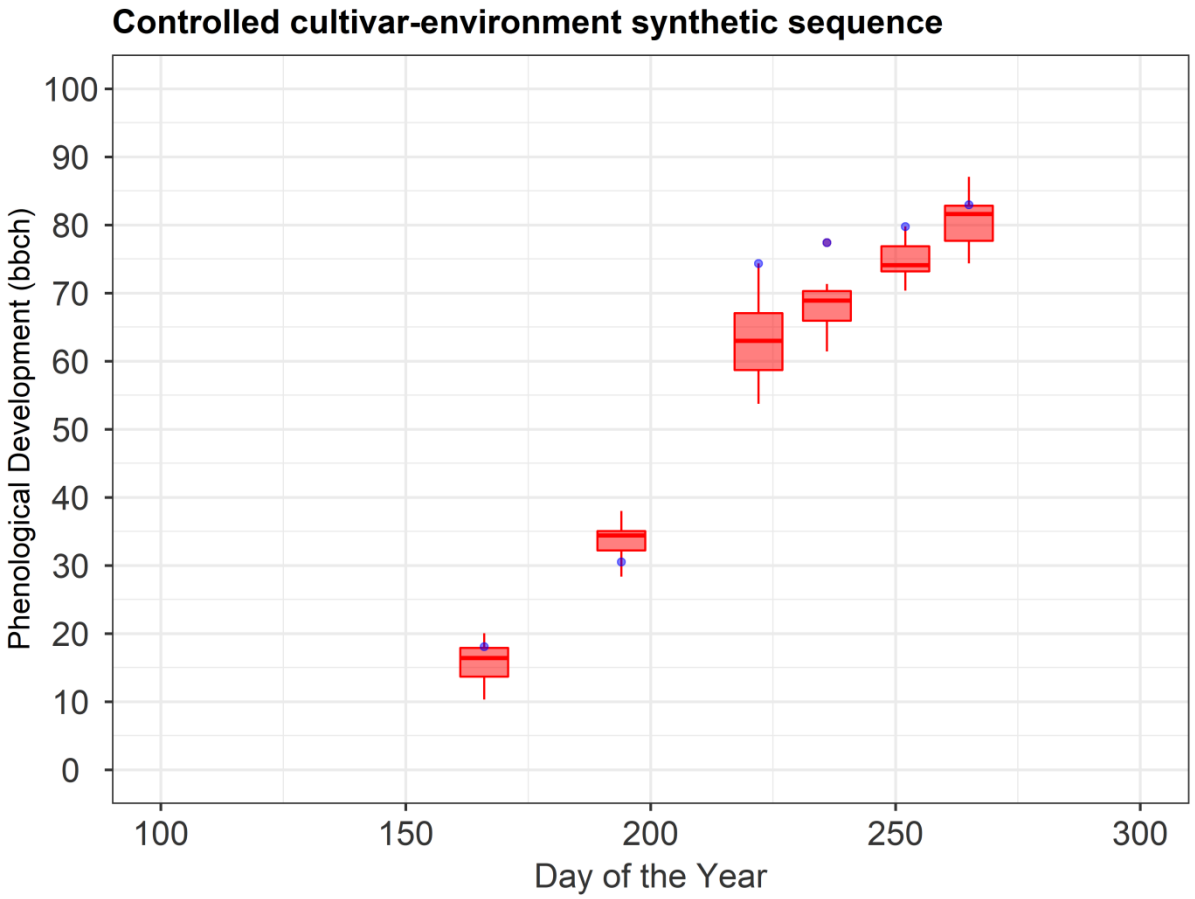

Figure S7: The boxplots show the phenological development (BBCH) of all the site-years used in calibration in the controlled cultivar-environment synthetic sequence. The blue point corresponds to the phenological development (BBCH) for site-year 9.

S3.2 True sequence in Swabian Alb

The residual plots for the sequential updates with greater than 3 calibration site-years show high residuals in the vegetative phase (simulated phenology<61BBCH) (Fig. S11, Fig. S12, Fig. S13). Residuals from site-years 6_2013 and 5_2016 cause this skewness in the distribution of the residuals (Fig. S14, Fig. S15, Fig. S16). This behaviour is attributed to the model's inability to capture the slow development seen in these site-years as evident from the single-site-year calibration results in Fig. S22.

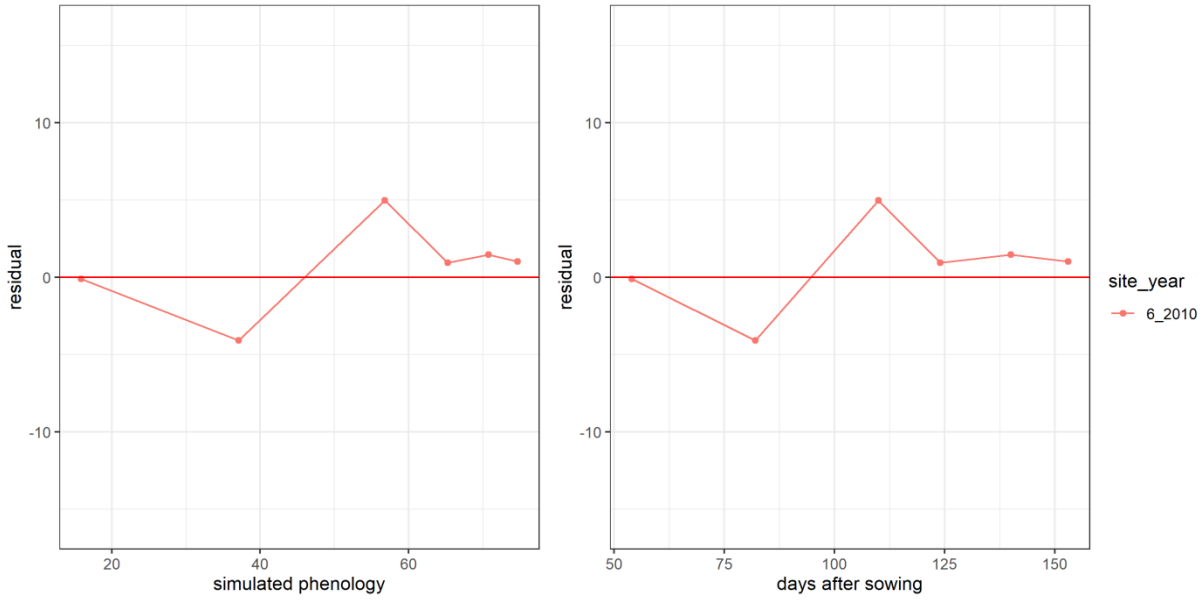

Figure S8: Residuals vs simulated phenology and days after sowing after calibration to site-year 6_2010

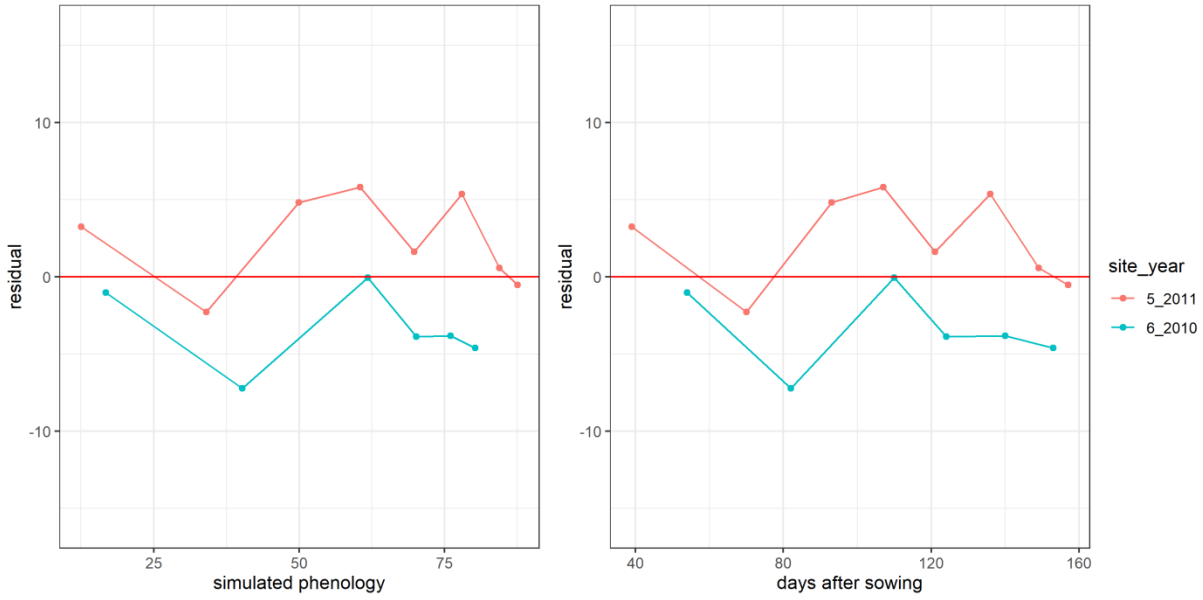

Figure S9: Residuals vs simulated phenology and days after sowing after calibration to site-years 6_2010 and 5_2011

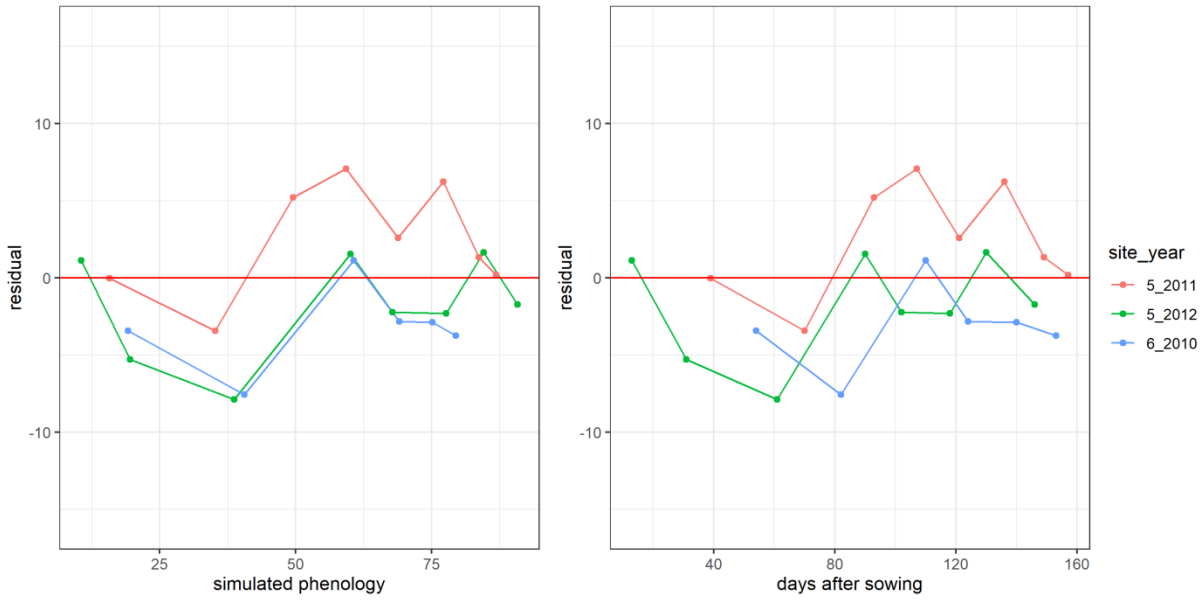

Figure S10: Residuals vs simulated phenology and days after sowing after calibration to site-years 6_2010, 5_2011, and 5_2012

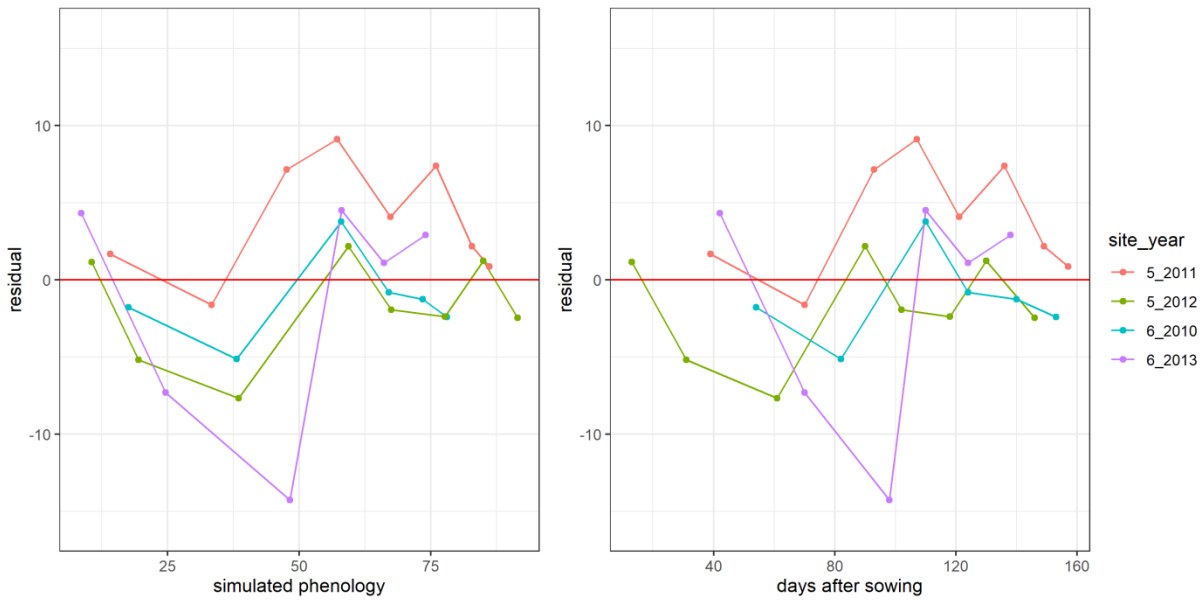

Figure S11: Residuals vs simulated phenology and days after sowing after calibration to site-years 6_2010, 5_2011, 5_2012, and 6_2013

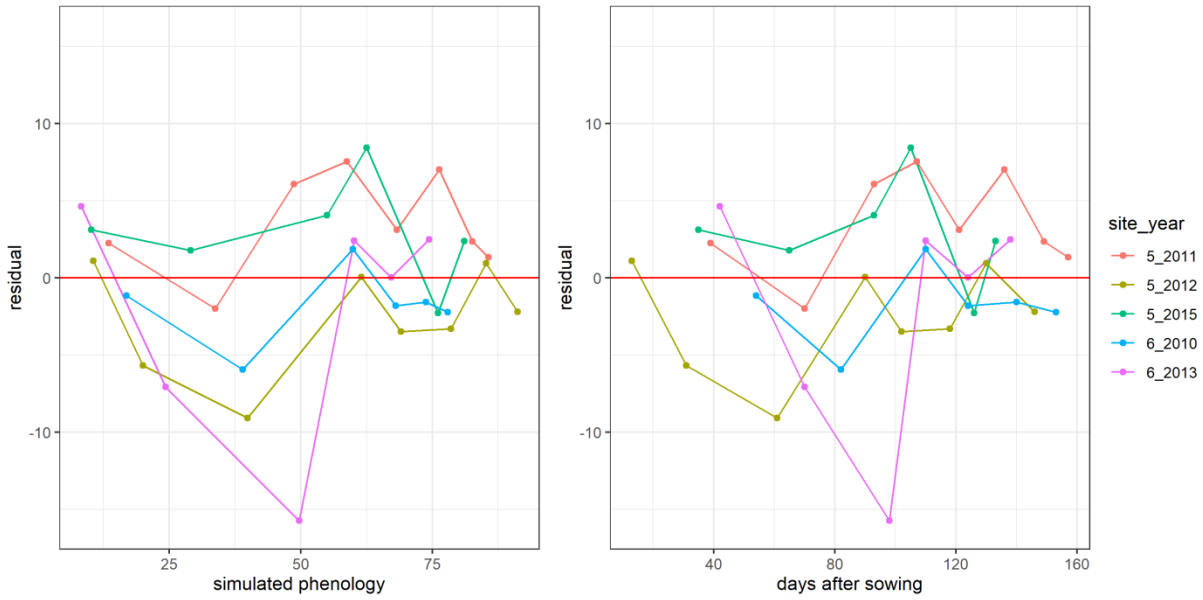

Figure S12: Residuals vs simulated phenology and days after sowing after calibration to site-years 6_2010, 5_2011, 5_2012, 6_2013, and 5_2015

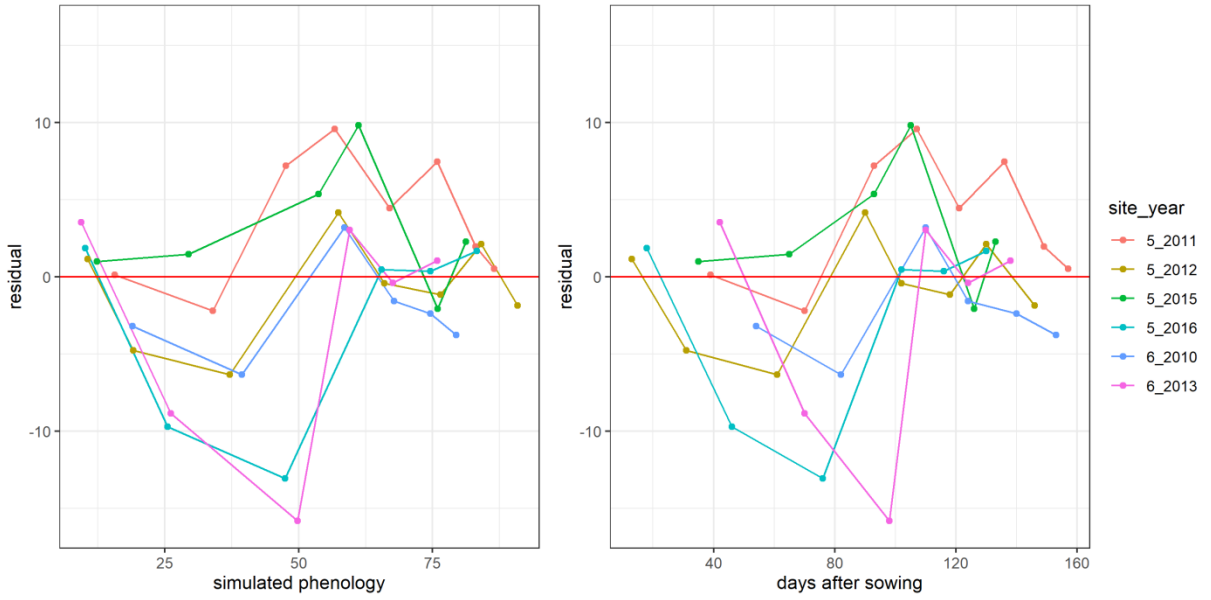

Figure S13: Residuals vs simulated phenology and days after sowing after calibration to site-years 6_2010, 5_2011, 5_2012, 6_2013, 5_2015, and 5_2016

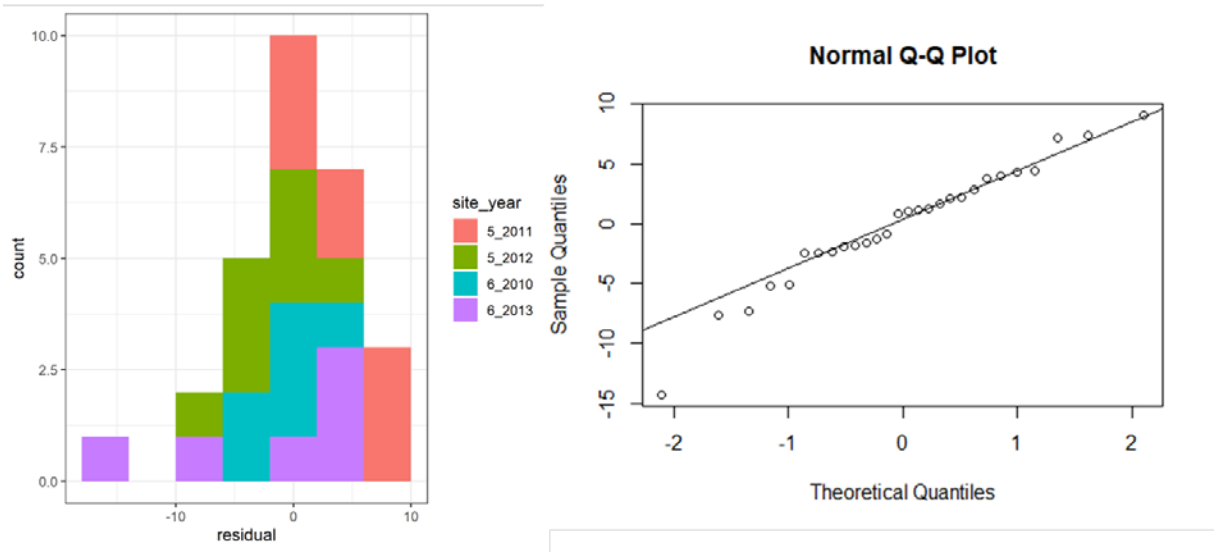

Figure S14: Histogram and quantile-quantile plots of the residuals after calibration to site-years 6_2010, 5_2011, 5_2012, and 6_2013

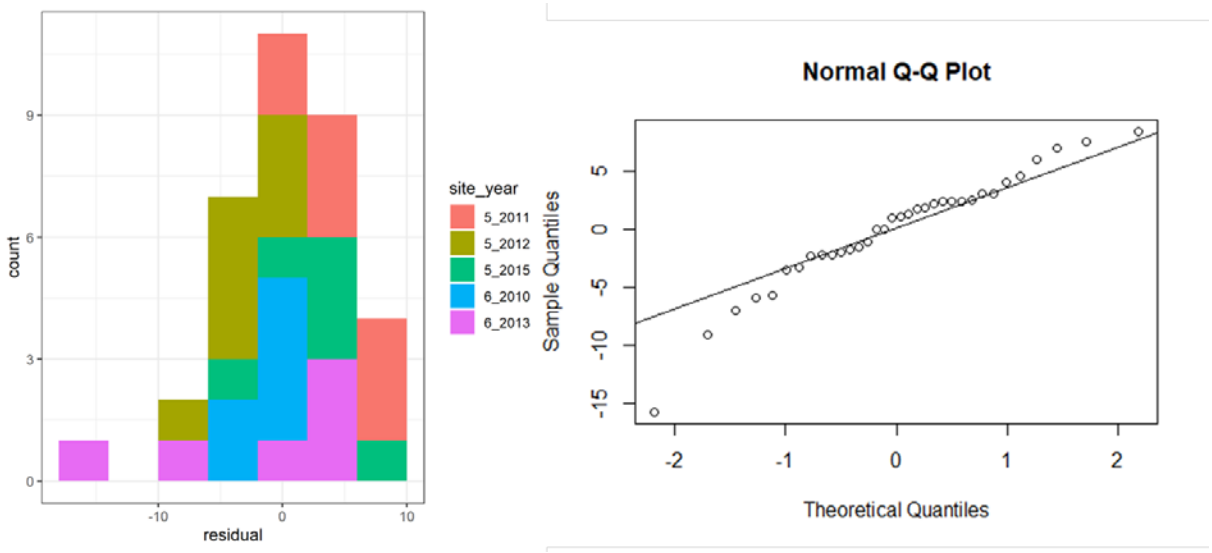

Figure S15: Histogram and quantile-quantile plots of the residuals after calibration to site-years 6_2010, 5_2011, 5_2012, 6_2013, and 5_2015

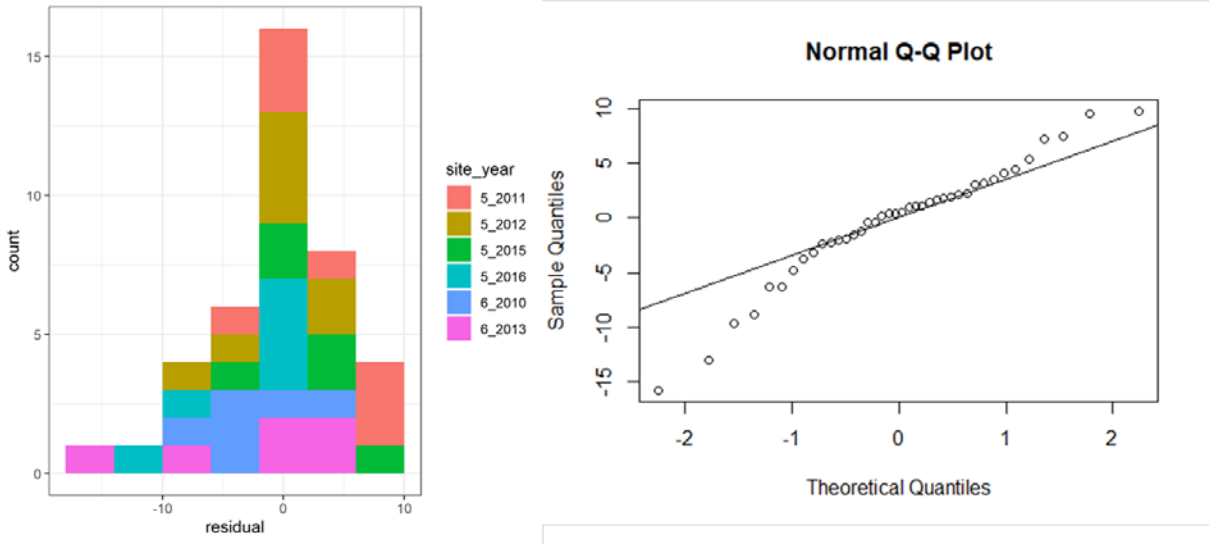

Figure S16: Histogram and quantile-quantile plots of the residuals after calibration to site-years 6_2010, 5_2011, 5_2012, 6_2013, 5_2015, and 5_2016

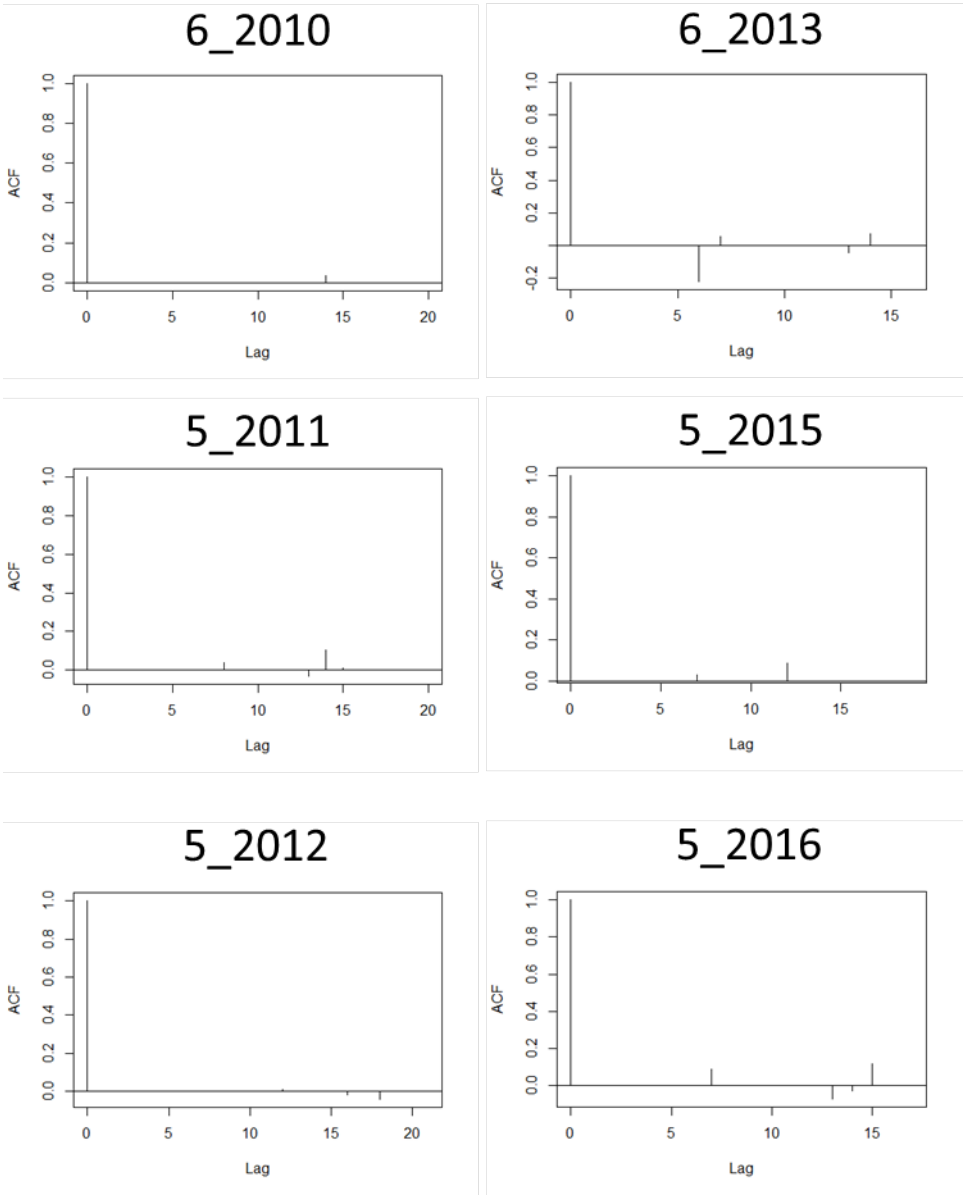

Figure S17: ACF (auto-correlation function) plots of the residuals after calibration to site-years 6_2010, 5_2011, 5_2012, 6_2013, 5_2015, and 5_2016

S3.3 True sequence in Kraichgau

The residual plots for Kraichgau with limited observations show no evidence of heteroscedasticity (Fig. S18, Fig. S19, Fig. S20) and a nearly normal distribution (Fig. S21).

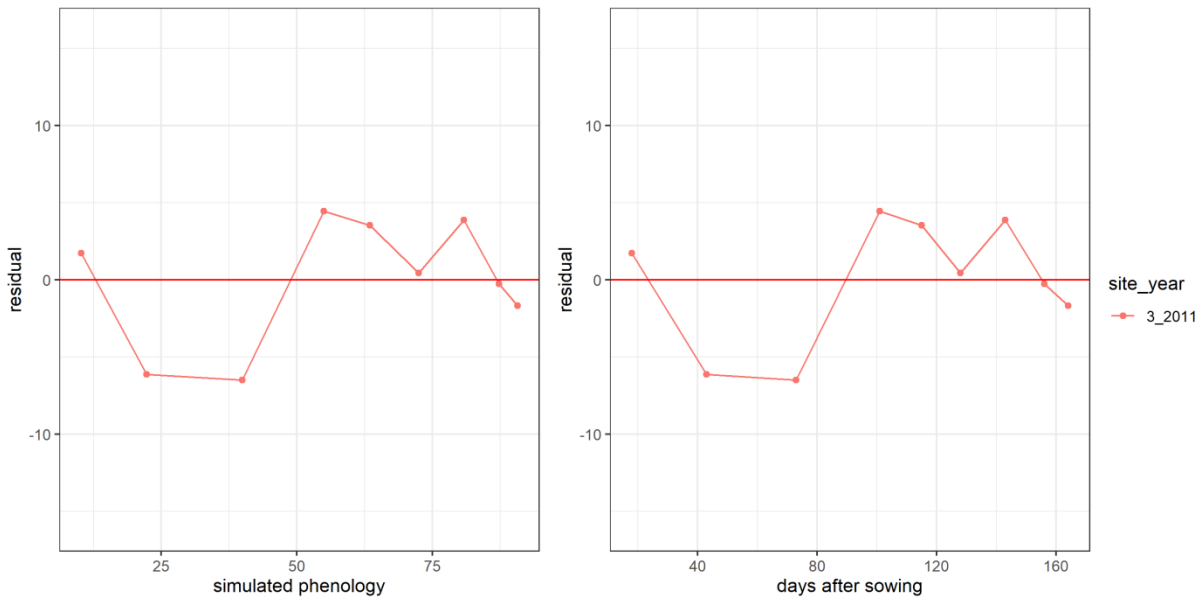

Figure S18: Residuals vs simulated phenology and days after sowing after calibration to site-years 3_2011

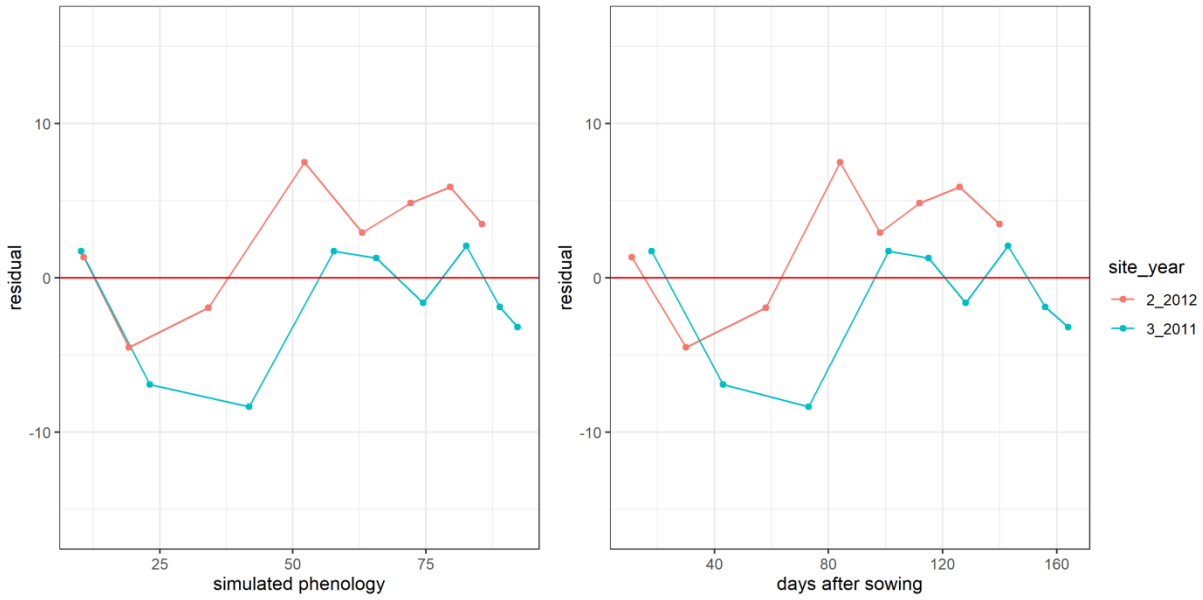

Figure S19: Residuals vs simulated phenology and days after sowing after calibration to site-years 3_2011 and 2_2012

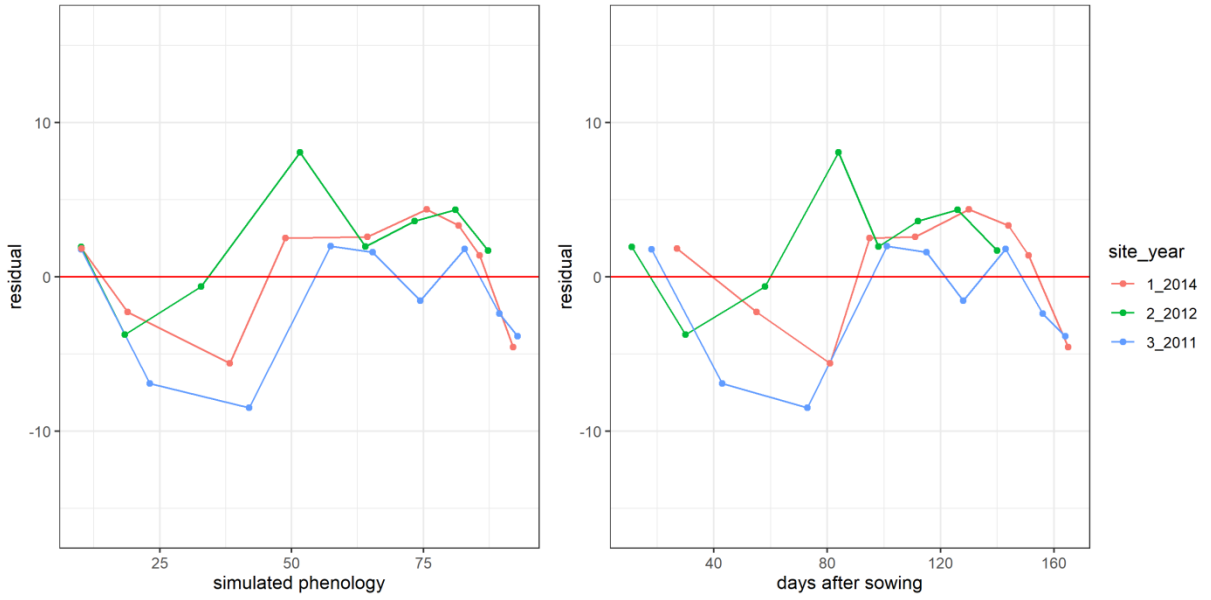

Figure S20: Residuals vs simulated phenology and days after sowing after calibration to site-years 3_2011, 2_2012, and 1_2014

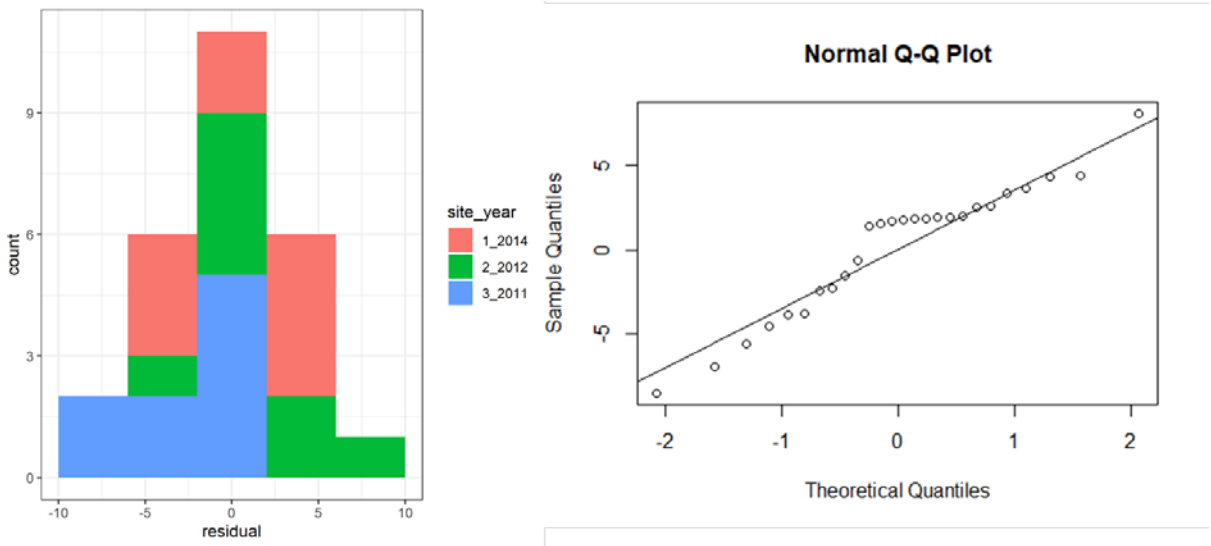

Figure S21: Histogram and quantile-quantile plots of the residuals after calibration to site-years 3_2011, 2_2012, and 1_2014

S4. Single-site-year calibration results

Observed and simulated phenology, after the SPASS model was calibrated individually to the site-years in the study, are plotted in Fig. S22.

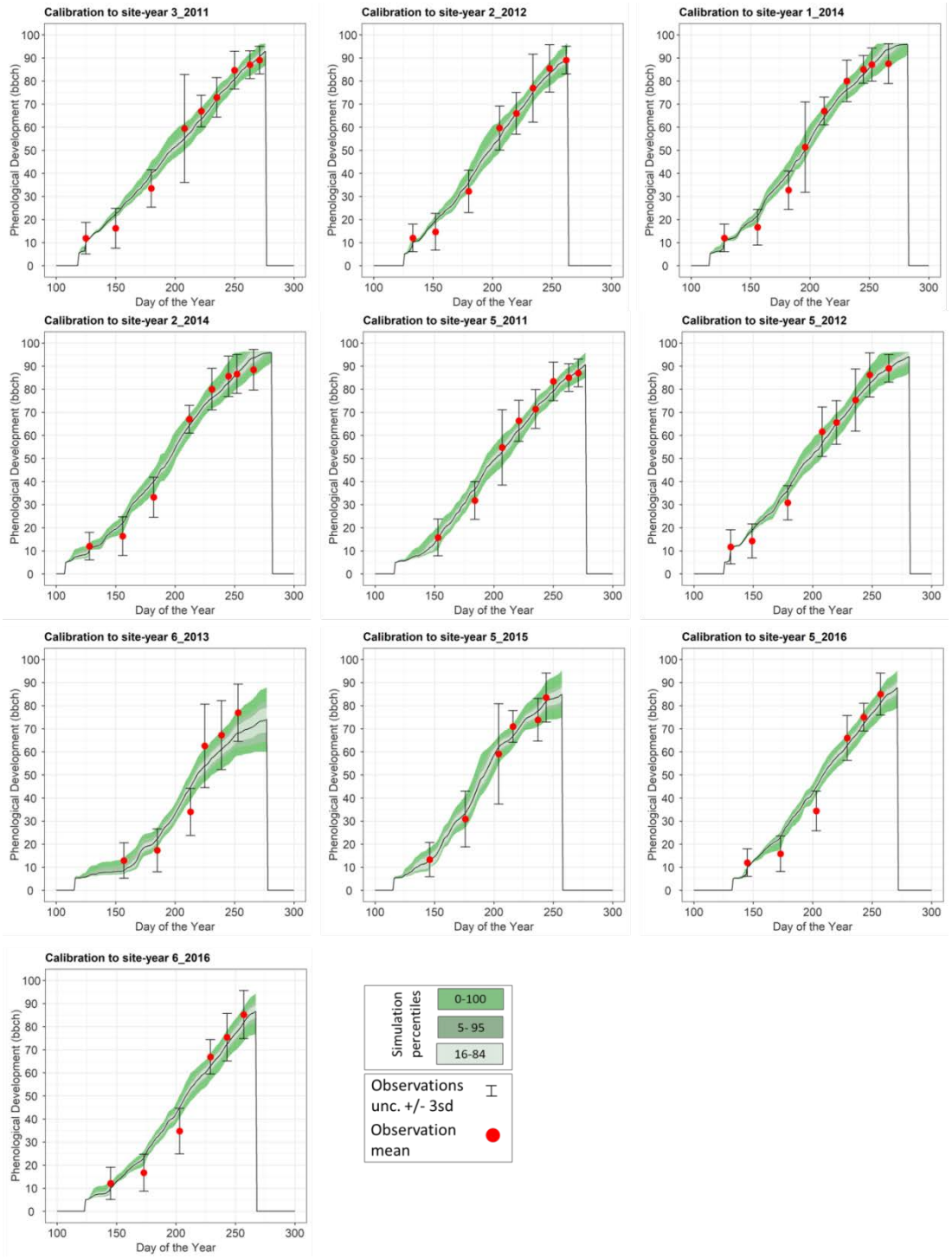

Figure S22: Observed and simulated phenological development after calibration, plotted against the day of the year. The red points are the mean observations, while the black error bars indicate +/- 3 standard deviations. The mean simulation is indicated by the continuous black line. The green bands represent the different percentiles of simulated phenology. It is

noted that for some site-years, the calibrated model is unable to capture the slow development rate during the vegetative phase.

S5. Parameter distributions and entropy: synthetic sequences

Marginal prior and posterior distributions for the 6 estimated parameters of the SPASS phenology model and the entropy estimates are plotted for the ideal (Fig. S23) and controlled cultivar-environment (Fig. S24) synthetic sequences.

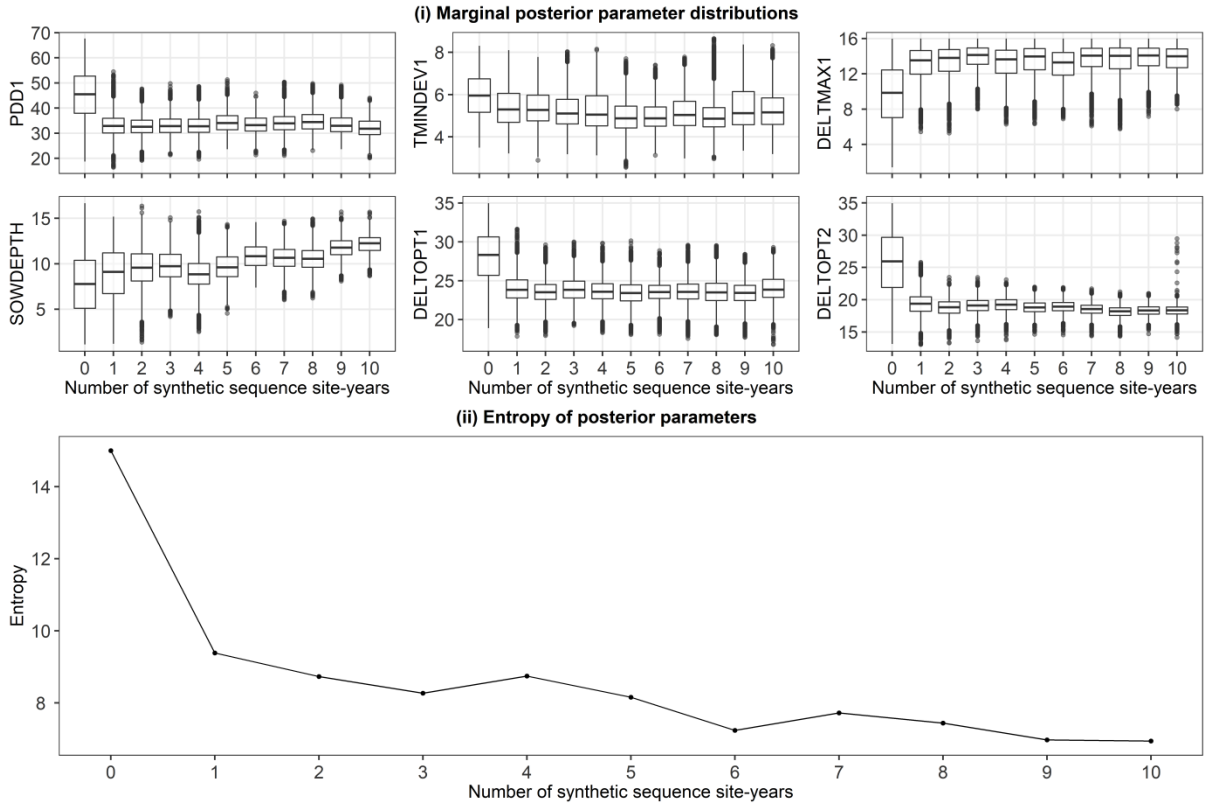

Figure S23: (i) Marginal prior and posterior parameter distributions of the 6 estimated parameters after BSU in the ideal synthetic sequence. Marginal posterior parameter values (y-axis) is plotted against the number of site-years used for calibration (x-axis), starting with the initial prior (0 on x-axis). (ii) Information entropy of the posterior parameter distributions after BSU was applied to the synthetic sequence. Length of the box represents the inter-quartile range (IQR), whiskers extend from the boxes up to 1.5 × IQR and values beyond this range are plotted as points. The ranges for parameters SOWDEPTH and DELTOPT2 narrowed through the sequential updates while the remaining parameters do not show a noticeable narrowing in range.

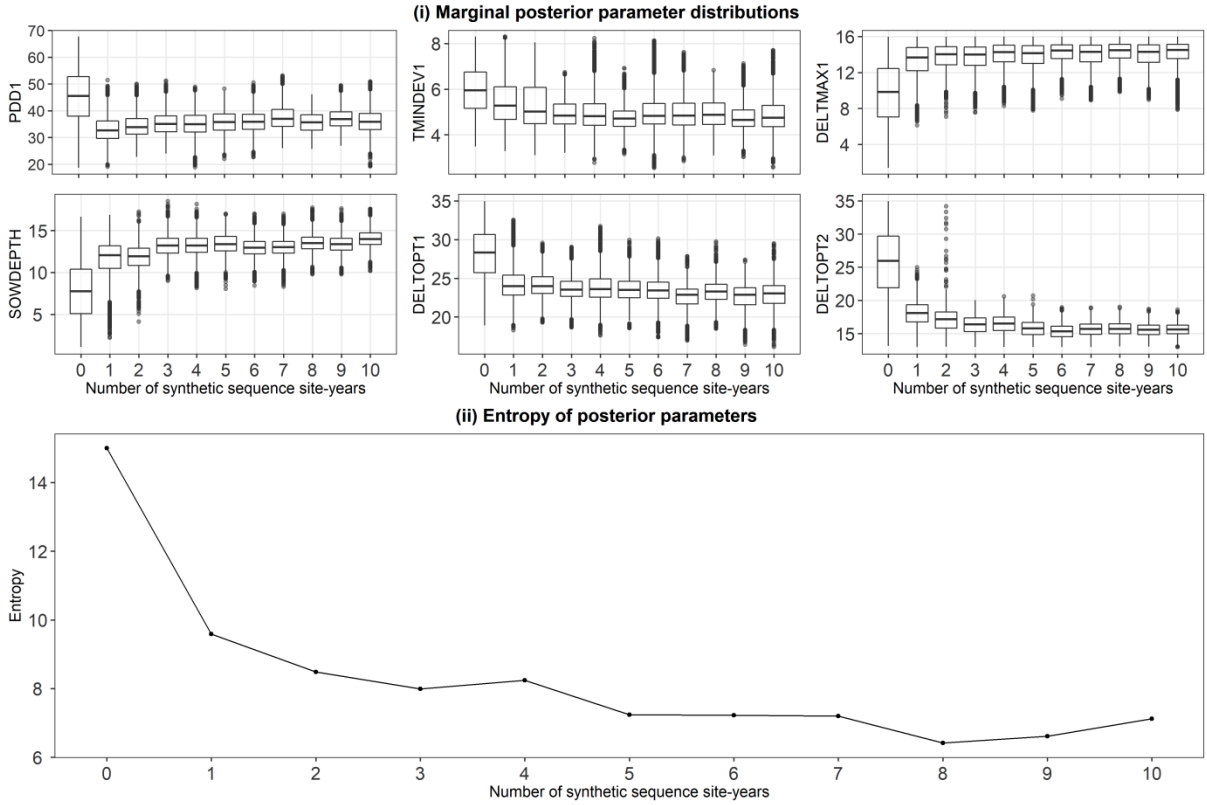

Figure S24: (i) Marginal prior and posterior parameter distributions of the 6 estimated parameters after BSU in the controlled cultivar-environment synthetic sequence. Marginal posterior parameter values (y-axis) is plotted against the number of site-years used for calibration (x-axis), starting with the initial prior (0 on x-axis). (ii) Information entropy of the posterior parameter distributions after BSU was applied to the synthetic sequence. Length of the box represents the inter-quartile range (IQR), whiskers extend from the boxes up to 1.5 × IQR and values beyond this range are plotted as points. The ranges for parameters SOWDEPTH and DELTOPT2 narrowed through the sequential updates while the remaining parameters do not show a noticeable narrowing in range.

S6. References

Beirlant, J., Dudewicz, E., Györfi, L., Dénes, I., 1997. Nonparametric entropy estimation. An overview. Int. J. Math. Stat. Sci. 6, 17–39.

Bertrand Iooss, Veiga, S. Da, Weber, A.J., Pujol, G., 2020. sensitivity: Global Sensitivity Analysis of Model Output.

Cuntz, M., Mai, J., Zink, M., Thober, S., Kumar, R., Schäfer, D., Schrön, M., Craven, J., Rakovec, O., Spieler, D., Prykhodko, V., Dalmasso, G., Musuuza, J., Langenberg, B., Attinger, S., Samaniego, L., 2015. Computationally inexpensive identification of noninformative model parameters by sequential screening. Water Resour. Res. 51, 6417–6441. https://doi.org/10.1002/2015WR016907

Duong, T., 2020. ks: Kernel Smoothing.

Morris, M.D., 1991. Factorial Sampling Plans for Preliminary Computational Experiments. Technometrics 33, 161–174.