Table S1. A summary of four field experimental datasets used in this study.

No.	Crop variable	Variety	Location	Season	Sowing	Note
1	Days to heading and maturity	Bohaine, Imam, Elneelain, Taghana	Wad Medani	2005/06 to 2006/07	8 sowing dates from 20 October to 30 December with a 10-day interval	3 replications; ARC's data
2	Days to heading and maturity, plant height	Debeira, Imam	Wad Medani	2008/09 to 2009/10, 2012/13 to 2017/18	23 November to 3 December; one sowing date per year; Sowing date varied by year	2 replications; Iizumi et al. (2021a)
3	Days to heading and maturity	Imam	Wad Medani	2015/16	25 November, 15 December	8 replications; ARC's data
4	Days to heading and maturity, plant height	Wadi Elneel, Debeira, Imam, Zakia	Dongola, Wad Medani, New Halfa	2015/16 to 2016/17	Mid-November to early December (15 and 30 November and 10 December were assumed since specific dates were not reported)	Average of 3 replications; Tahir et al. (2018)

Table S2. Wheat coefficient values used for the seven varieties. *GDD*c, the crop total thermal requirement from sowing to harvesting; $fr_{GDD,ant}$, the fraction of growing season at which flowering occurs; H_{max}, the maximum plant height at an optimal condition; and $T_{u,heat}$, the maximum temperature for the LAI growth and yield formation.

Variety	*GDD*c (°C days)	$fr_{GDD,ant}$ (-)	H_{max} (cm)	$T_{u,heat}$ (°C)
Bohaine	2073	0.628	87.5	35.5
Debeira	2380	0.607	82.0	35.0
Elneelain	2346	0.638	87.5	35.5
Imam	2443	0.643	86.0	36.0
Taghana	2402	0.659	87.5	35.5
Wadi_Elneel	2497	0.572	88.0	35.5
Zakia	2553	0.600	94.0	35.5

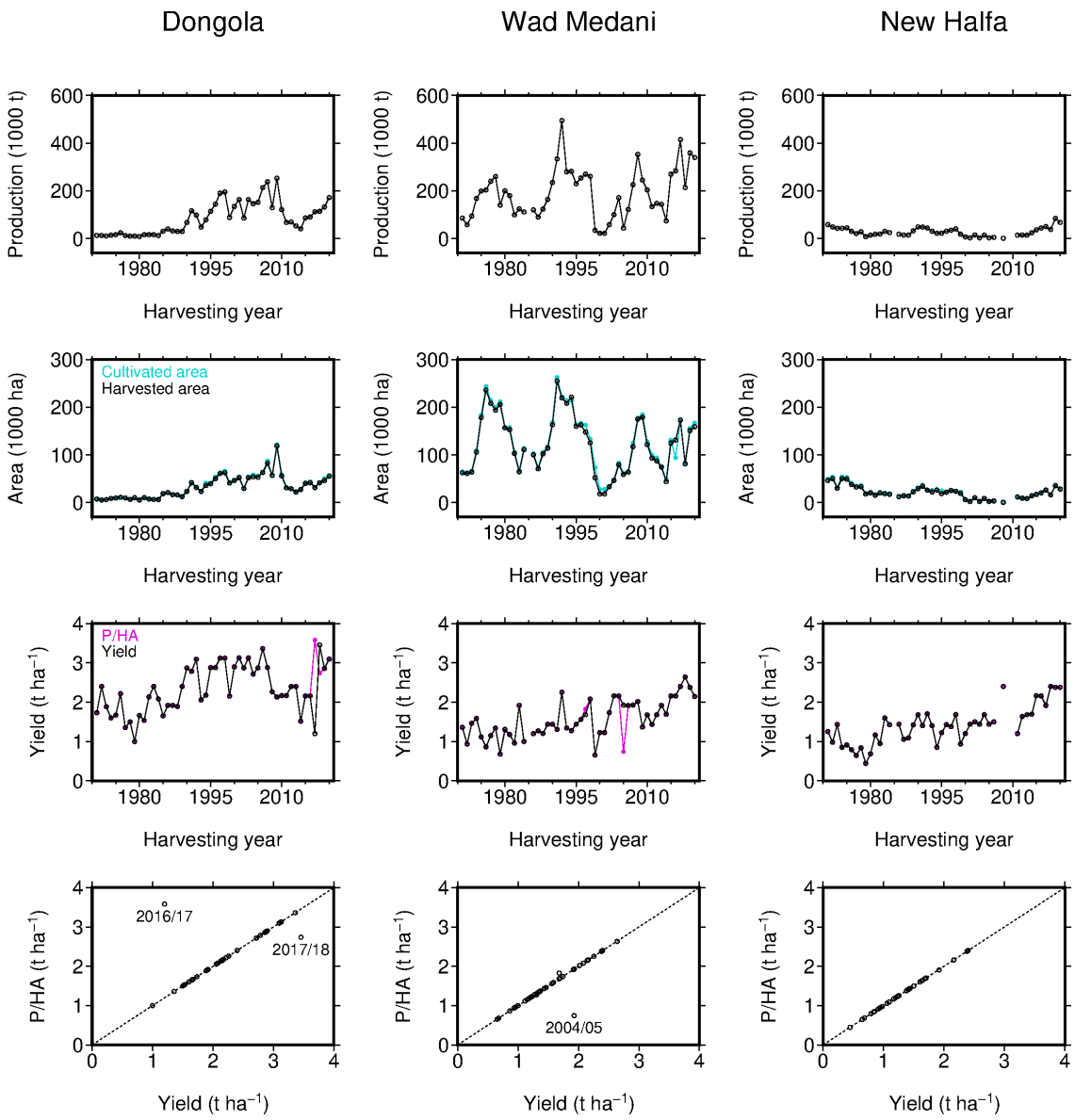

Fig. S1. Annual wheat production, cultivated and harvested area, production divided by harvested area (P/HA) and yield in the studied regions from 1970/71 to 2019/20 seasons and comparisons between P/HA and yield. Seasons of the samples excluded from the analysis are indicated in the bottom panel.

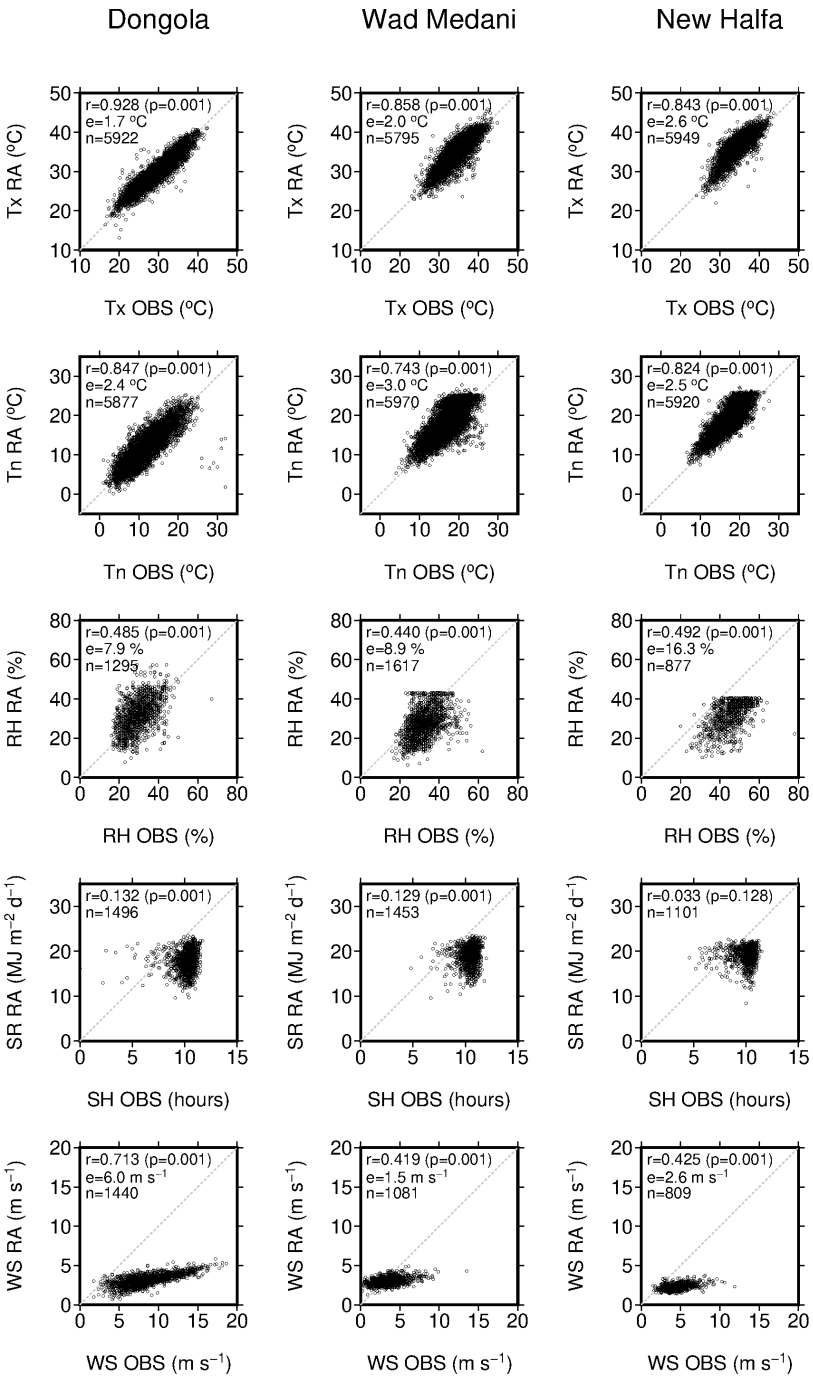

Fig. S2. Comparisons between the local daily observations (OBS) and reanalysis-based product (RA) in the studied regions during the wheat season (NDJF) for the 1970–2020 period. Tx, daily maximum air temperature (°C); Tn, daily minimum air temperature (°C); RH, relative humidity (%); SH, sunshine hours (hours); SR, solar radiation (MJ m^{-2} d^{-1}); and WS, wind speed (m s^{-1}). r, Pearson correlation coefficient; p, p-value; e, root-mean-squared error (RMSE); and n, sample size. The observed sunshine hours are compared with the reanalysis-based solar radiation due to the lack of the corresponding observations. RMSEs are not shown for the reanalysis-based solar radiation because of the inconsistent unit with the observed sunshine hours.

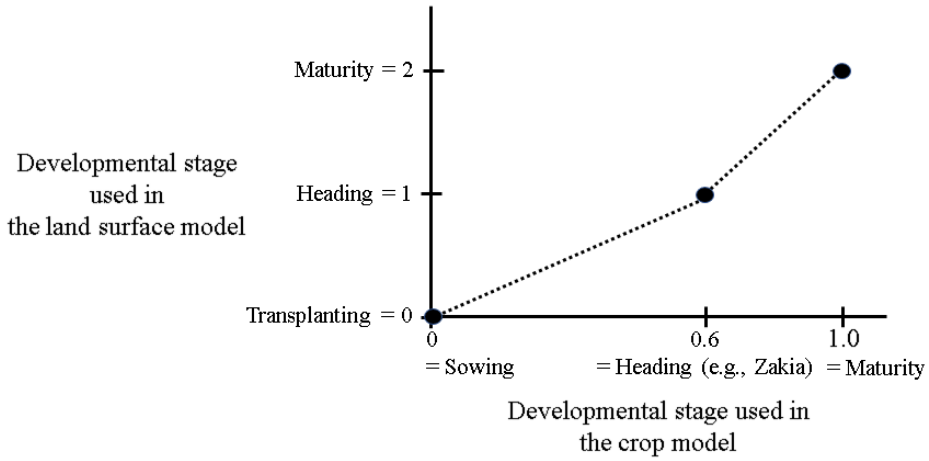

Fig. S3. Schematic illustrating the scaling of the development stage values simulated by the crop model to input to the land surface model.

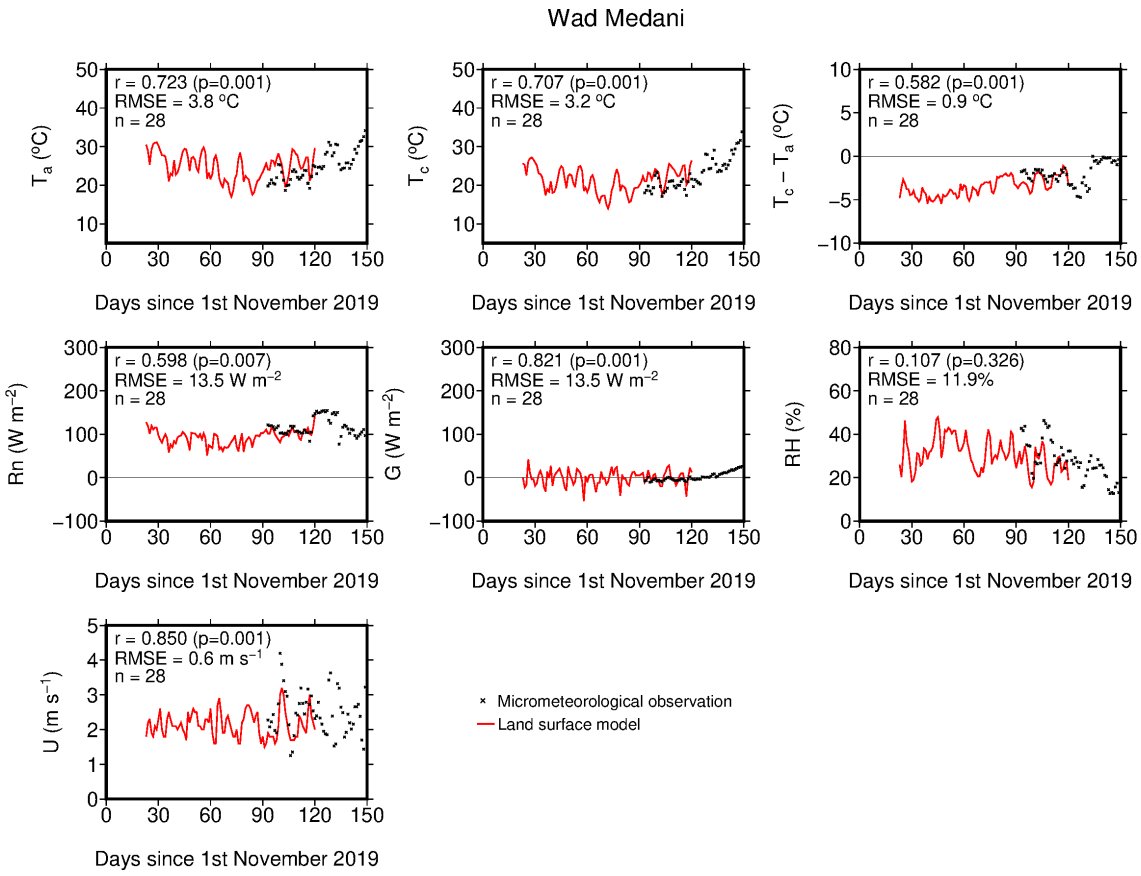

Fig. S4. Comparisons between the micrometeorological observations and land surface model simulation at the wheat field growing Imam in the Gezira research station, Wad Medani for the 2019/2020 season. Ta, the 2-m air temperature; Tc, the wheat canopy temperature; Tc–Ta, the difference between canopy and air temperature; Rn, the net radiation; G, the heat flux into the ground (the sum of heat flux into the water layer and soil heat flux); RH, the relative humidity; and U, the 2-m wind speed. Ta, RH and U are from the reanalysis-based product and input to the land surface model, whereas the other weather variables are simulated by the land surface model. RH was converted from the specific humidity input to the land surface model for consistent comparison with the observations. r, Pearson correlation coefficient; p, p-value; e, root-mean-squared error (RMSE); and n, sample size.

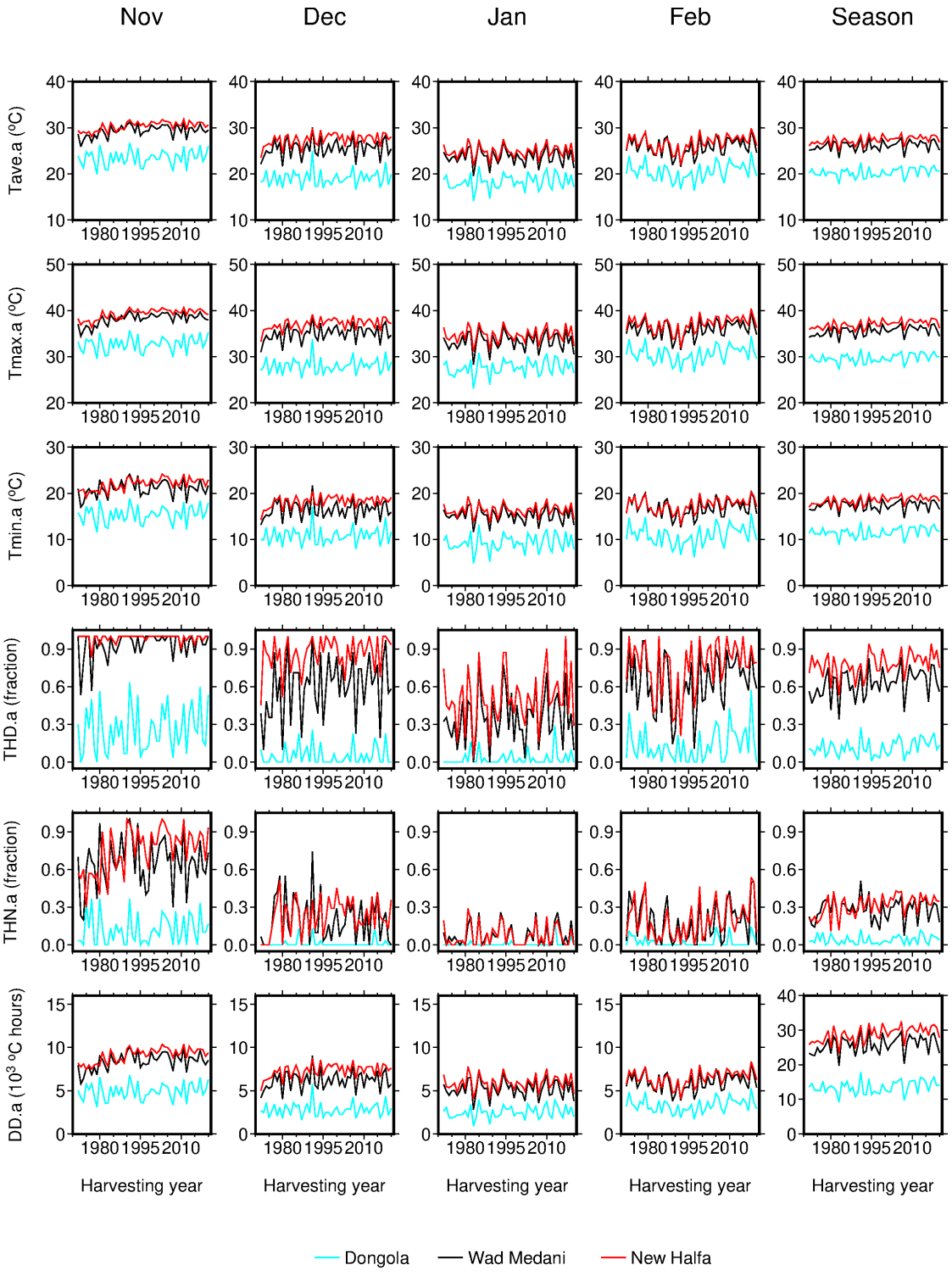

Fig. S5. Ta-based high temperature indicators for the months and season for the 49 seasons (1971/72 to 2019/20).

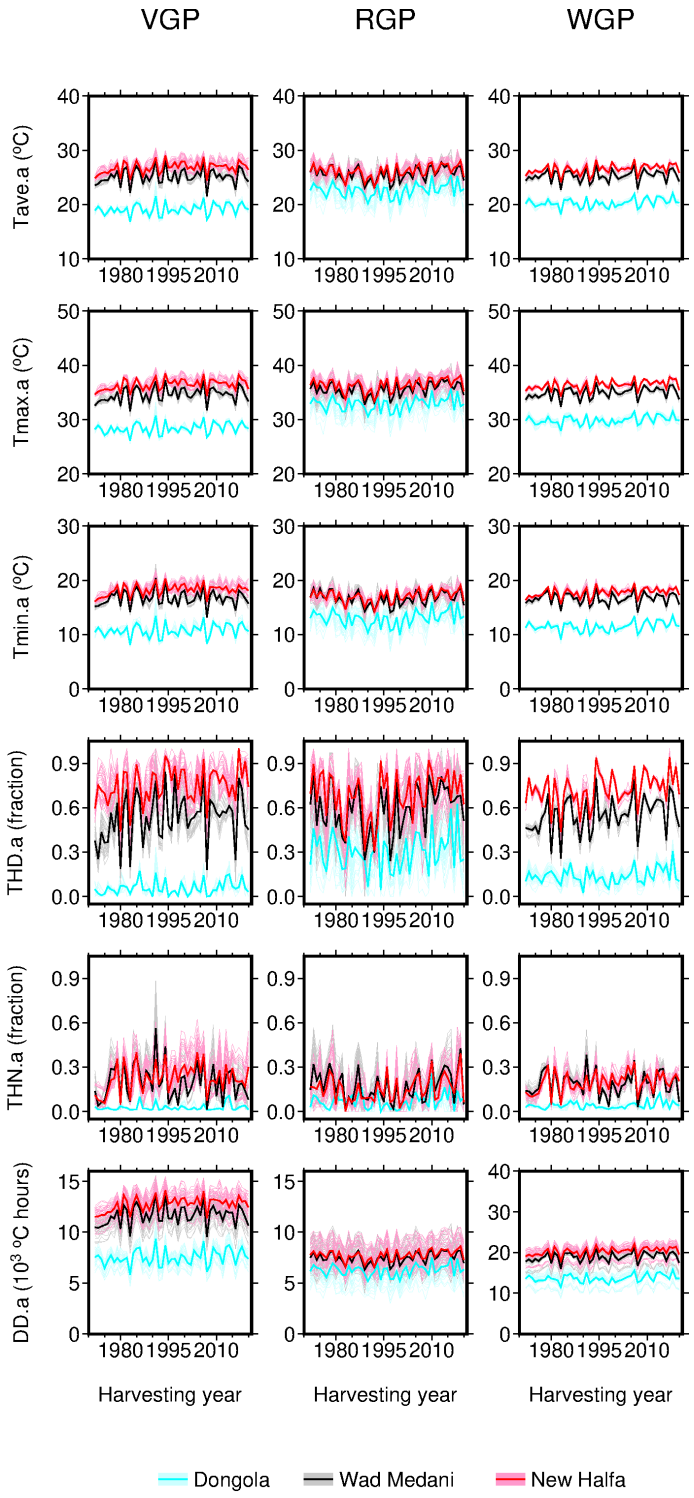

Fig. S6. Ta-based high temperature indicators for the phenological periods for the 49 seasons (1971/72 to 2019/20). Faded and solid lines denote the individual variety-sowing date combinations and their average, respectively.

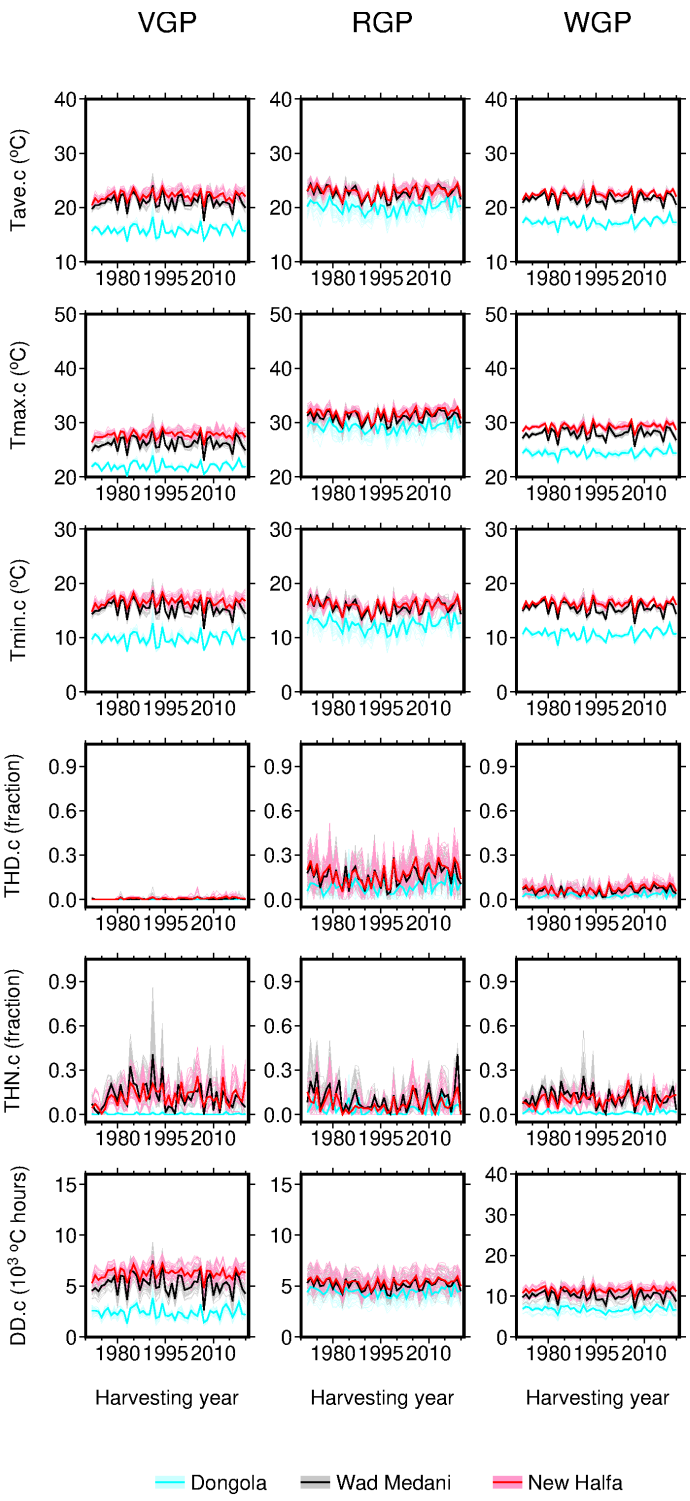

Fig. S7. Same as Fig. S6 but for Tc-based high temperature indicators.

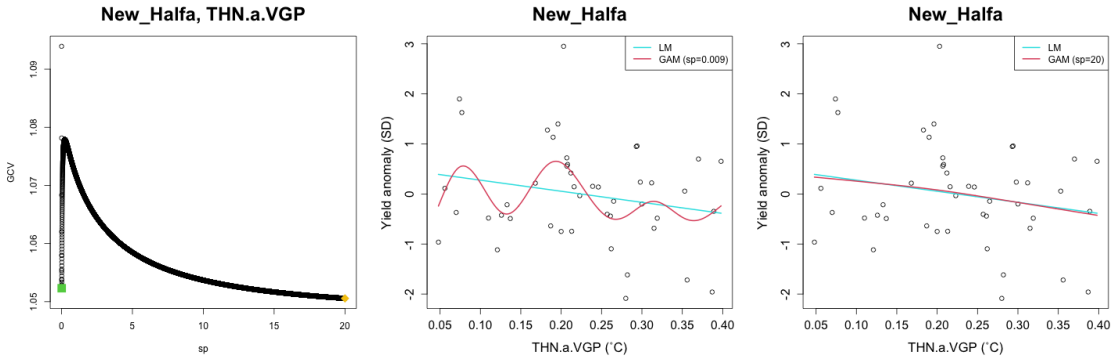

Fig. S8. The selection of the smoothing parameter (*sp*) values in the nonlinear specification. (left) Changes in GCV score in response to increase in the *sp* value. The relationships between the Ta-based VGP THN and yield anomaly in New Halfa estimated using the nonlinear specification with (middle) *sp*=0.009 and (right) *sp*=20. The linear specification between the middle and right panels is identical. These *sp* values and the corresponding GCV score values are indicated in the left panel (the green and yellow symbols).

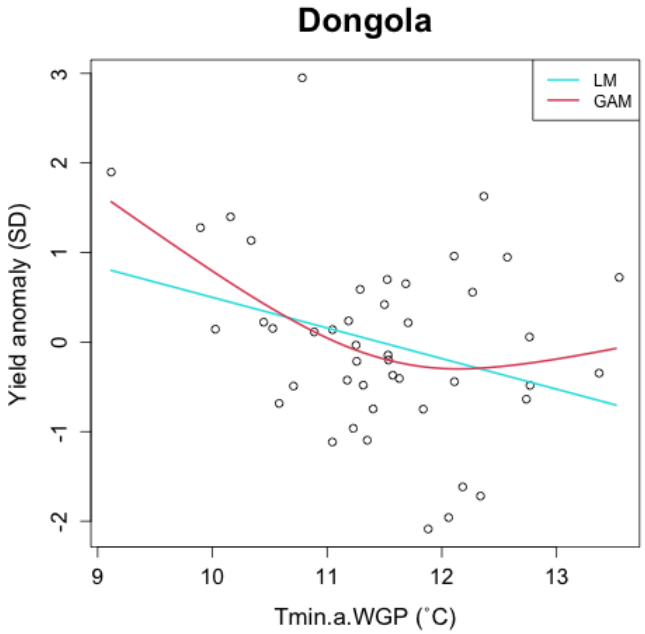

Fig. S9. Relationships between the Ta-based WGP Tmin and yield anomaly in Dongola estimated using the linear (LM) and nonlinear (GAM) specifications.

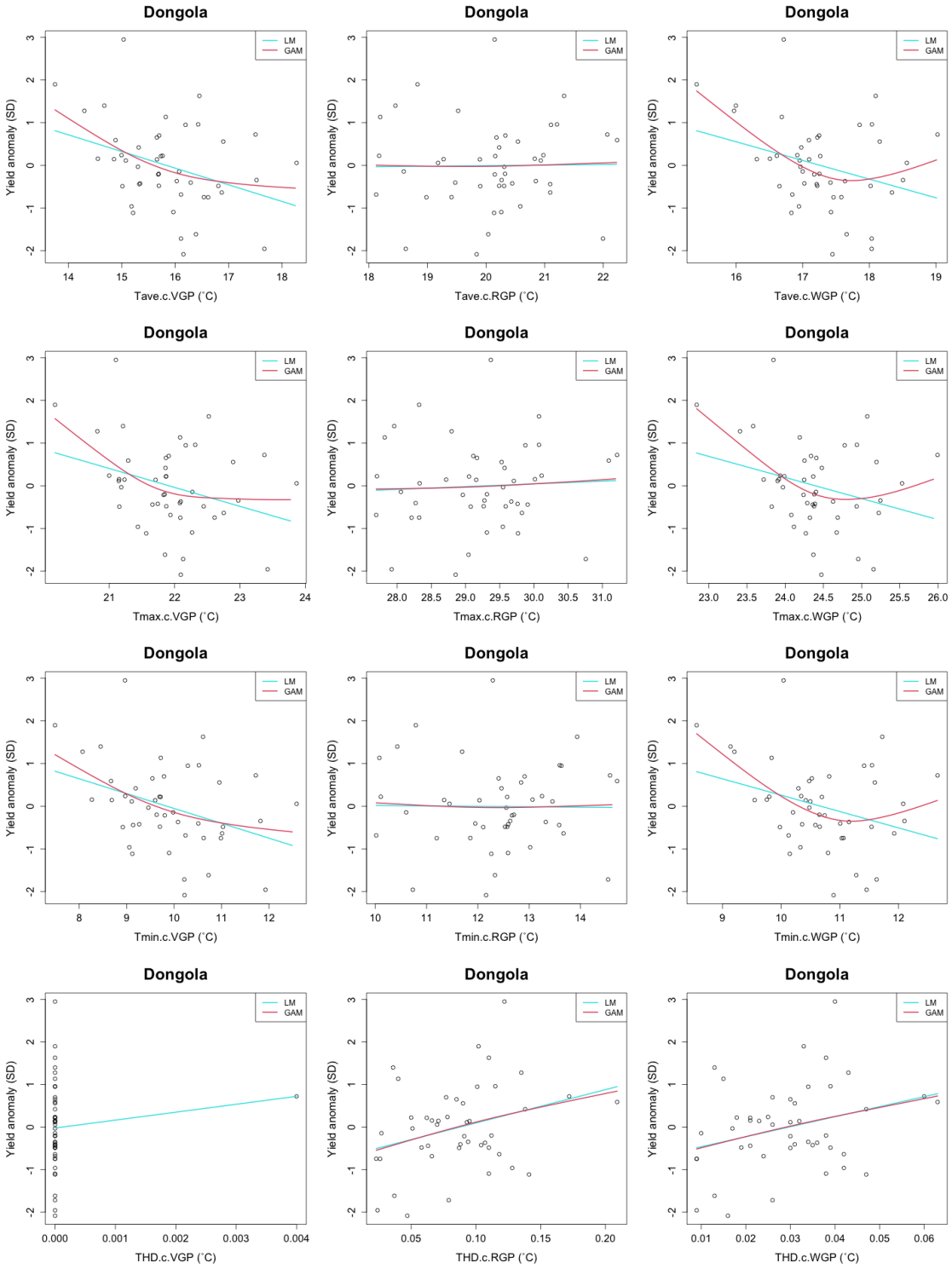

Fig. S10. Relationships between the Tc-based high temperature indicators and yield anomaly in Dongola estimated using the linear (LM) and nonlinear (GAM) specifications. The nonlinear specification was not successfully fitted for VGP THD due to the insufficient number of nonzero samples.

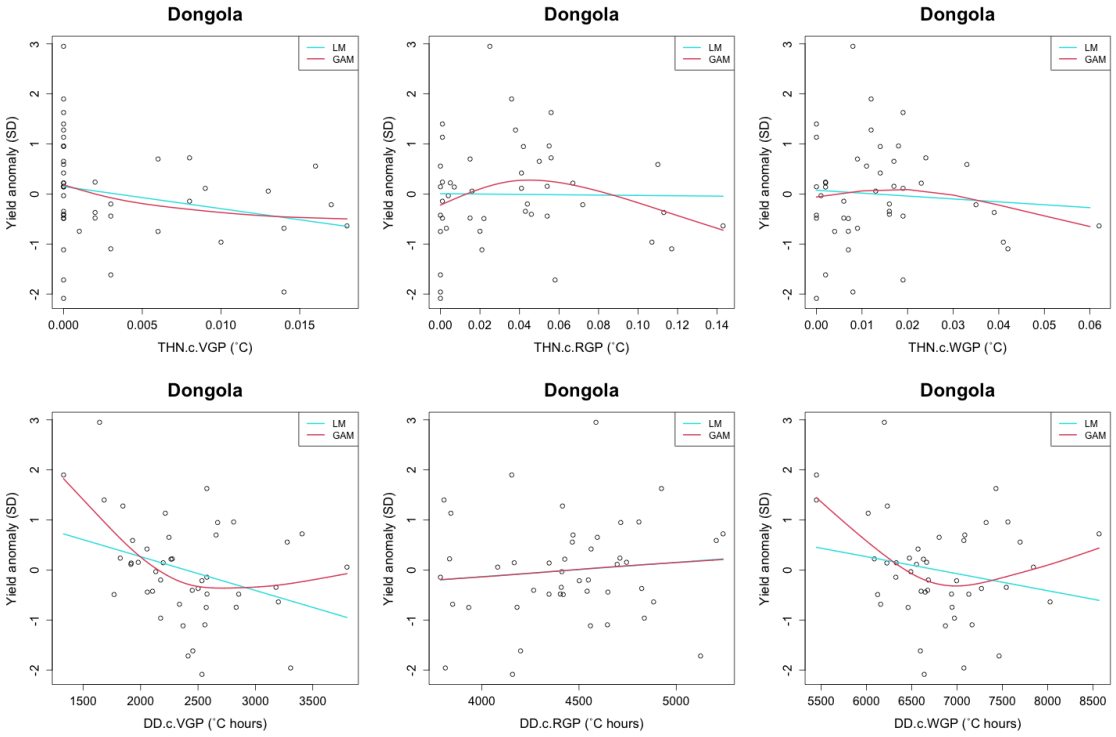

Fig. S10. (continued)

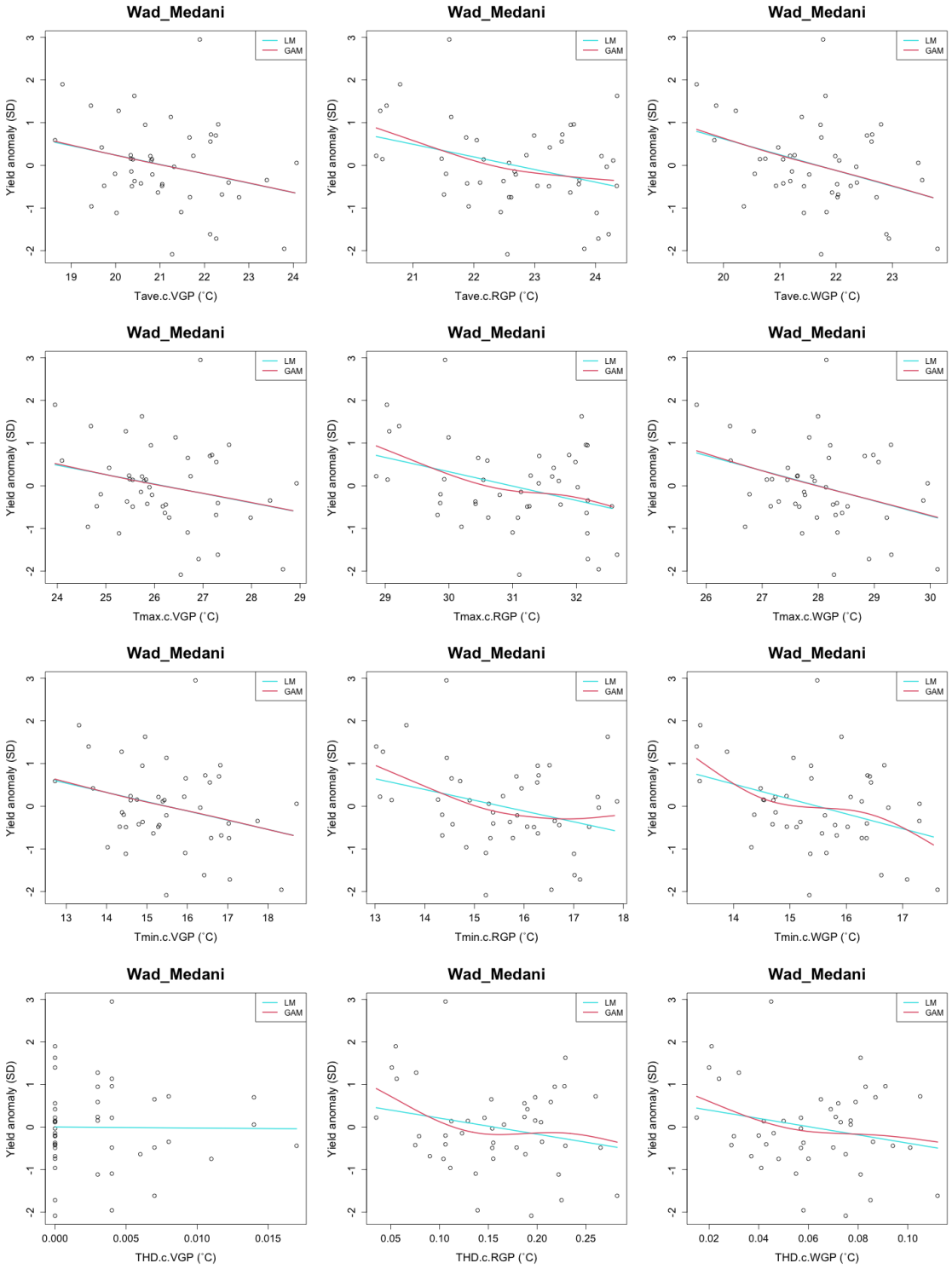

Fig. S11. Same as Fig. S9 but for Wad Medani. The nonlinear specification was not successfully fitted for VGP THD, as in Dongola.

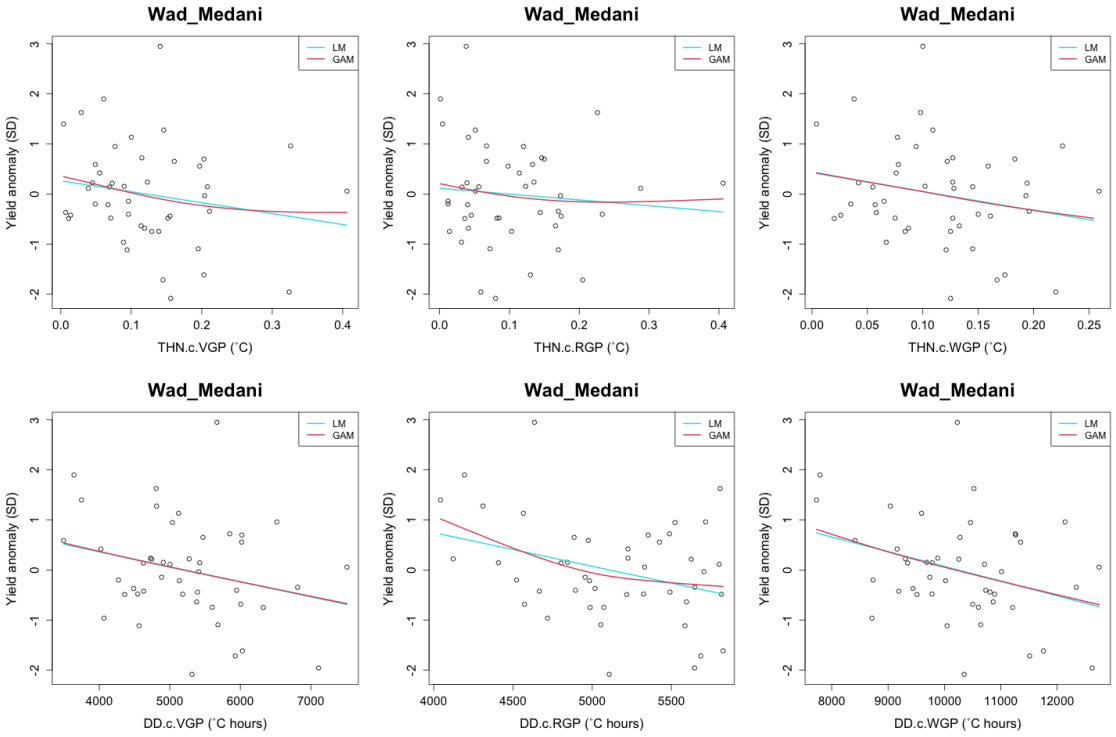

Fig. S11. (continued)

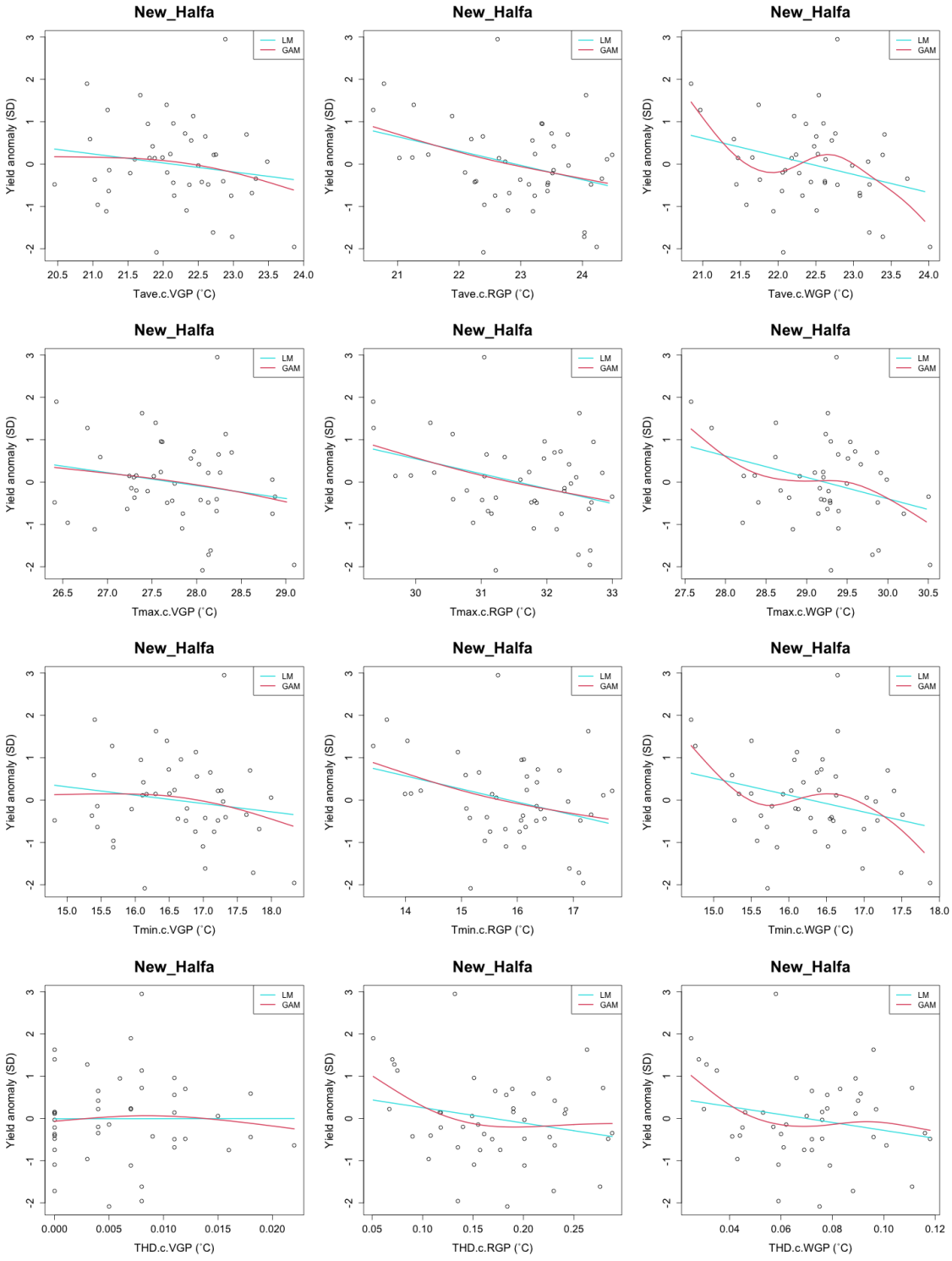

Fig. S12. Same as Fig. S10 but for New Halfa.

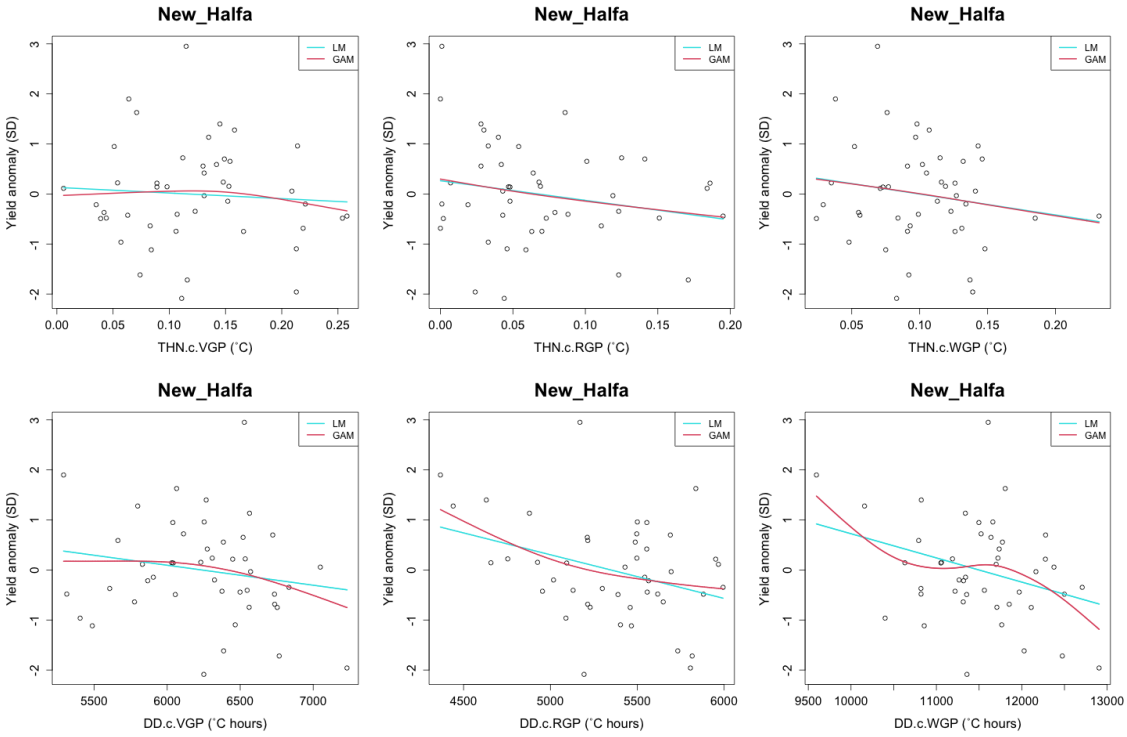

Fig. S12. (continued)