Title: Simultaneous calibration of grapevine phenology and yield with a soil-plant-atmosphere system model using the frequentist method

Journal Name: Agronomy

Authors: Chenyao Yang*, Christoph Menz, Helder Fraga, Samuel Reis, Nelson Machado, Aureliano C. Malheiro and João A. Santos

Corresponding Author: Chenyao Yang

Affiliation of corresponding author: Centre for the Research and Technology of Agro-Environmental and Biological Sciences (CITAB)/ Inov4Agro (Institute for Innovation, Capacity Building and Sustainability of Agri-Food Production), University of Trás-os-Montes and Alto Douro (UTAD), 5000-801, Vila Real, Portugal

Email: chenyao_yang@outlook.com

Supplementary Material

Supplementary information on processing weather and soil data into required model inputs:

The daily mean surface (2-m) temperature (°C) and precipitation sum (mm) were directly obtained using the recent release of E-OBS gridded dataset (v21.0 e) at an enhanced spatial resolution (0.1° × 0.1°) from 1950 to 2019 (Cornes et al., 2018). Daily mean solar radiation (MJ m^{-2} day^{-1}), wind speed (m/s) and actual vapour pressure (mbar) were required by the model, which were retrieved and processed from the new ERA5-Land hourly reanalysis dataset at a 0.1° × 0.1° horizontal resolution covering a period from 1981 to present (Hersbach et al., 2020). The hourly 10-m wind components data (u, v) from ERA5-Land was firstly used to obtain the hourly wind speed using Pythagorean theorem, before calculating the daily mean and then being adjusted to the standard height of 2 m following the logarithmic wind speed profile (Allen et al., 1998). Daily mean vapour pressure was estimated as the mean between minimum and maximum vapour pressure, which were respectively calculated using the conventional exponential function of daily minimum and maximum dew point temperature (Allen et al., 1998). The daily minimum and maximum dew point temperatures were both derived from the 2-m dewpoint temperature dataset from ERA5-Land. For soil inputs, the data source and associated horizontal resolution (~0.01°), as well as estimation methods for required soil parameters (e.g. soil texture, surface dry albedo, soil volumetric moisture at field capacity and wilting point), were the same from our previous modelling works, which had already been described in details in Yang et al., (2020). These four vineyard plots indeed featured the same soil type of Dystric Regosol with a loam-dominant texture, which represented about 63% of soil types in DDR (Fraga et al., 2017).

References:

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration: Guidelines for computing crop requirements. Irrig. Drain. Pap. No. 56, FAO. https://doi.org/10.1016/j.eja.2010.12.001

Cornes, R.C., van der Schrier, G., van den Besselaar, E.J.M., Jones, P.D., 2018. An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets. J. Geophys. Res. Atmos. 123, 9391–9409. https://doi.org/https://doi.org/10.1029/2017JD028200

Fraga, H., Costa, R., Santos, J.A., 2017. Multivariate clustering of viticultural terroirs in the Douro winemaking region. Cienc. e Tec. Vitivinic. https://doi.org/10.1051/ctv/20173202142

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J.N., 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. https://doi.org/10.1002/qj.3803

Yang, C., Fraga, H., van Ieperen, W., Santos, J.A., 2020. Assessing the impacts of recent-past climatic constraints on potential wheat yield and adaptation options under Mediterranean climate in southern Portugal. Agric. Syst. https://doi.org/10.1016/j.agsy.2020.102844

Supplementary information on description of STICS growth and yield formation modules:

The phenology-driven leaf area growth, integrating the abiotic stress and source/sink (leaf *vs* fruit) competition (trophic competition), defined the canopy geometry (width, height). The canopy geometry affected canopy light interception, which was estimated using a radiation transfer method to calculate the distribution of radiation fluxes between rows. The intercepted photosynthetically active radiation was then converted into the above-ground dry matter, following the radiation use efficiency approach, also taking into account the effects of atmospheric CO_2 concentration. Subsequently, the yield formation was essentially a process of dry matter allocation to fruits, from fruit-setting until maturity. A boxcartrain technique was used to account for the asynchronous nature of berry maturation by using the fruit age class concept, i.e. number of growth compartments through which berries went. The time spent on each compartment depended on the temperature, corresponding to increasing physiological ages. In each compartment, fruit number, fruit growth (elementary fruit sink strength) and source/sink competition, as well as possibly the thermal stress effects, were calculated, which determined the fruit allocation coefficient from accumulated dry matter. Fruit number depended on planting density, harvestable cluster number/vine (affected by the training system and viticulture practices), potential fruit number formation per cluster (temperature-driven), as well as the frost and trophic stress effects that possibly limited the fruit number formation. The fruit sink strength function was derived from the potential fruit growth function, which combined the berry cell division (exponential curve) and expansion (logistic curve) process, and taking into account the genetic potential of berry growth.

Table S1. Summary of yearly measured cluster number per vine and individual cluster weight at harvest in four experimental vineyard plots over 2014–2019. The indicated figures and the figures in the brackets are the median and the standard deviation (SD) of measurements over 20 random replicates (vines) in each plot, respectively. The representative values are determined by using additional plot measurements on top of current plot data. The representative value of median harvestable cluster number (bold) have been prescribed and fixed in STICS model (i.e. nbinflo parameter) in all subsequent simulation tasks to represent the general situation under a given variety-training system combination.

Years and representative values	PlotS		PlotD		PlotO		PlotM	
	Touriga Nacional with Single-Cordon		Touriga Nacional with Double-Cordon		Touriga Franca with Single-Cordon		Touriga Franca with Double-Cordon	
	Cluster number	Cluster weight (kg)	Cluster number	Cluster weight (kg)	Cluster number	Cluster weight (kg)	Cluster number	Cluster weight (kg)
2014	16 (6)	0.088 (0.041)	21 (6)	0.123 (0.040)	7 (2)	0.184 (0.056)	7 (4)	0.150 (0.069)
2015	8 (6)	0.100 (0.056)	21 (7)	0.123 (0.116)	7 (2)	0.229 (0.084)	10 (4)	0.165 (0.043)
2016	14 (3)	0.139 (0.033)	17 (6)	0.133 (0.037)	7 (3)	0.225 (0.060)	9 (5)	0.238 (0.050)
2017	9 (4)	0.088 (0.024)	19 (8)	0.110 (0.024)	10 (4)	0.160 (0.057)	9 (7)	0.186 (0.059)
2018	10 (5)	0.112 (0.047)	15 (7)	0.067 (0.022)	10 (6)	0.131 (0.046)	10 (4)	0.219 (0.051)
2019	12 (5)	0.126 (0.037)	22 (5)	0.092 (0.026)	9 (3)	0.152 (0.052)	13 (4)	0.153 (0.041)
Representative values	**13**	/	**18**	/	**6**	/	**10**	/
Measurement size for determining representative values (including the 20 replicates)	780 measurements from 3 plots in (average) 7 years		780 measurements from 3 plots in (average) 7 years		360 measurements from 2 plots in (average) 5 years		860 measurements from 4 plots in (average) 5 years	

Table S2. List of parameters that respectively minimize the sum of nRMSE (%) based on the multivariate objective function (combine phenology and yield), the nRMSE for flowering and harvest date, yield separately (univariate function). Since RG is the only parameter that affects flowering date, possible difference of values that minimize nRMSE between flowering and harvest date is indicated. These parameter values are obtained based on the same calibration procedures (see section 2.3 for details) using additional evaluation data at the experimental vineyard (Lat: 41.15°N, Lon: –7.75°W) in DDR over 2012–2014 for Touriga Nacional (double-cordon) and over 2012–2013 for Touriga Franca (double-cordon) (Fraga et al., 2015). Parameters that minimize the multivariate objective function with identical values in both calibration and evaluation datasets are highlighted with bold. Detailed parameter description is available in Table 2.

Grapevine parameter abbreviation	Touriga Nacional with double-cordon			Touriga Franca with double-cordon		
	Minimize objective function	Minimize nRMSE (%) for phenology	Minimize nRMSE (%) for yield	Minimize objective function	Minimize nRMSE (%) for phenology	Minimize nRMSE (%) for yield
FS	50	350	50	200	275	125
FN	2.0	1.5	2.0	1.0	0.5—2.5	1.0
FF	1100	1500	1100	**1300**	1300	900
FW	**0.5**	0.8	0.5	**0.5**	0.5—1.7	0.8
VG	**600**	1200	600	1000	600	1400
RG	400	400 (both flowering & harvest date)	400	**300**	300 (flower)/ 400 (harvest)	450
WD	**250**	100—250	100—300	250	100—250	100—300
BN	**15**	5	15	**10**	5	10
SS	0.5	0.75	0.5	0.25	0.75	0.25

Table S3. Goodness-of-fit comparison of selected parameter vectors that minimize the sum of nRMSE (%) based on the objective function (combine phenology and yield) with parameter vectors previously selected with trial-and-error approach (Fraga et al., 2015). This is based on the additional evaluation dataset obtained from the experimental vineyard (Lat: 41.15°N, Lon: –7.75°W) in DDR over 2012–2014 for Touriga Nacional (double-cordon) and over 2012–2013 for Touriga Franca (double-cordon).

Evaluation statistics of studied variables		Touriga Nacional with double-cordon		Touriga Franca with double-cordon	
		Minimize objective function	Trial-and-error	Minimize objective function	Trial-and-error
Flowering Date	MBE (days)	–3	–6	3	–6
	MAE (days)	5	6	3	6
	RMSE (days)	6	8	3	6
	nRMSE (%)	4%	6%	2%	4%
	Rsd	2.0	2.1	1.5	1.8
Harvest Date	MBE (days)	–3	1	–3	–1
	MAE (days)	9	1	10	1
	RMSE (days)	11	2	10	1
	nRMSE (%)	4%	1%	4%	0.4%
	Rsd	1.2	0.7	1.2	1.0
Yield	MBE (kg/ha)	491	–1200	–15	–1550
	MAE (kg/ha)	491	1200	15	1550
	RMSE (kg/ha)	493	1257	20	1818
	nRMSE (%)	7%	17%	0.3%	28%
	Rsd	1.0	0.9	1.0	1.9

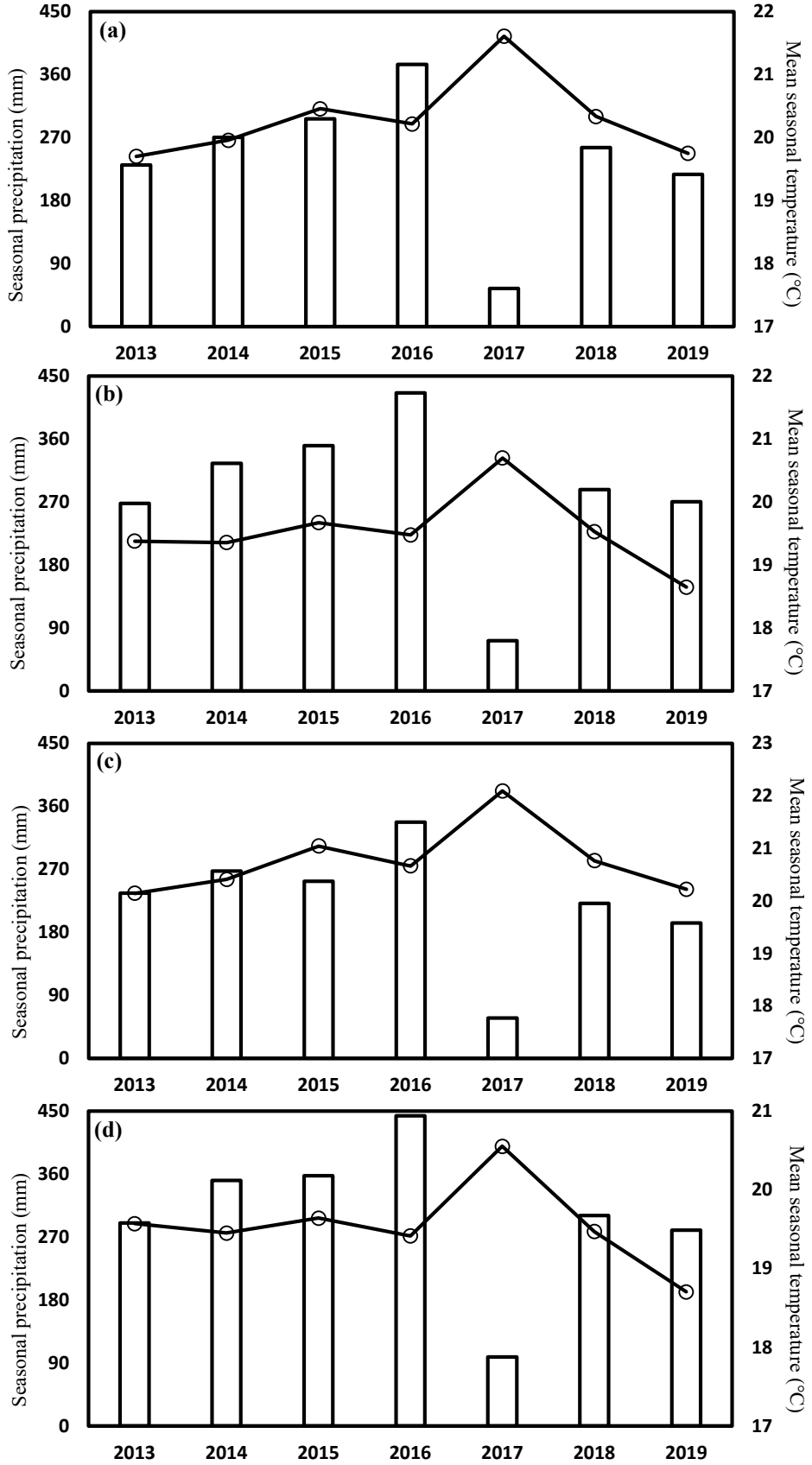

Figure S1. Mean temperature (°C) and precipitation sum (mm) during growing season (between April and October) over 2013—2019 in the studied experimental vineyard plots. Temperature is represented by the line plot with circle markers, while precipitation is indicated by the column plot. (a) Touriga Nacional with single-cordon in plot S; (b) Touriga Nacional with double-cordon in plot D; (c) Touriga Franca with single-cordon in plot O; (d) Touriga Franca with double-cordon in plot M;

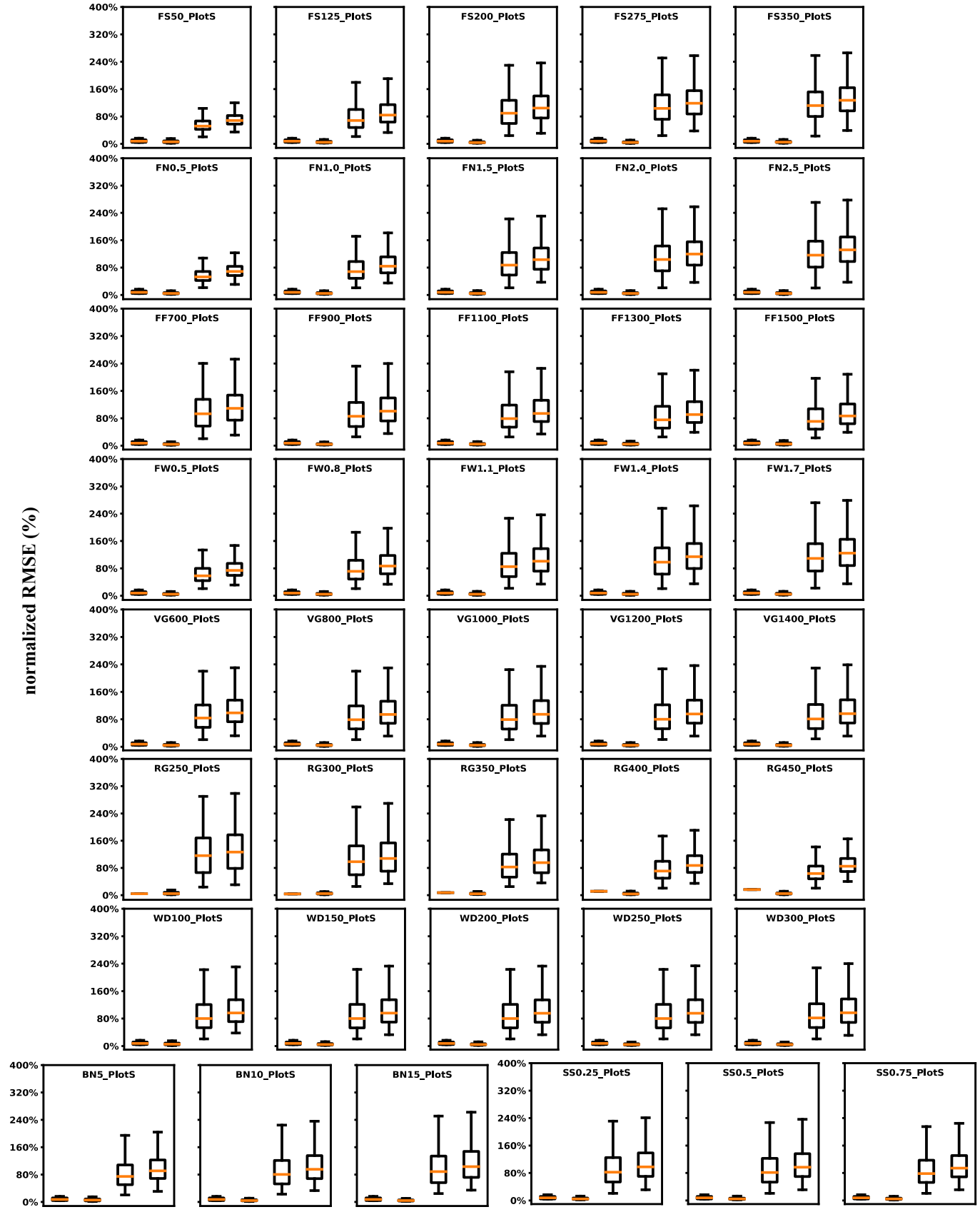

Figure S2. Spread of prediction uncertainties (nRMSE, %) on phenology, yield and the combined variable (i.e. objective function) under individual parameter values (fix a given parameter value and test the remaining combinations) for **Touriga Nacional under single-cordon training system** (observations in plotS over 2014—2019). The parameter abbreviation (see table 2 for detailed descriptions) and corresponding tested values, along with the vineyard plot, are presented as each subplot title. The two generic parameters (only test 3 values) are placed in the bottom. The box from left to right in each subplot indicates flowering date, harvest date, yield and the corresponding combined variable, respectively. The box horizontal lines respectively represent first, second (median) and third quartile, while lower and upper whiskers are positioned according to Tukey's original definition of boxplots (excluding outliers). These definitions were applied for all boxplots throughout.

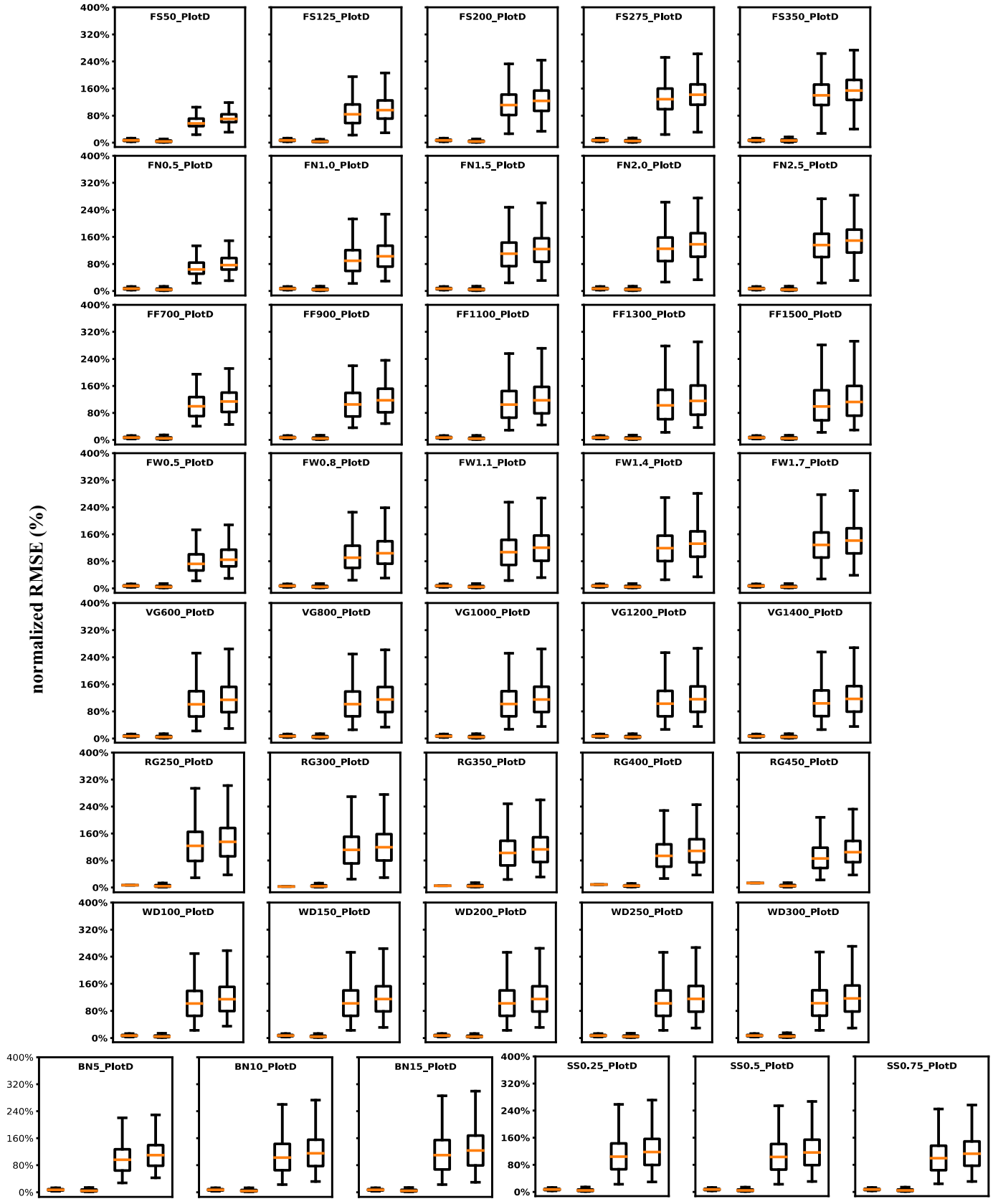

Figure S3. Spread of prediction uncertainties (nRMSE, %) on phenology, yield and the combined variable (i.e. objective function) under individual parameter values (fix a given parameter value and test the remaining combinations) for **Touriga Nacional under double-cordon training system** (observations in plotD over 2014—2019). The box from left to right in each subplot indicates flowering date, harvest date, yield and the corresponding combined variable, respectively.

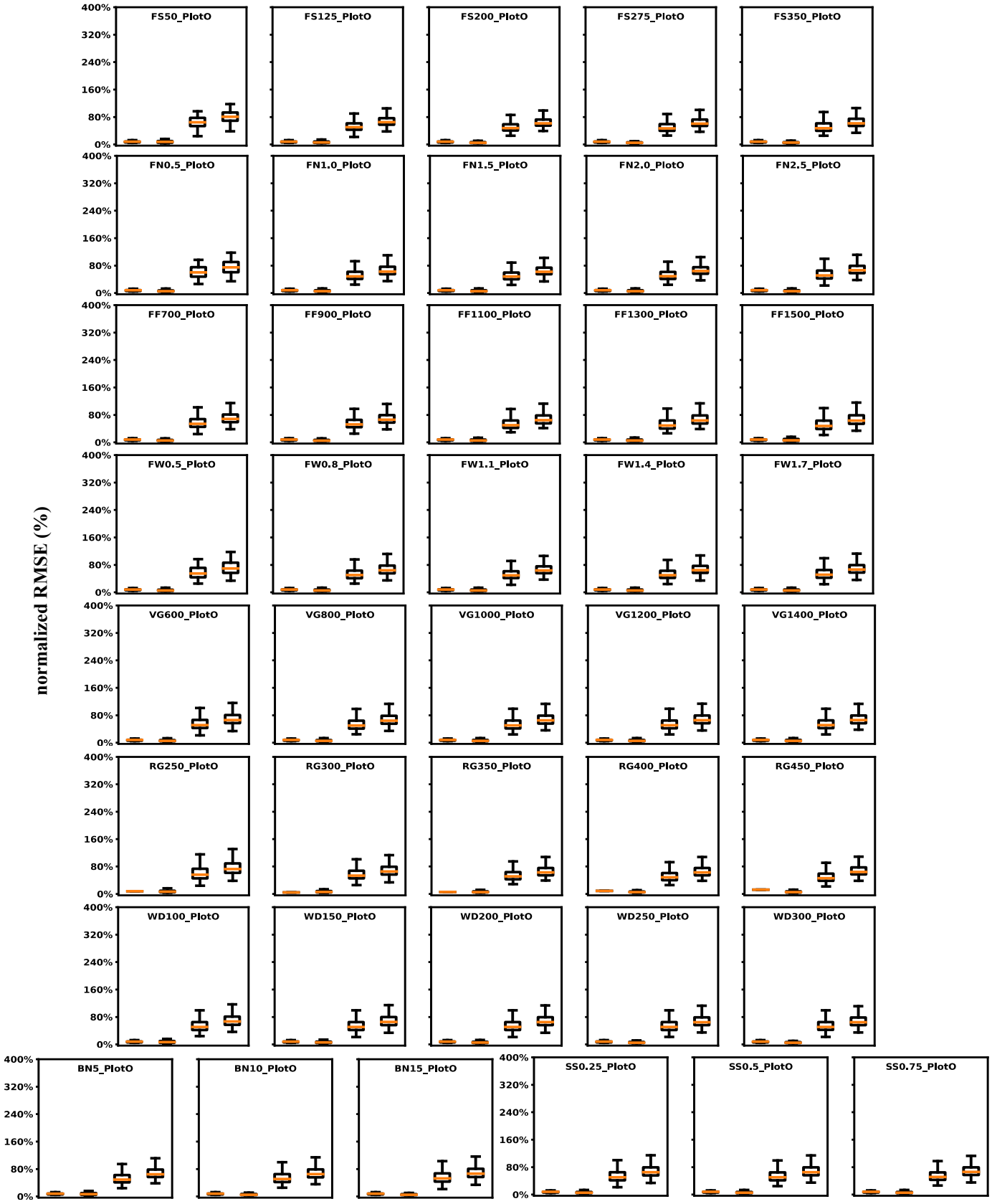

Figure S4. Spread of prediction uncertainties (nRMSE, %) on phenology, yield and the combined variable (i.e. objective function) under individual parameter values (fix a given parameter value and test the remaining combinations) for **Touriga Franca under single-cordon training system** (observations in plotO over 2014—2019). The box from left to right in each subplot indicates flowering date, harvest date, yield and the corresponding combined variable, respectively.

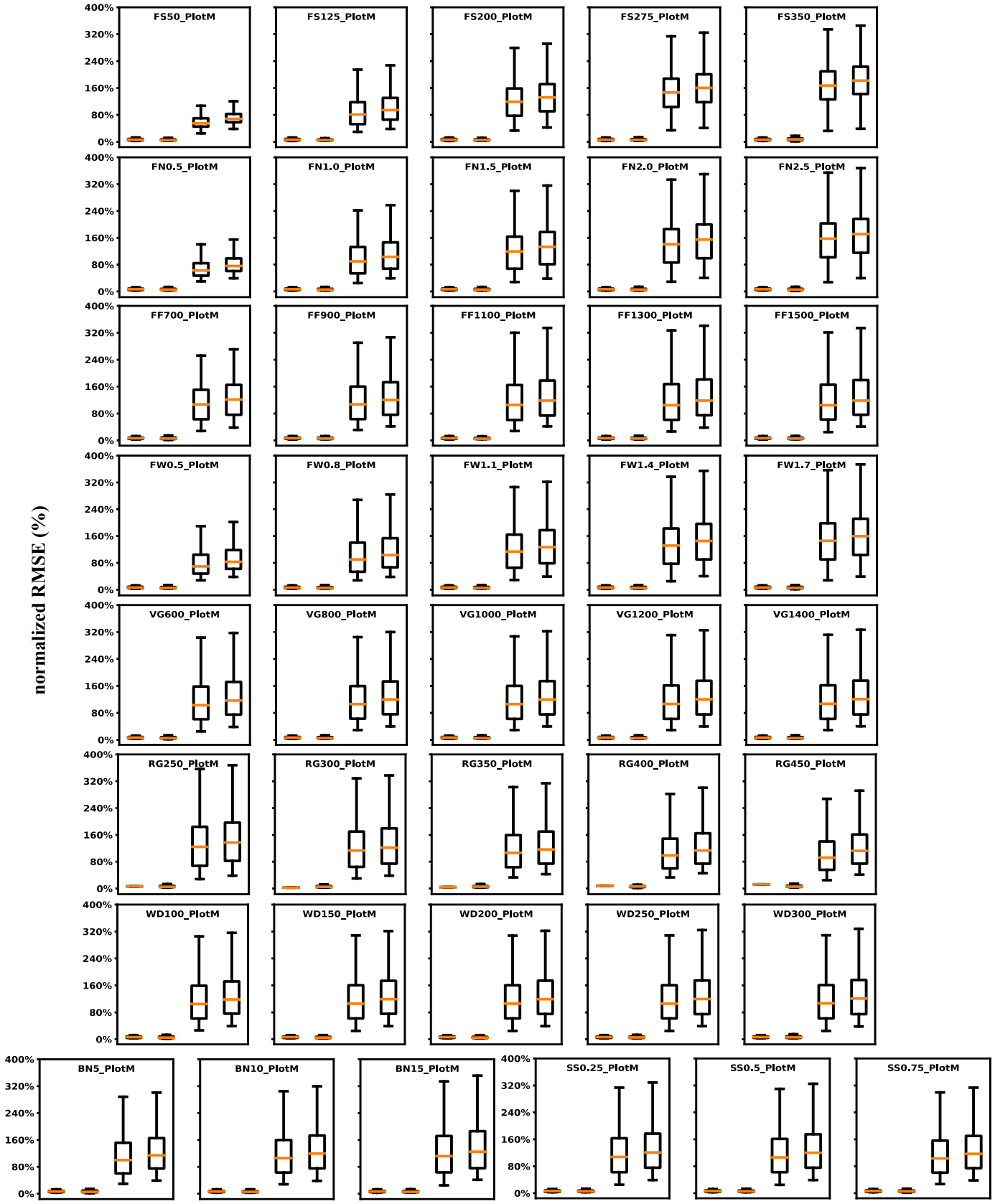

Figure S5. Spread of prediction uncertainties (nRMSE, %) on phenology, yield and the combined variable (i.e. objective function) under individual parameter values (fix a given parameter value and test the remaining combinations) for **Touriga Franca under double-cordon training system** (observations in plotM over 2014—2019). The box from left to right in each subplot indicates flowering date, harvest date, yield and the corresponding combined variable, respectively.

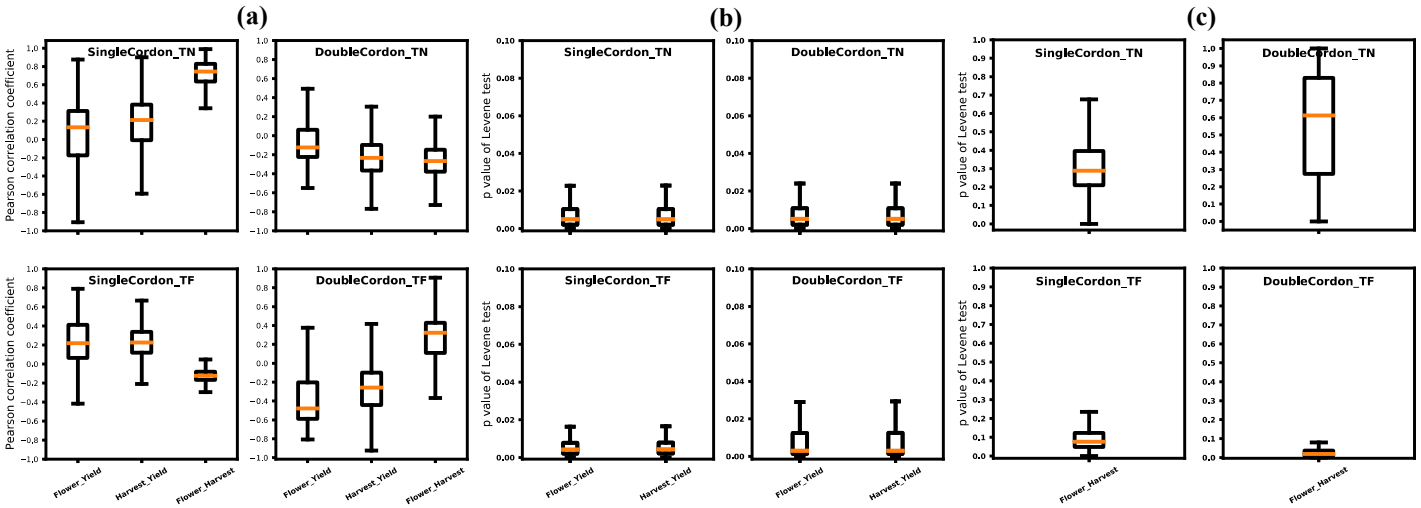

Figure S6. Boxplots of (a) error (observation-simulation) correlations between studied variables (flower and harvest date and yield) among all tested parameter vectors (684,375), (b) *p* values of Levene´s test on homoscedasticity between prediction errors in flowering/harvest date and in yield among all parameter vectors, (c) *p* values of Levene´s test between prediction errors in flowering and harvest date. The underscore appearing in the labels of horizontal axis indicates each paired two variables.

The results for error correlation between different response variables, at the same site in different years, are complicate, generally showing a large spread (e.g. vary from negative to positive) among all tested parameter vectors (**Fig. S6a**). This indicates error correlation is not only affected by multiple measurements from the same field (Wallach et al. 2017), but also depends on the testing model parameters. The median values of error correlation also show a large variability among different variety-training systems (random site effect), in which there are negative to weak positive correlations (Pearson coefficient *r* varies from –0.5 to 0.2) between phenology variables and yield, and negative weak to significant positive correlation between flowering and harvest DOY (*r* varies from –0.3 to 0.8) (**Fig. S6a)**. Single-cordon training system (both varieties) constantly presents a weak positive median correlation (around 0.2) between phenology variables and yield, whereas double-cordon training system (both varieties) only show negative values (*r* varies from –0.5 to –0.2) (**Fig. S6a)**. This may indicate the random site effect of errors mainly come from the uncertainties in the type of training systems rather than the choice of varieties (similar distribution of *r* between phenology variables and yield is found between varieties), which again illustrate the importance of making distinctions between grapevine training systems. For Levene´s test on homoscedasticity, error variance is significantly different ($p<0.05$) between phenology and yield irrespective of testing parameters and variety-training systems (**Fig. S6b)**. This is as expected since inter-annual variability between phenology and yield data is considerably different, as characterized by CV in **Table 1**. However, this is not the case between two phenology variables, showing homogeneity of variance in majority of situations (parameter vectors) despite there is relatively a larger spread for TN than for TF (**Fig. S6c)**.

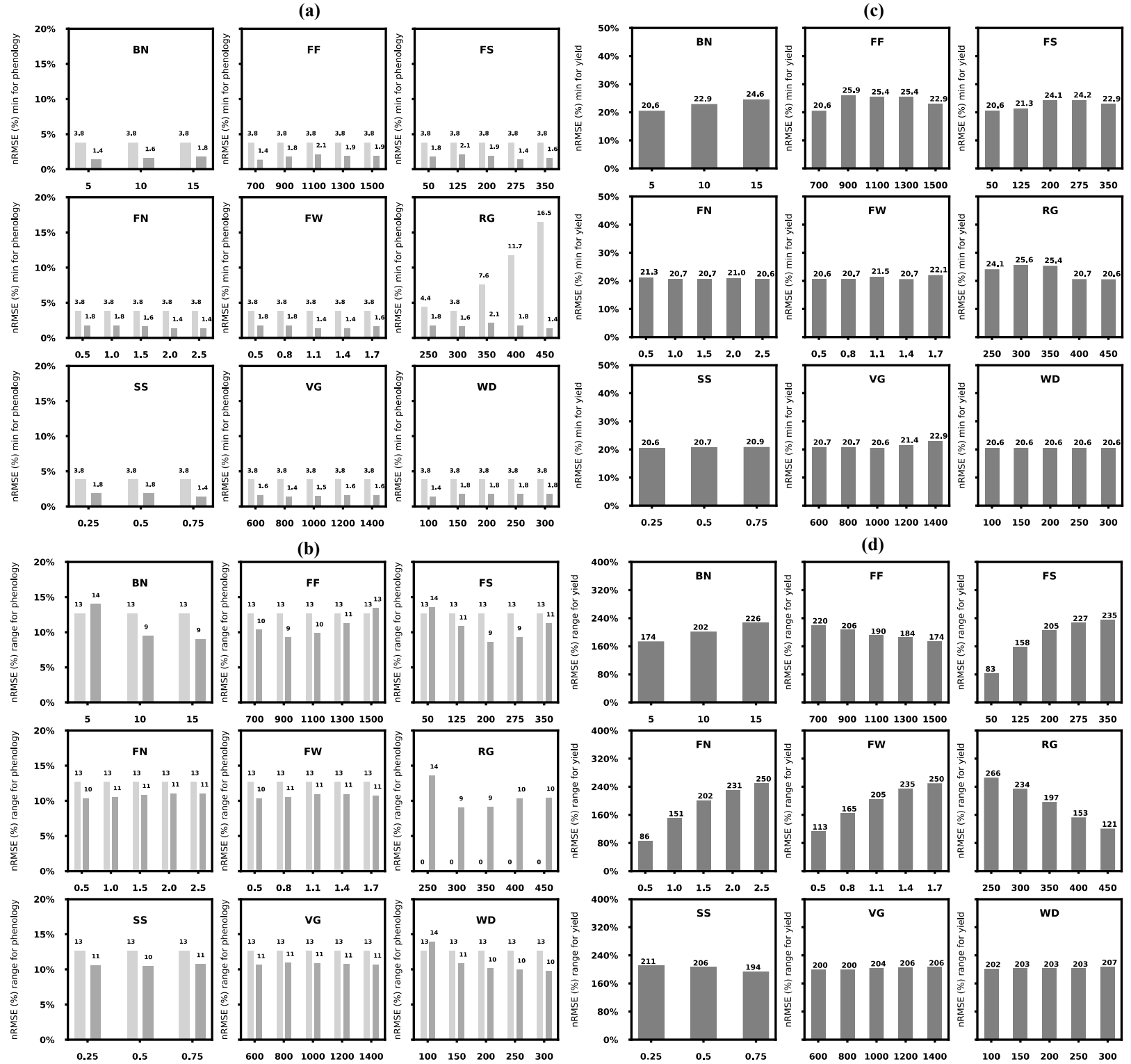

Figure S7. Minimum and range values (spread) of STICS model prediction uncertainties (nRMSE, %) under individual parameter values (fix a given parameter value and test the remaining combinations) on (a—b) phenology (flowering date: light grey bar; harvest date: dark grey bar) and (c—d) yield, respectively. The minimum and range values are respectively withdrawn from the minimum and difference (four times the inter-quartile range) between upper and lower whiskers in boxplots in Figure S2. The analysis concerns the grapevine variety of **Touriga Nacional (TN) at single-cordon** training system (observations in plotS over 2014—2019). The parameter abbreviation has been denoted for each subplot (Refer to table 2 for detailed parameter descriptions).

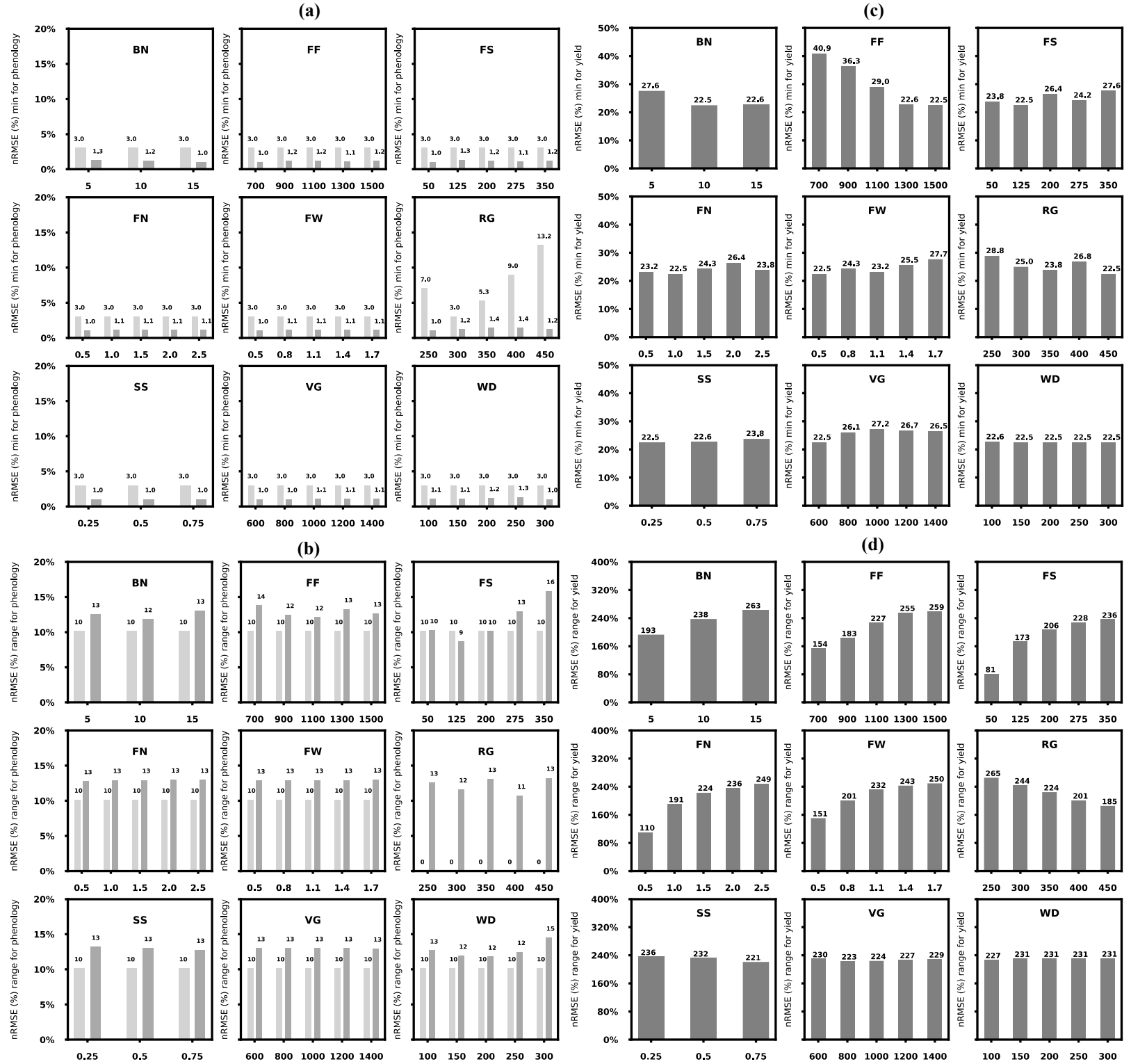

Figure S8. Minimum and range values (spread) of STICS model prediction uncertainties (nRMSE, %) under individual parameter values (fix a given parameter value and test the remaining combinations) on (a—b) phenology (flowering date: light grey bar; harvest date: dark grey bar) and (c—d) yield, respectively. The minimum and range values are respectively withdrawn from the minimum and difference (four times the inter-quartile range) between upper and lower whiskers in boxplots in Figure S3. The analysis concerns the grapevine variety of **Touriga Nacional (TN) at double-cordon** training system (observations in plotD over 2014—2019). The parameter abbreviation has been denoted for each subplot.

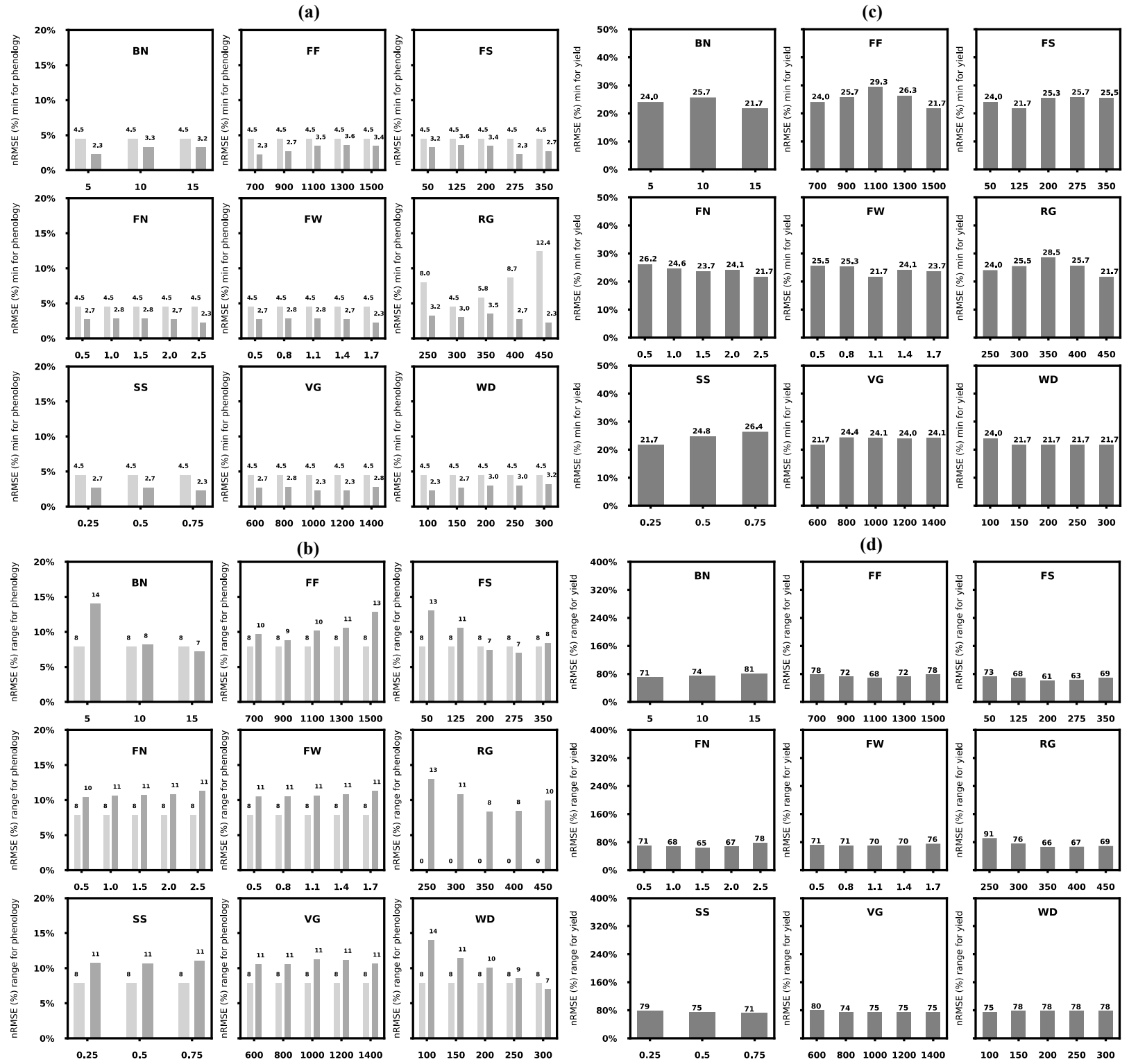

Figure S9. Minimum and range values (spread) of STICS model prediction uncertainties (nRMSE, %) under individual parameter values (fix a given parameter value and test the remaining combinations) on (a—b) phenology (flowering date: light grey bar; harvest date: dark grey bar) and (c—d) yield, respectively. The minimum and range values are respectively withdrawn from the minimum and difference (four times the inter-quartile range) between upper and lower whiskers in boxplots in Figure S4. The analysis concerns the grapevine variety of **Touriga Franca (TF) at single-cordon** training system (observations in plotO over 2014—2019). The parameter abbreviation has been denoted for each subplot.

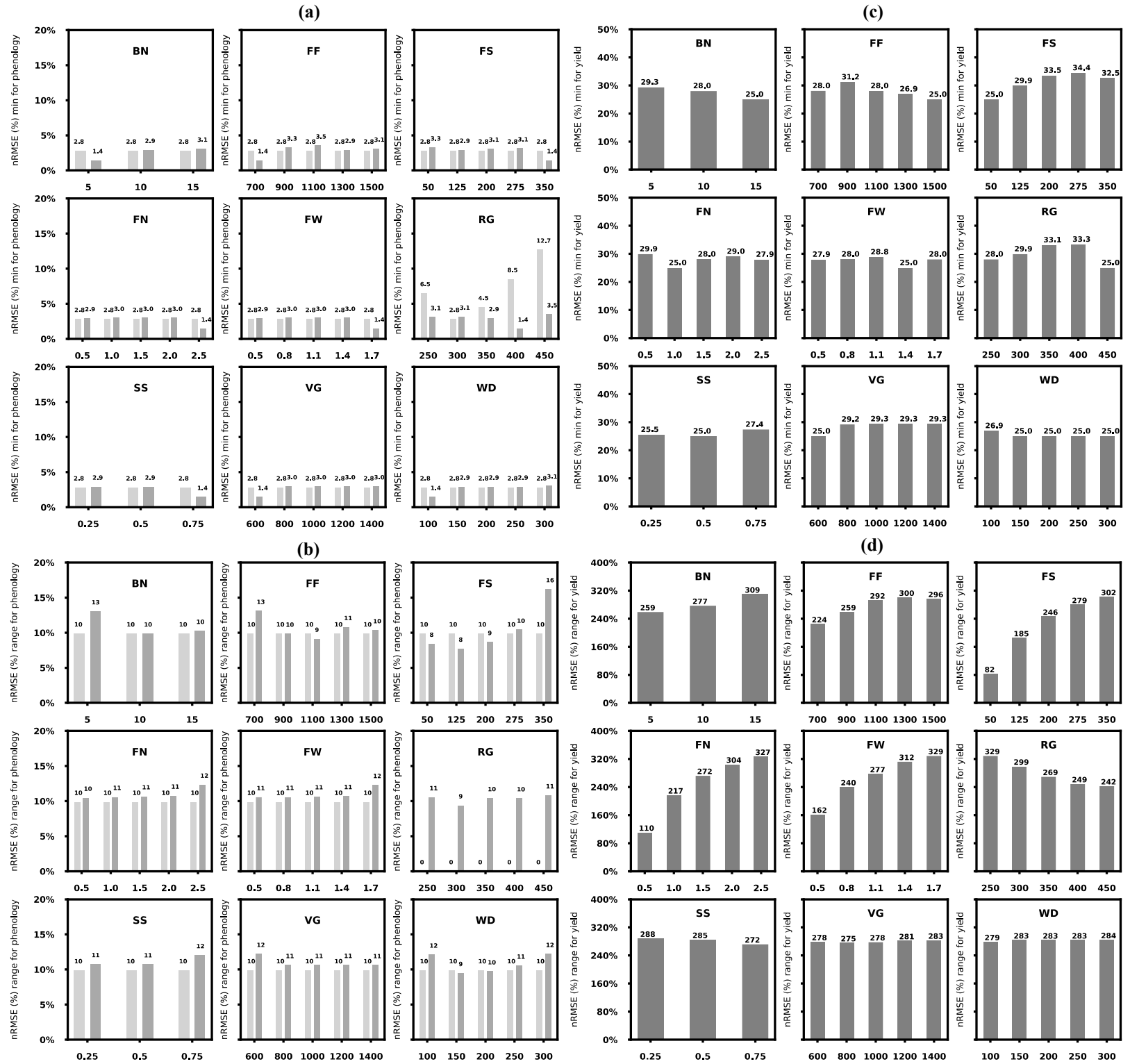

Figure S10. Minimum and range values (spread) of STICS model prediction uncertainties (nRMSE, %) under individual parameter values (fix a given parameter value and test the remaining combinations) on (a—b) phenology (flowering date: light grey bar; harvest date: dark grey bar) and (c—d) yield, respectively. The minimum and range values are respectively withdrawn from the minimum and difference (four times the inter-quartile range) between upper and lower whiskers in boxplots in Figure S5. The analysis concerns the grapevine variety of **Touriga Franca (TF) at double-cordon** training system (observations in plotM over 2014—2019). The parameter abbreviation has been denoted for each subplot.