A generalized phenological model for durum wheat: application to the Italian peninsula

Published Version:
A generalized phenological model for durum wheat: application to the Italian peninsula / Di Paola, Arianna; Ventura, Francesca; Vignudelli, Marco; Bombelli, Antonio; Severini, Maurizio. - In: JOURNAL OF THE SCIENCE OF FOOD AND AGRICULTURE. - ISSN 0022-5142. - ELETTRONICO. - 100:11(2020), pp. 4093-4100. [10.1002/jsfa.9864]

This version is available at: https://hdl.handle.net/11585/692769 since: 2019-07-24

Published:

DOI: http://doi.org/10.1002/jsfa.9864

A generalised phenological model for durum wheat. Application to the Italian peninsula.

Di Paola Arianna[1], Francesca Ventura[2], Marco Vignudelli[2], Antonio Bombelli[1], Maurizio Severini [3]

1. Impacts on Agriculture, Forests and Ecosystem Services Division, Euro-Mediterranean Center on Climate Change (CMCC), Viterbo, Italy

2. Department of Agricultural and Food Sciences, University of Bologna, Bologna, Italy

3. Department of Ecology and Sustainable Economic Development, University of Tuscia, Viterbo, Italy.

Corresponding author: arianna.dipaola@cmcc.it

ABSTRACT

BACKGROUND: A likely increasing demand for varieties mixtures, landraces and genetic diversity in cropping systems will underpin calls for models able to generalise phenological development at the species level, while providing the expected range of phenological variability. In the present article, we aimed to obtain a generalised phenological model of durum wheat (*Triticum durum*, Desf.).

RESULTS: By using a large phenological dataset embracing field data collected under different sowing dates, varieties, and locations over the Italian peninsula, we searched for the phenophases enabling the best linear approximations between developmental rates and air temperature, in order to minimize the residual variability from drivers other than temperature, as genetic and environmental diversity. The developmental rates of the resulting phases were then examined with respect to the mean daylength, to determine possible additional relations with photoperiod. If a correlation with daylength was also present, the developmental rate is calibrated by multiple linear regression, otherwise by simple linear regression of temperature. The resulting calibration, tested on an independent data subset, proves that the model is able to generalise wheat development over the Italian peninsula with high accuracy (MAE =3-8 days; R^2= 0.75-0.98), regardless of the wheat variety.

CONCLUSION: The generalised phenological model is potentially suitable for many agro-ecological and large-scale applications. It is hoped that the model will aid situations where phenological observations to parameterize a model are still lacking, as is probably the case for landraces and underutilized crop varieties.

Keywords: *durum wheat*, phenological model, developmental rates, agro-ecology, landraces

1. Introduction

The way the plants progress through their life cycle (i.e. their phenology) represents one of the most important plant/environment adaptation strategy [1,2]. Phenological models are essential tools for organic farming, and other types of crop management, as they allow for scheduling crop practices and irrigation [3]. Moreover, such models can help practices aimed at reducing climate risks and optimizing external resources, and enhancing pest and weed [4,5,6]. However, contemporary modelling tools, of which the phenological module represents a key component, are typically designed for optimising the productivity of monoculture at the field scale [7]. Phenological models for crop mixtures, large-scale simulations and, overall, for situations where no data are available, as may happen for landraces (locally adapted varieties) and underutilized varieties remain absent in literature.

Of the well-established predictor variables for wheat development, namely temperature, vernalization and photoperiod [2,8,9,10,11], temperature is considered the most important [2,8].

Vernalization is the physiological mechanism that plants use to compensate for winter season and to flower in spring [12,13]. Wheat cultivars requiring vernalization become sensible to photoperiod after prolonged exposure to cold temperature, although the amount of cold requirement in the field is still uncertain [14,15]. Plants sensitive to photoperiod grow faster under increasing daylength [16].

A versatile way to model plant development is to regress its rate (the reciprocal of the time to mature a given phase, d^{-1}) against the mean value(s) of the predictor variable(s) experienced during that phase [17]. Here we refer to the resulting equation, whether linear or not, as the Developmental Rate (DR) function.

In the present work, we aimed to obtain a generalised phenological model for durum wheat, valid over the Italian peninsula. By *generalised,* we mean a model that could be used to simulate any wheat variety, climate and agricultural regimes. Such a model would allow reliable applications over numerous case studies, such as those involving landraces and underutilized crop varieties, cultivar mixtures used in agro-ecology to increase the resilience of the field [18], and large-scale simulations of phenological development.

The development of a generalised model implies the use of field observations representative of a wide range of environmental, climatic and genotypic variability. It also requires an approach that minimizes the variability in the developmental rates due to such heterogeneity. To treat with this, we used a large phenological database on durum wheat embracing data collected from diverse years, varieties, sowing dates, and experimental sites across the Italian peninsula to identify the wheat phases with developmental rates better approximated to a linear function of the primary driver (i.e. temperature). After suitable phases were identified, we also examined the developmental rates with respect to the mean daylength, searching for an additional explanatory power from the photoperiod. In the phases where a correlation with daylength also emerged, we estimated developmental rate functions by multiple linear regression with respect to both temperature and photoperiod, otherwise by simple linear regression of temperature.

2 Methods

We opted for linear functions as the related errors are constant, allowing for a robust estimate of critical values of temperature. However, the same approach could be followed also using non-linear functions. The underlying principle of our approach is that developmental rate functions may change (being linear or not) among phases [19]. Thus, by inspecting many phases it is possible to find out those where the best linear temperature responses hold, if any, and, in turns, where the residual variability in the developmental rates explainable by variables other than temperature, as environmental[2] and genotypic[20] diversity, is minimal.

2.1 Developmental Rate functions of Temperature (T) and Photoperiod (P)

The developmental rate of a given phase could be described, in the first instance, as a linear function of temperature, as early suggested by [2,21,22]:

$$DR[T] = a + bT \qquad (1)$$

where

- DR is the developmental rate, i.e. the reciprocal of the time to mature the phase [d^{-1}];
- T is the mean air temperature experienced during the phase [°C];
- a is the intercept[d^{-1}]and b is the slope [$°C^{-1}d^{-1}$] of the linear function, respectively.

The intersection of the linear DR function with the abscissa returns the value for the base temperature T_0[°C] [20,21]:

$$T_0 = -a/b \qquad (2)$$

T_0 represents the critical temperature below which plant development is assumed nil, since the DR would assume negative values.

In the same way, following the approach adopted by [17], we suggest that whenever an additional linear relation between developmental rates and daylength holds, a multiple linear regression can be considered as:

$$DR[T, P] = a + bT + cP \qquad (3)$$

where

- P is the mean daylength during the phase in hours [hr];
- a, b and c the coefficients of the multiple regression.

Formally, even in case of multiple linear regression, the development rate stops when temperature and photoperiod fall below critical values. We name these critical values as T_{0m} and P_{0m} in analogy with the symbol used for base temperature, where the subscript m stands for "multiple". Similarly, the values of T_{0m} and P_{0m} are formally given by the intersection of the linear DR function (projected on a two-dimensional scatter plot) with the corresponding abscissa.

Using phenological observations, i.e. a phase time length and the related mean air temperature and daylength experienced during that phase, it is possible to ascertain whether the development rate is a linear function of mean temperature, or temperature and daylength, and if so, perform a least squares regression to estimate the parameters of Eq. (1) or (3), respectively.

2.2 Using DRs to simulate wheat development

When only the temperature is the explanatory variable, the prediction of a given phenological event, i.e. the number of days to complete a phase, could be achieved with the only inputs of sowing date and daily mean air temperatures as follow:

$$\sum_{j=1}^{S} DR_j[T_j] = 1 \qquad (4)$$

Where DR_j is the daily developmental rate of the phase, S is the phase duration in days and T_j is the mean air temperature [°C] of the *j*-day. The linear behaviour of the phase is analytically expressed as;

$$DR_j = a + bT_j \ (if T_j > T_0) \qquad (5)$$

$$DR_j = 0 \quad (if T_j \leq T_0) \qquad (6)$$

When the sum of the daily rates reaches 1, the end of the phase S is achieved [17]. The starting date (*j=1*) of the next phenological phase is on the day following the end of the current, except the last phase which defines the end of the annual crop life cycle.

Similarly, when both temperature and daylength are explanatory variables, *S* can be obtained as:

$$\sum_{j=1}^{S} DR_j[T_j, P_j] = 1 \qquad (7)$$

Where

$$DR_j = a + bT_j + cP_j \ (if\ T_j > T_{0m}\ and\ P_j > P_{0m}) (8)$$

$$DR_j = 0 \qquad (if\ T_j < T_{0m}\ and\ P_j < P_{0m}) \qquad (9)$$

and P_j is the daily mean daylength [hr] of the *j*-day.

2.3 Data source

In this work, phenological field observations were retrieved from the PHEANGRI database (http://phenagri.entecra.it/) and the Agrophenological Station of Cadriano (University of Bologna, DISTAL, Italy).

PHENAGRI provides a free database of field observations on both weather and crop phenological development, collected from several experimental sites widely spread over Italy, during the period 1996-1999. Weather data provided records from *in situ* meteorological stations (when present) or from the nearest reference station to the experimental field. Reference stations were those belonging to the national networks of the Italian Council for Agricultural Research and Analysis of Agricultural Economics (CREA) or to the Air Force Met service (AFM).

Available observations on durum wheat included dates of several phenological events obtained from the scalar sowing dates (ranging from the beginning of November to the beginning of March), several varieties, and five experimental sites placed in: S. Angelo Lodigiano (LO) and Garica di Podenzano (PC), northern Italy; Vasto (CH) and Foggia (FG), south-central Italy; Cassibile (SR), southern Italy (Fig. 1). Further details on the experimental sites and reference meteorological stations are available at the PHEANGRI project website (http://phenagri.entecra.it/).

Observed wheat phenological events, in BBCH centesimal scale [23] were: sowing (BBCH 00), emergence (BBCH 09), three leaves unfolded (BBCH 13-14), beginning of stem elongation (BBCH 30), second node detectable (BBCH 32), beginning of booting (BBCH 41), beginning/end of heading (BBCH 51/59), beginning/end of anthesis (BBCH 61/69), milk maturity (BBCH 73-77), and physiological maturity (BBCH 89). The dates of each phenological event refer to the median date between the sampled plants (at least ten per variety). Further details on the operational protocol of the PHENAGRI project are reported in [24].

The Agrophenological Station of Cadriano (BO) is an experimental site led by the University of Bologna, collecting data from 2003 to the present, following the same operational protocol of PHENAGRI. Observations enclosed the same phenological events reported above, obtained from autumn sowing dates (October-November), for durum wheat cv. Duilio. Weather data were retrieved from *in situ* agrometeorological station. Details on both the agro-phenological and agro-meteorological station of Cadriano are available in [25,26,27].

Further available data from the Experimental Farm of Cadriano, collected in the period 1972-1978, were also used (shared by personal communication). Observed events, which correspond to BBCH values, were: beginning of stem elongation (BBCH 30), beginning of heading (BBCH 51) and physiological maturity (BBCH 89).

2.4 Data analysis

We defined two subsets of data, namely:

- *Calibration dataset*: Data on durum wheat varieties Creso and Simeto, obtained from the PHENAGRI experimental sites of Garica di Podenzano (PC), Vasto (CH) and Cassibile (SR).
- *Validation dataset*: observations on durum wheat varieties Ares, Cirillo, Colosseo, and Zenit, obtained from the PHENAGRI experimental sites of S. Angelo Lodigiano (LO) and Foggia (FG); observations on varieties Duilio (collected during 2003-2016), and Sansone and Valgerardo (collected during 1972-1978), from the experimental farm of Cadriano (BO).

Figure 1 shows the geographical distribution of the calibration and validation sites.

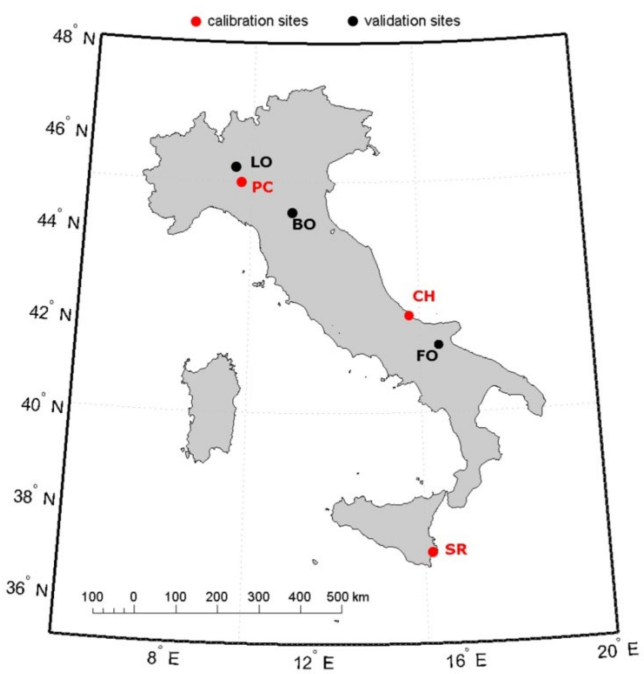

Figure 1. Geographical distribution of the experimental sites. Red dots: experimental sites whose data were used to calibrate the phenological model (*calibration sites*). Black dots: experimental sites whose data were used to test the model performances (*validation sites*).

From the calibration dataset, we estimated the developmental rate of many phenophases and related mean air temperature experienced during each phase. Among suitable combinations to define the whole wheat life cycle, we empirically identified the wheat phases where the relationships between developmental rates and mean air temperature were better approximated to a linear function (p-values<0.01). This selection was achieved by looking at the Pearson linear correlation (*r*) coefficients. Then, we checked for further correlations with daylength. Mean daylength was computed according to the FAO guideline [28] on a daily basis and then averaged over the phase time length. If a further correlation with daylength held, and in the case of no collinearity between temperature and daylength, DR functions were regressed using ordinary least squares technique in the form of Eq. (3), otherwise in the form of Eq. (1). For each phase we provided *i*) the Pearson correlation coefficient (*r*) ;*ii*) the coefficients of the linear functions and the critical values for temperature and daylength (if any); *iii*) the residuals to check that the correlations are unbiased and homoskedastic; iv) the error variable(ε).

The coefficients estimated for each DR function were used to simulate wheat development at the validation sites, according to eq. (7) or (4) depending on whether the developmental rates are also correlated to photoperiod or not, respectively.

Results were compared with the observations (*validation set*) looking at the Mean Absolute Errors (MAE, [d]), normalised MAE (NMAE, [%]) and the model efficiency (EF, dimensionless).

The statistical indices are defined as:

$$MAE = \frac{\sum_{i=1}^{n}(|S_i - O_i|)}{n} \qquad (11)$$

$$NMAE = \frac{MAE}{\bar{O}} \tag{12}$$

$$EF = 1 - \frac{\sum_{i=1}^{n}(S_i - O_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2} \tag{13}$$

Where S_i and O_i are the days after sowing of the i^{th} prediction and observation, respectively, n the number of observations and $\bar{O}$ is the mean observation from the whole dataset.

MAE [29] indicates the mean absolute values of errors (predictions minus observations) in absolute terms. The lower the values of MAE, the higher the agreement of the model prediction with the observations. NMAE expresses the MAE with respect to the observed phase length, which permits the expression of errors in relative terms. Generally, simulations are considered excellent when relative errors are less than 10%, good when ranging from 10 and 20%, fair between 20-30% and poor when greater than 30% [30].

The EF [31] compares the deviance of the errors (described by the numerator) with that of the observations (described by the denominator). Its maximum value is 1 and indicates complete agreement between predictions and observations. The EF decreases with decreasing predictive power of the model until reaching negative values, meaning that the model describes the data less well than the arithmetical mean of the observations.

3. Results

The phases whose relationship between DR and mean temperature (T) is better approximated by a linear function were:

1) sowing to three leaves unfolded (S-3L, BBCH 0-13);

2) three leaves unfolded to second node detectable (3L-2N, BBCH 13-32);

3) second node detectable to the beginning of heading (2N-H, BBCH 32-51);

4) beginning of heading to physiological maturity (H-M, BBCH51-89).

Linear relationships between developmental rates and mean air temperature for single sub-phases are shown in Figure 2. Coefficients and related statistics for the simple linear DR functions are given in Table 1.

Table 1. Statistics of the simple linear regression for the selected phases. Phases as in Fig.1.

DR = a + bT (eq. 1)				
phase	***T_0 [°C]***	***a[d^{-1}]***	***b[$°C^{-1}d^{-1}$]***	***ε[d^{-1}]***
S-3L	-3.3 ± 2.3	0.0055	0.0017	0.0040
3L-2N	3.3 ± 1.5	-0.0119	0.0036	0.0054
2N-H	9.2 ± 1.5	-0.1018	0.0110	0.0169
H-M	11.2 ± 0.9	-0.0323	0.0029	0.0026

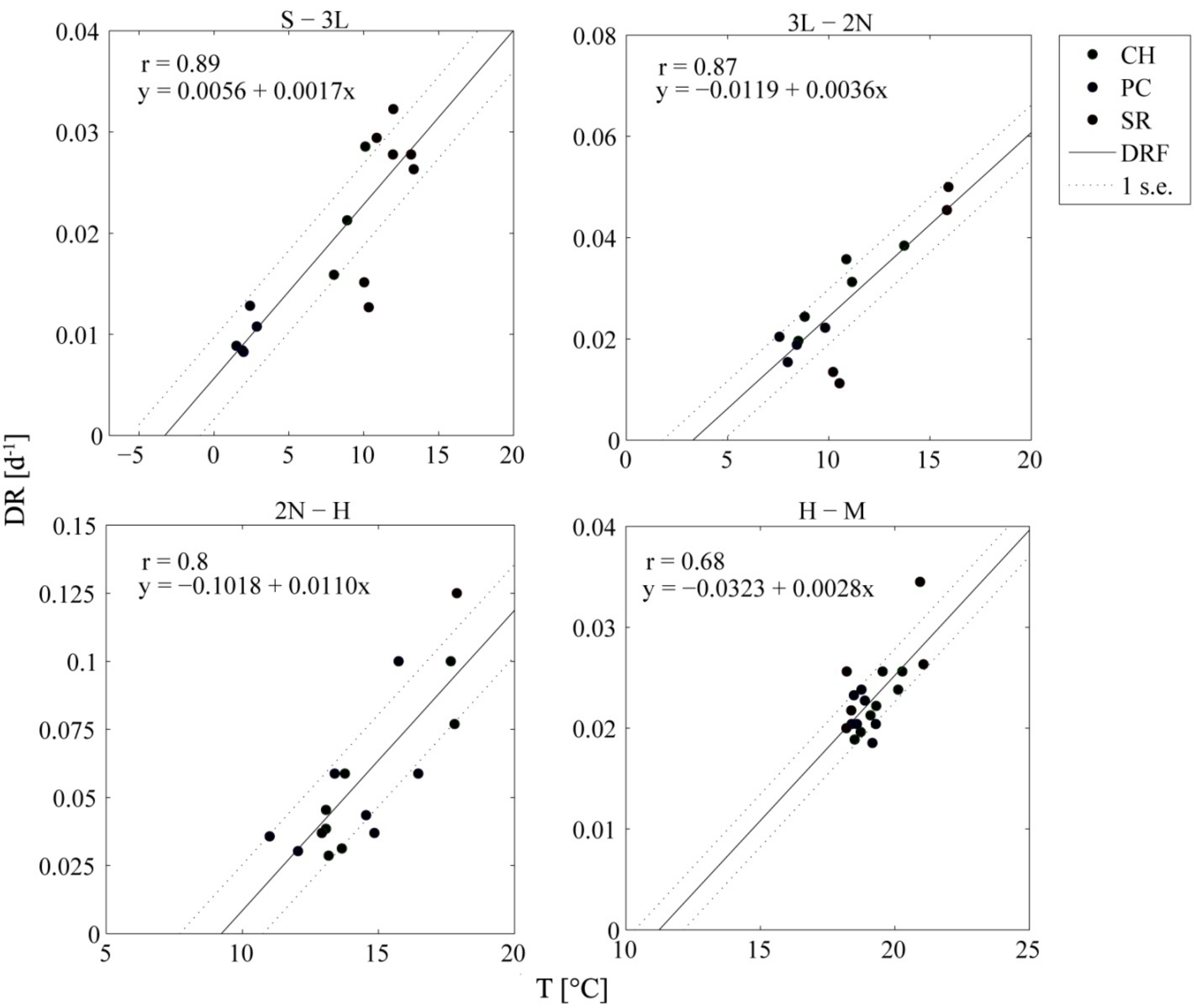

Figure 2. Relationshis between wheat developmental rates and mean air temperature (from the calibration dataset) for the selected wheat phases. Black lines: DR function, dotted lines: 1 standard error (ε), used to track back to the uncertainty in the base temperature. S-3L: from sowing to three leaves unfolded; 3L-2N: from three leaves unfolded to second node detectable; 2N-H from second node detectable to the beginning of heading; H-M: from the beginning of heading to physiological maturity.

In the first three phases the correlations between development rate and temperature were very high (*r*=0.80-0.89), whilst in H-M the correlation was lower (*r* = 0.68) and the range of mean temperatures experienced was narrower (about 4°C) when compared to the other phases (ca. 15°C in S-3L; 10°C in 3L-2N; 7°C in 2N-H). Temperature alone explained 64-79% of developmental rates (values from r^2) in the first three phases, 46% in H-M.

Base temperatures increased throughout the wheat crop life cycle, ranging from -3.3°C in S-3L to 11.2°C in H-M (Tab.1 and Fig S1). Uncertainties in the base temperature, quantified by ε (see dotted line in Fig. 1) were relatively large in S-3L and decreased in the subsequent phases (Tab.1). The slopes of DRs regularly increased throughout the vegetative phases and slow down after heading (Fig. S1). Residuals (Fig. S2) revealed homogeneous variance and no bias.

The phase 3L-2N was the only showing a significant ($p<0.01$) correlation with daylength (Fig. 3, left panel). In this phase, no correlation resulted between temperature and daylength, suggesting no collinearity between explanatory variables (Fig 3, middle panel). The multiple linear regression (Fig. 3, right panel) between developmental rate, temperature and daylength increased the explanatory power of the regression from 0.75 (r^2) to 0.82 (adjusted r^2). Related coefficients and statistics for the improved DR are summarized in Table 2. Values for T_{0m} and P_{0m} (Fig S3) were 5.6 °C and 9.5 Hr, respectively.

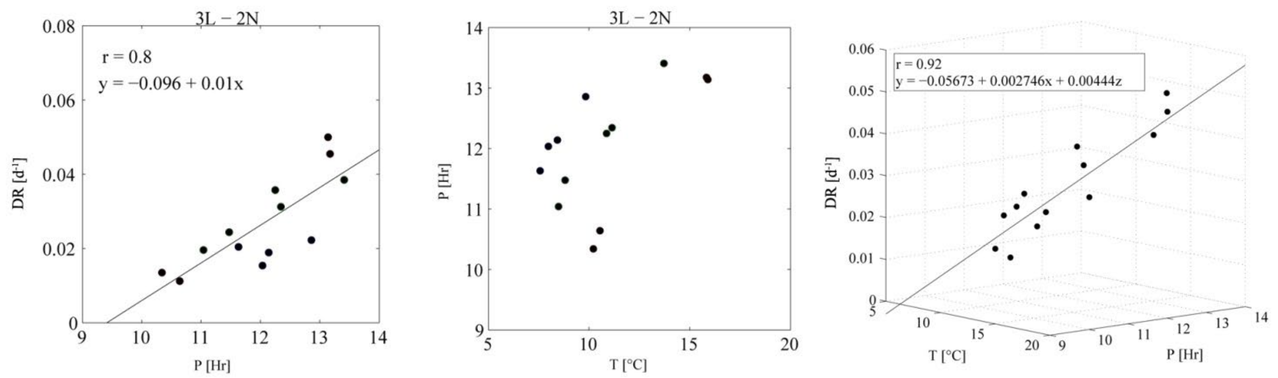

Figure 3. Left: relationship between wheat developmental rates and mean daylength (P [Hr]); middle: mean temperature vs. mean daylength scatter plot (showing no collinearity); right: multiple linear regression defining developmental rates in 3L-2N as a function of temperature and daylength.

Table 2. Statistics of the multiple linear regression for 3L-2N.

DR = a + bT + cP (eq. 4)						
phase	T_{0m}	P_{0m}	***a***	***b***	**c**	***adj. r²***
3L-2N	5.6	9.5	-0.0567	0.0027	0.0044	0.84

The performance of the generalised phenological model, which uses simple linear functions of temperatures in the three phases (S-3L; 2N-H; H-M) and multiple linear function of temperature and photoperiod in 3L-2N, is shown in Figure 4, while the corresponding statistics for each phase are summarised in Table 3. Overall, simulations were able to catch the general pattern of wheat development (EF=0.80-0.99) and mean errors remained below eight days.

Table 3. Statistical evaluation of the phenological model over the validation dataset for single phases. *n*: sample size; $\bar{O}$(std): mean observed phase length (1 standard deviation); MAE: Mean Absolute Error; NMAE: normalized MAE; r^2 determination coefficient; correlation between observations and predictions were always significant ($p<0.01$).

Phenological phases	***N***	$\bar{O}$ **(std) [d]**	**MAE [d]**	**NMAE [%]**	**EF**	r^2
S-3L	22	55(23)	7.8	14	0.80	0.75
3L-2N	23	135(35)	7.0	5	0.91	0.83
2N-H	32	162(44)	4.5	2	0.98	0.96

H-M	30	200 (40)	3	1	0.99	0.98

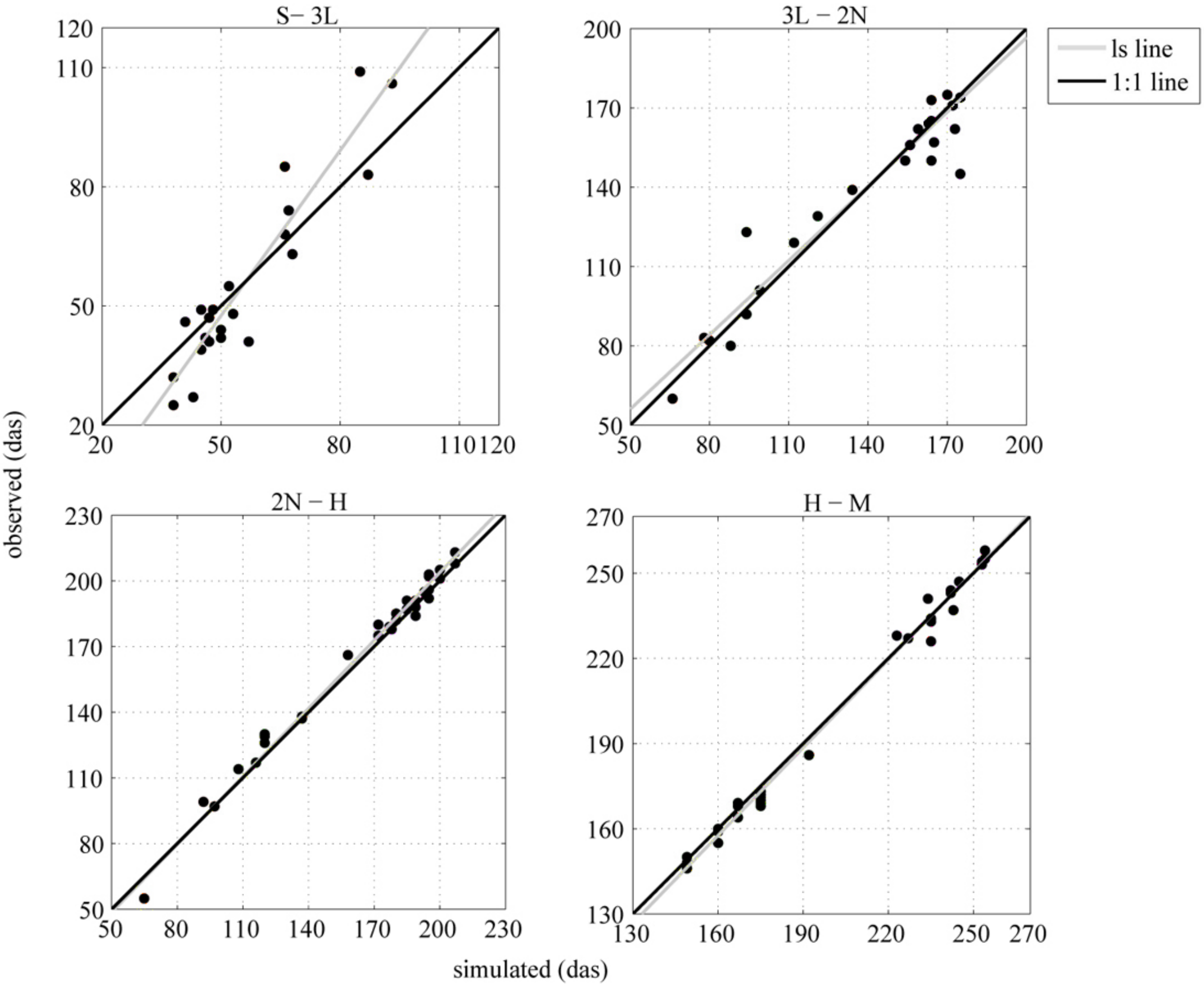

Figure 3. Simulated vs. observed durum wheat phenological events (3L, 2N, H and M) after sowing (S), from the independent validation dataset. Black line: 1:1 line, grey line: least square line.

Largest deviations from the observations were found in 3L, where MAE and NMAE are high (MAE = 7.8 days; NMAE = 14%). The simulation of the subsequent events improved gradually. In 2N, the model efficiency was very good (EF = 0.91), albeit few predictions were far from observations, keeping MAE relatively high (MAE=7). In H and M predictions and observations are strongly correlated ($r^2 \geq 0.98$). Indeed, the efficiency of the model to predict H and M results high (EF ≥ 0.98), NMAE excellent (NMAE≤2%) and MAE 4.5 and 3 days, respectively. Predictions of M are the most accurate.

In Table S2 we also provide the results that would be obtained if predictions were made by the model not improved by daylength. Results obtained without considering the photoperiod have higher inaccuracy, especially in the predictions of H where the mean errors would be doubled (around 8 days).

4. Discussion and Conclusion

A generalised phenological model for durum wheat valid for the Italian peninsula was obtained by using a large phenological database and searching for the phases where the residual variability in the developmental rates were minimized among linear temperature responses. The resulting DR functions gave satisfactory results (EF = 0.89-0.99) over different Italian temperature regimes and wheat varieties.

The generalised model has a practical advantage of being usable under a wide range of environmental conditions where the reference to single wheat variety, climate and the agronomic regime could be reductive. To date, most contemporary crop models are developed for monoculture systems, where the specific crop variety, environmental condition and management practices are well defined in the model. The present model has several examples of potential applications, such as regional simulations as, for instance, long-term impact analysis due to climate change (e.g. how plant phenology is shifting due to global climate change, [32]) and land suitability analysis [33], including the identification of the optimal sowing window to minimise the risk of spring frost and late-season drought. Interesting applications of the generalised model could also be proposed for agro-ecological purposes where, for instance, landraces, intra-specific crop mixtures and crop diversification are recommended to improve the resilience of the system, promote pest regulation and enhance nutrient recycling [18,34,35]. Although the model is yet to be tested on ancient varieties, we recommend it as a suitable tool for modeling the phenology of landraces and underutilized variety (i.e. situations where observational data are few) due to its high level of generalisation. Indeed, the consequence of transition from landraces to modern varieties in the phenology of Italian durum wheat is still poorly understood. In some cases, the transition appears to be a steady advance in anthesis date [36], yet in others no significant changes are observed [37].

Our results were achieved following a method distinct from that of scholars who typically develop phenological model, since the wheat phases were not established *a priori* and the model calibration was carried out only after the strongest linear responses, involving different wheat varieties, had been identified.

Following our approach, temperature alone can explain approximately 64-79% of the variability in the developmental rates from sowing to the beginning of heading, observed from different cultivars. Indeed, the early phenological models, which were based on air temperature, could explain most of the observed developmental-time variability [21, 22, 38]. Later, photoperiod and vernalization were proven to further explain the observed variability in wheat development [9,10,11].

Accordingly, our results show a significant correlation between developmental rates and daylength in 3L-2N, and the ability to explain the observed variability in that phase increases from 75% to 85% when introducing the photoperiod. Photoperiod increases the accuracy of the model, mostly on the predictions of the beginning of heading. A similar improvement was also reported in McMaster and Smika [10].

Overall, we argue that there will always be a variability in wheat development rates not explained by temperature and photoperiod (e.g. due to the environmental heterogeneity, proximity of the meteorological stations, uncertainty in the measurements, genotypic differences, etc.), but, in our case, such variability was minimized by two concomitant strategies: *i*) using data from different experimental sites and sowing periods, which provided a wide range of explored temperatures; *ii*) identifying phenological phases with variability in the developmental rates better explained by a linear dependence on temperature. The obtained result is a set of simple linear relationships describing the widely-recognised general rule that plants grow faster when the temperature is warmer, enhanced by an increasing photoperiod, but with a unique parameterisation for durum wheat over the Italian peninsula.

The largest errors were found in the predictions of 3L (7.8 days) and 2N (7.0 days), which was consistent with the larger uncertainties in the base temperatures in S-3L and 3L-2N (Tab. 1). Our results are in agreement with previous studies reporting that wheat phenological events facing the winter and falling into the period of tillering (up to the beginning of stem elongation) generally show large variability and are also the hardest to predict [39,40]. Moreover, the resulting errors are also comparable with those found elsewhere [39,40, 41,42,43,44] with values ranging from 3 to 11 days, depending on the phenological phase.

Errors on H(4.5) and M(3.0) are also in line with the observed variability from 193 durum wheat varieties, including landraces, representative of the Mediterranean basin, as reported by Soriano et al. [45].

Our results also show that base temperature progressively increased throughout the crop life cycle (Tab. 1, Fig.2 and Fig. S1), with values are in line with the base temperatures reported in Porter and Gawith's review of the literature [15]. The progressive increase of base temperatures has already been documented elsewhere [21, 22] and used to explain non-linear temperature responses observed over long wheat phases [47]. In particular, Slafer and Rawson [46] report that long phases, as seedling to anthesis, can show curvilinear temperature response, but also a clearly linear relationship during shorter phases.

Our work does feature some simplification and limitations that should be described.
First, phenophases, generally, reflect some scientific or managerial interest, such as periods when the plant is particularly vulnerable or demanding for nutrient, where a farmer can, for example, intervene or purchase insurance. Similarly, some phases are adequate for the analysis of climate change impacts, pests, and diseases. From an agronomical point of view, our approach could lead to significant results for unattractive phases.
Second, in our work vernalization is not explicitly modeled. This can lead to some errors in the parameterisation. For instance, if some plants requiring vernalization are not satisfied, the variability in the observed developmental rates will increase (as data could be more scattered). In our case, the strong correlation between developmental rates and daylength in 3L-2N suggests that most of the plants have likely satisfied their vernalization requirement, or, at least, might not have required any vernalization, being sensitive to photoperiod. The wide range of mean temperature needed for vernalization (roughly from -1° to 15 °C, optimally from 4° to 6° C, see Porter and Gawith. [15]), as well as the possibility of seeds not requiring field vernalization, helped our approach; however, an implicit limitation remains.
Third, the model could not reflect the impact of prolonged extreme temperatures since linear rate functions may underestimate the development times under extremely high temperatures. However, in our data and due to the adopted approach, which search for the best linear temperature responses, there are no observations (i.e. points in the scatter plots of Fig. 2) that suggest a likely drop in the rates in the selected phases under the Italian growing season. Indeed, observations subjected to few days with extremely high temperatures would unlikely result in a drop in the developmental rates since data are average values over several days. Here, we argue that when extreme high temperatures arise for only a few days, the model's result would likely be within the expected error. Otherwise development times could be underestimated.

In conclusion, a likely increasing demand for genetic diversity in monoculture will call for models that are able to generalise phenological development at the species level, and provide the expected range of variability. To date, such models are still poorly developed. Our work could be a step forward toward a new modelling approach that to support agro-ecology applications.

Acknowledgements

We deeply thank Prof. Giovanni Toderi (UNIBO) for providing us with a database of historical phenological data.

Declaration of interests

The authors have no competing interests to declare

Author Contributions

A.D.P. and M.S. conceived and planned the study. M.V. and F.V. collected the data. ADP performed the analysis. All the authors contributed to analyzed and discussed the results and wrote the paper.

References

1. Shorter, R., Lawn, R. J., & Hammer, G. L. (1991). Improving genotypic adaptation in crops–a role for breeders, physiologists and modellers. *Experimental Agriculture*, *27*(2), 155-175.
2. Slafer, G. A., & Rawson, H. M. (1994). Sensitivity of wheat phasic development to major environmental factors: a re-examination of some assumptions made by physiologists and modellers. *Functional Plant Biology*, 21(4), 393-426.
3. McMaster, G. S., & Wilhelm, W. W. (2003). Phenological responses of wheat and barley to water and temperature: improving simulation models. *The Journal of Agricultural Science*, *141*(2), 129-147.
4. Andres L.A., Goeden R.D. (1971) The Biological Control of Weeds by Introduced Natural Enemies. In: Huffaker C.B. (eds) Biological Control. Springer, Boston, MA
5. Ghersa, C. M., & Holt, J. S. (1995). Using phenology prediction in weed management: a review. *Weed Research*, *35*(6), 461-470.
6. Landis, D. A., Wratten, S. D., &Gurr, G. M. (2000). Habitat management to conserve natural enemies of arthropod pests in agriculture. *Annual Review of Entomology*, *45*(1), 175-201.
7. Malézieux, E., Crozat, Y., Dupraz, C., Laurans, M., Makowski, D., Ozier-Lafontaine, H., ... &Valantin-Morison, M. (2009). Mixing plant species in cropping systems: concepts, tools and models: a review. In *Sustainable Agriculture* (pp. 329-353). Springer, Dordrecht
8. McMaster, G. S., White, J. W., Hunt, L. A., Jamieson, P. D., Dhillon, S. S., & Ortiz-Monasterio, J. I. (2008). Simulating the influence of vernalization, photoperiod and optimum temperature on wheat developmental rates. *Annals of Botany*, 102(4), 561-569
9. Masle, J., Doussinault, G., Farquhar, G. D., & Sun, B. (1989). Foliar stage in wheat correlates better to photothermal time than to thermal time. *Plant, Cell & Environment*, 12(3), 235-247.
10. McMaster, G. S., &Smika, D. E. (1988). Estimation and evaluation of winter wheat phenology in the central Great Plains. *Agricultural and Forest Meteorology*, 43(1), 1-18.
11. Saiyed, I. M., Bullock, P. R., Sapirstein, H. D., Finlay, G. J., & Jarvis, C. K. (2009). Thermal time models for estimating wheat phenological development and weather-based relationships to wheat quality. *Canadian Jjournal of Plant Science*, *89*(3), 429-439.

12. Amasino, R. (2004). Vernalization, competence, and the epigenetic memory of winter. *The Plant Cell*, *16*(10), 2553-2559.
13. Sung, S., &Amasino, R. M. (2004). Vernalization and epigenetics: how plants remember winter. *Current Opinion in Plant Biology*, *7*(1), 4-10.
14. Brooking, I. R., & Jamieson, P. D. (2002). Temperature and photoperiod response of vernalization in near-isogenic lines of wheat. *Field Crops Research*, *79*(1), 21-38.
15. Porter, J. R., & Gawith, M. (1999). Temperatures and the growth and development of wheat: a review. *European Journal of Agronomy*, *10*(1), 23-36.
16. Slafer, G. A., & Rawson, H. M. (1995a). Photoperiod× temperature interactions in contrasting wheat genotypes: time to heading and final leaf number. *Field Crops Research*, 44(2), 73-83.
17. Angus, J. F., Mackenzie, D. H., Morton, R., & Schafer, C. A. (1981). Phasic development in field crops II. Thermal and photoperiodic responses of spring wheat. *Field Crops Research*, 4, 269-283.
18. Lin, B. B. (2011). Resilience in agriculture through crop diversification: adaptive management for environmental change. *BioScience*, *61*(3), 183-193.
19. Atkinson, D., & Porter, J. R. (1996). Temperature, plant development and crop yields. *Trends in Plant Science*, *1*(4), 119-124.
20. Slafer, G. A. (1996). Differences in phasic development rate amongst wheat varieties independent of responses to photoperiod and vernalization. A viewpoint of the intrinsic earliness hypothesis. *The Journal of Agricultural Science*, 126(04), 403-419.
21. Slafer, G. A., &Savin, R. (1991). Developmental base temperature in different phenological phases of wheat (Triticum aestivum). *Journalof Experimental Botany*, 42(8), 1077-1082.
22. Del Pozo, A. H., García-Huidobro, J., Novoa, R., & Villaseca, S. (1987). Relationship of base temperature to development of spring wheat. *Experimental Agriculture*, 23(1), 21-30.
23. Meier, U. (1997). BBCH-Monograph. Growth stages of plants–Entwicklungsstadien von Pflanzen–Estadios de las plantas–Développement des Plantes. Berlin undWien: Blackwell Wissenschaftsverlag, 622.
24. Botarelli, L., Brunetti, A., Pasquini, A., &Linoni, F. (1999). Aspetti generali delle osservazioni agrofenologiche. Collana di Agrofenologia, MiPAF, PF Phenagri, Fenologia per l'Agricoltura. Vol, 1, 110.
25. Ventura, F., Traini, S., Gaspari, N., Rossi Pisa, P., Marletto, V., &Zinoni, F. (2006). La prima stazione agrofenologica italiana: installazione e risultati preliminari. *Rivista Italiana di Agrometereologia*. 11 (1), 41, 45.
26. Ventura F.; Marletto V.; Traini S.; Tomei F.; Botarelli L.; Rossi Pisa P. (2009). Validation of development models for winter cereals and maize with independent agrophnological observations in the BBCH scale, *RivistaItaliana di Agrometereologia*. 14(3), pp. 17 - 26.
27. Matzneller, P., Ventura, F., Gaspari, N., & Pisa, P. R. (2010). Analysis of climatic trends in data from the agrometeorological station of Bologna-Cadriano, Italy (1952–2007). *Climatic Change*, 100(3-4), 717-731.
28. Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. FAO, Rome, 300(9), D05109.
29. Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing mean model performance. *Climate Research*, 30(1), 79-82.

30. Jamieson, P. D., Porter, J. R., & Wilson, D. R. (1991). A test of the computer simulation model ARCWHEAT1 on wheat crops grown in New Zealand. *Field crops research*, 27(4), 337-350.
31. Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—a discussion of principles. J Hydrol, 10, 282-290.
32. Olesen, J. E., Børgesen, C. D., Elsgaard, L., Palosuo, T., Rötter, R. P., Skjelvåg, A. O., ... & Siebert, S. (2012). Changes in time of sowing, flowering and maturity of cereals in Europe under climate change. *Food Additives&Contaminants: Part A*, *29*(10), 1527-1542.
33. Di Paola, A., Caporaso, L., Di Paola, F., Bombelli, A., Vasenev, I., Nesterova, O. V., ... & Valentini, R. (2018). The expansion of wheat thermal suitability of Russia in response to climate change. *Land Use Policy*, *78*, 70-77.
34. Finckh, M. R., & Wolfe, M. S. (2006). Diversification strategies. In The Epidemiology of Plant Diseases (pp. 269-307). Springer, Dordrecht.
35. Altieri, M. A., & Nicholls, C. I. (2005). Agroecology and the search for a truly sustainable agriculture. United Nations Environmental Programme, Environmental Training Network for Latin America and the Caribbean.
36. Motzo, R., & Giunta, F. (2007). The effect of breeding on the phenology of Italian durum wheats: from landraces to modern cultivars. *EuropeanJournalofAgronomy*, *26*(4), 462-470.
37. Isidro, J., Álvaro, F., Royo, C., Villegas, D., Miralles, D. J., & García del Moral, L. F. (2011). Changes in duration of developmental phases of durum wheat caused by breeding in Spain and Italy during the 20th century and its impact on yield. *Annals of Botany*, *107*(8), 1355-1366.
38. Gallagher, J. N. (1979). Field studies of cereal leaf growth: I. Initiation and expansion in relation to temperature and ontogeny. *Journal of Experimental Botany*, *30*(4), 625-636.
39. McMaster, G. S., Wilhelm, W. W., & Morgan, J. A. (1992). Simulating winter wheat shoot apex phenology. *The Journal of Agricultural Science*, 119(1), 1-12.
40. Xue, Q., Weiss, A., &Baenziger, P. S. (2004). Predicting phenological development in winter wheat. *ClimateResearch*, 25(3), 243-252.
41. Donatelli, M., Stöckle, C., Ceotto, E., & Rinaldi, M. (1997). Evaluation of CropSyst for cropping systems at two locations of northern and southern Italy. *European Journal of Agronomy*, 6(1), 35-45.
42. Dettori, M., Cesaraccio, C., Motroni, A., Spano, D., & Duce, P. (2011). Using CERES-Wheat to simulate durum wheat production and phenology in Southern Sardinia, Italy. *Field Crops Reresearch*, 120(1), 179-188.
43. Palosuo, T., Kersebaum, K. C., Angulo, C., Hlavinka, P., Moriondo, M., Olesen, J. E., ... &Trnka, M. (2011). Simulation of winter wheat yield and its variability in different climates of Europe: a comparison of eight crop growth models. *European Journal of Agronomy*, 35(3), 103-114.
44. Cola, G., Pieri, L., Salvatorelli, F., & Ventura, F. (2016). Agro-phenological observation and modeling of cereals in Padana Plain in the period 2003-2012. *Italian Journal of Agrometeorology* - RivistaItaliana Di Agrometeorologia, 21(2), 5-14.
45. Soriano, J. M., Villegas, D., Aranzana, M. J., del Moral, L. F. G., &Royo, C. (2016). Genetic structure of modern durum wheat cultivars and Mediterranean landraces matches with their agronomic performance. *PloS one*, 11(8), e0160983.
46. Slafer, G. A., & Rawson, H. M. (1995b). Base and optimum temperatures vary with genotype and stage of development in wheat. *Plant, Cell & Environment*, *18*(6), 671-679.

SUPPLEMENTARY INFORMATION

Table S1. As Table 1 in the main text but showing results obtained by a simple TS model (not considering the effect of photoperiod).

Phenological phases	***n***	$\bar{O}$ **(std) [d]**	**MAE [d]**	**NMAE [%]**	**EF**	R^2	**p**
S-3L	22	55(23)	7.8	14	0.80	0.75	<0.01
3L-2N	23	135(35)	7.0	5	0.91	0.83	<0.01
2N-H	32	162(44)	4.5	2	0.98	0.96	<0.01
H-M	30	200 (40)	3	1	0.99	0.98	<0.01

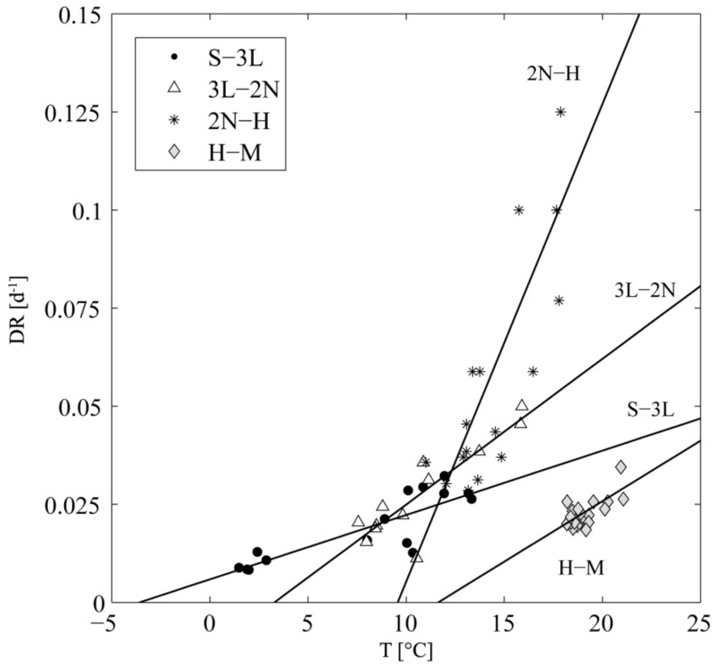

Figure S1. Comparison between DRs. Black lines: DRs; data points: data from the calibration dataset (symbols in legend). S-3L: from sowing to three leaves unfolded; 3L-2N: from three leaves unfolded to second node detectable; 2N-H: from second node detectable to beginning of heading; H-M: from beginning of heading to physiological maturity.

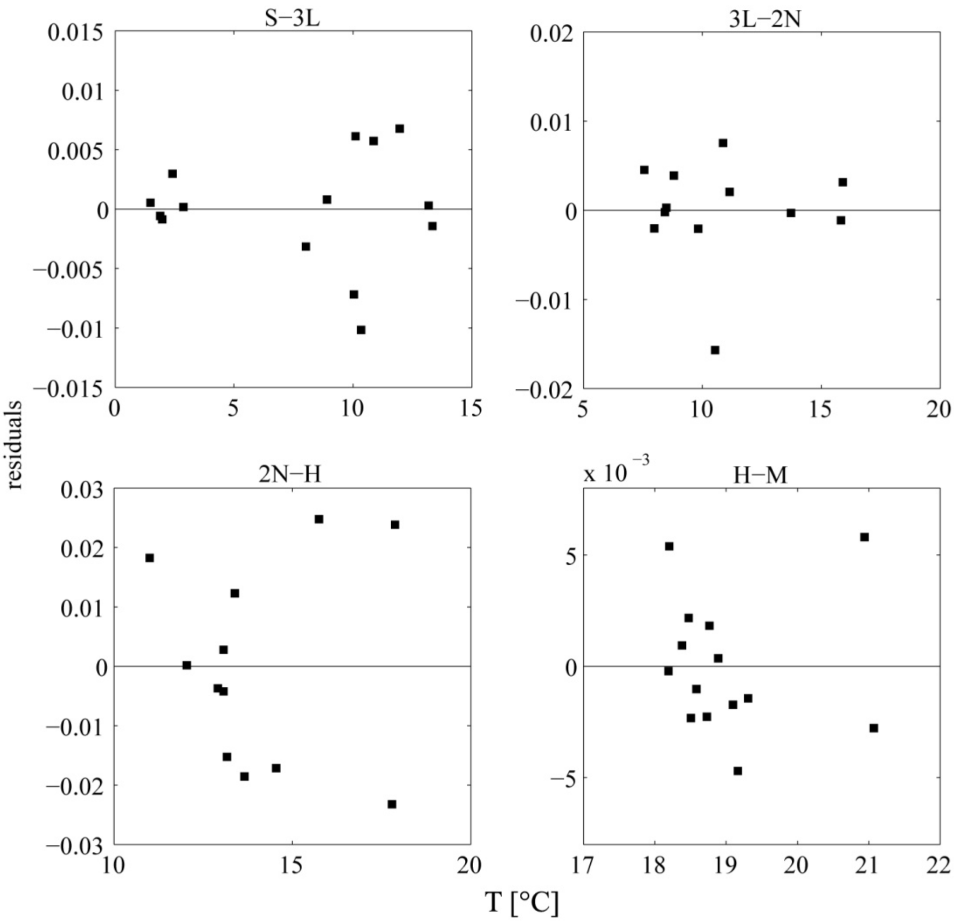

Figure S2. Residuals from simple linear regression considering only mean temperature. Dispersion of data reveals no biased relationships and almost homogenous variance.

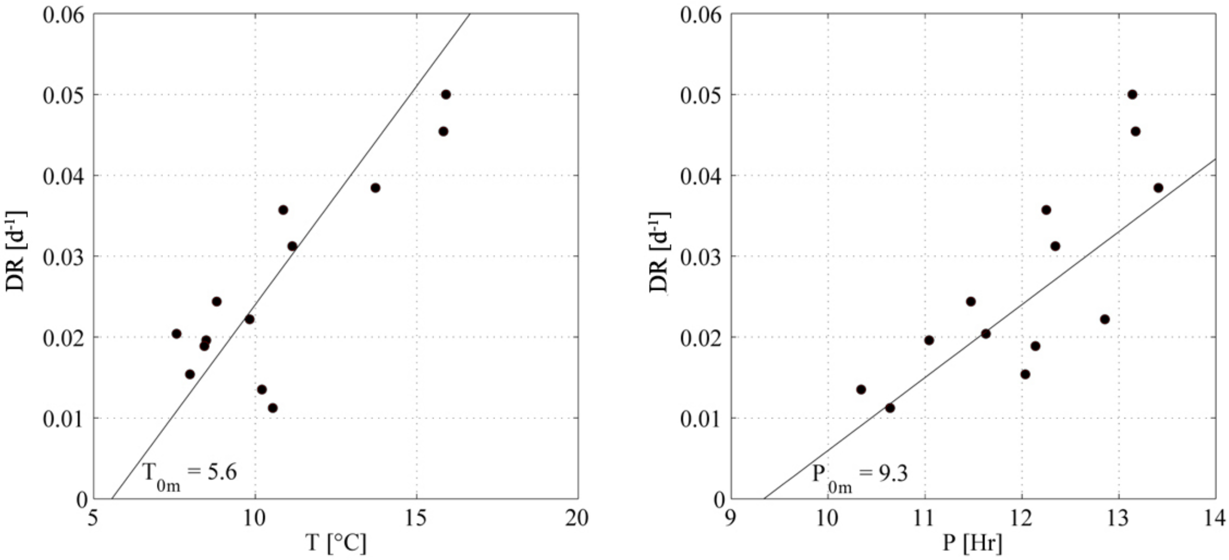

Figure S3. Multiple linear regression (as shown in the right panel of Fig. 3) projected on two-dimensional scatter plot. The intersection of the linear function with the abscissa return the threshold values for T_{0m} and P_{0m}.

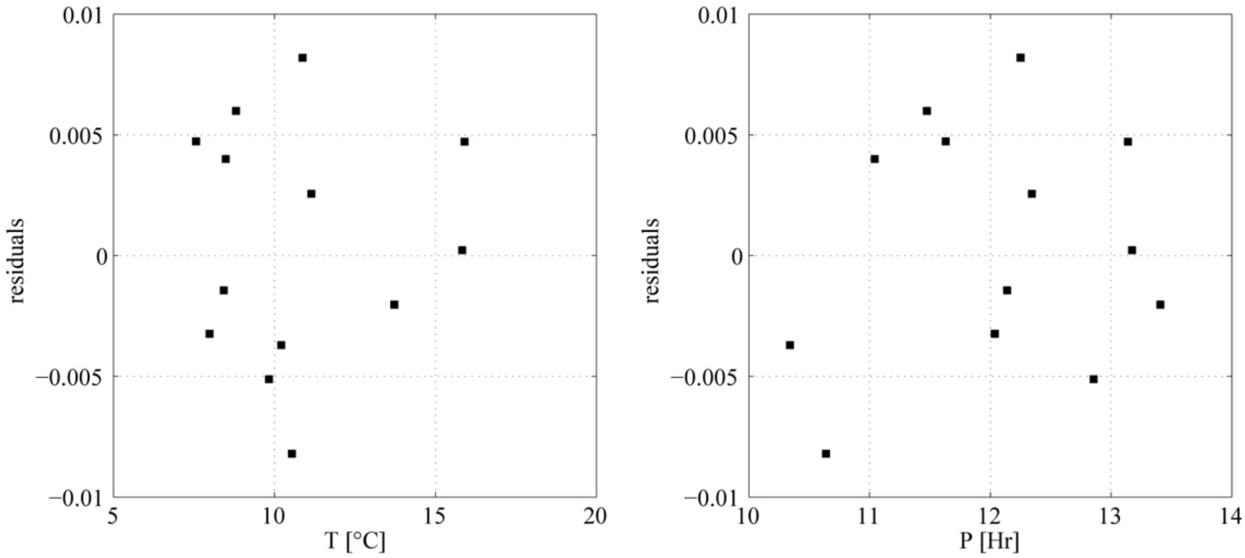

Figure S4. Residuals from multiple linear regression in 3L-2N. Dispersion of data reveals no biased relationships and almost homogenous variance.