Field Crops Research 133 (2012) 167–175

Contents lists available at SciVerse ScienceDirect

Field Crops Research

journal homepage: www.elsevier.com/locate/fcr

Improved modeling of grain number in winter wheat

Arne M. Ratjen*, Ulf Böttcher, Henning Kage

University of Kiel (Christian-Albrechts-Universität zu Kiel), Institute for Crop Science and Plant Breeding, Hermann-Rodewald-Str. 9, 24098 Kiel, Germany

ARTICLE INFO

Article history:
Received 18 March 2011
Received in revised form 4 April 2012
Accepted 5 April 2012

Keywords:
Grain number
Wheat
NNI
Photo thermal quotient
CERES Wheat
Growth curve
KNO
GPSM
Verhulst curves

ABSTRACT

Four models were compared for estimating grain number per square meter (GPSM) of winter wheat (*Triticum aestivum* L.). As a data-base for model comparison, values of GPSM were determined in a field trial in northern Germany ranging from 8.3 to 25 thousand grains per square meter under the influence of three years (2003/04 to 2005/06), five cultivars, and varying N supply (0–320 kg/ha). The comparison was repeated using a published independent dataset collected in the Netherlands in 1983 and 1984 (Wageningen dataset) with a cultivar differing from the German trial grown across sites and N treatments. Both datasets included several measurements of shoot dry weight (DM), shoot N concentration (cN) and stage of development (BBCH) during a vegetation period.

Simulations of phenological development (BBCH scale) were performed with a separate model and used for all four models. Input values of all models were obtained from experimental data, using fitted logistic growth curves to estimate DM, whereas cN was linearly interpolated. Three of the four models (M1–3) had been published before: M1 uses shoot dry weight at flowering (DM_{65}), M2 uses shoot dry weight increase between end of leaf growth and flowering (ΔDM_{39-65}), whereas M3 is a multiple regression with log transformed nitrogen nutrition index at anthesis (NNI_{60}) and average photothermal quotient from 45 days preceding anthesis (Q45) as explanatory variables. The fourth model (M4) was developed in this study based on the data observed in the German trial and considers the product of DM_{65}, NNI_{60} and Q45.

The relation between explanatory variables and GPSM did not vary greatly between the modern bread wheat cultivars of the German dataset, but there were considerable differences to the cultivar used in the Netherlands dataset. Thus, a genotype specific fit parameter (G) was added to the models and calibrated over each dataset. The Wageningen dataset was used for a *ceteris paribus* comparison between models and as validation of the new model. M4 shows best results for both datasets ($n = 45$), whereas the relative root mean square error (rRMSE) of simulated GPSM over all crops of the *ceteris paribus* comparison ($n = 9$) could be reduced to 8%, compared to 12–17% obtained from the existing approaches. The number of grains per unit shoot weight is influenced by NNI_{60} and Q45. This relation is considered by M4 and founded its improvement.

1. Introduction

A sufficiently accurate prediction of grain number per square meter (GPSM) is highly important for most wheat crop simulators, e.g. CERES-Wheat (Ritchie and Otter, 1985). According to the CERES-Wheat approach, the potential grain-filling rate for carbohydrates and protein is calculated for a single grain, which is then multiplied with the number of grains per plant to obtain the daily plant-grain weight increase. For this reason, potential sink size during grain growth is directly proportional to GPSM. A strong correlation between measured GPSM and yield has been reported by many authors (e.g. Midmore et al., 1984; Savin and Slafer, 1991; Fischer, 1993; Sayre et al., 1997; Gonzalez et al., 2003; Sinclair and Jamieson, 2006), even though, the role of GPSM as a sink limiting factor for yield is still controversial (Sinclair and Jamieson, 2006, 2008; Fischer, 2008). Nevertheless, the yield component GPSM is more closely related to yield than thousand kernel weight (Sayre et al., 1997; Duggan et al., 2000; Brancourt-Hulmel et al., 2003), and there is evidence that a larger grain number can increase radiation use efficiency as a result of a higher assimilate demand (Reynolds et al., 2005). In contrast, growing grains are also strong sinks for nitrogen (N) and trigger N remobilization from the vegetative organs, which decreases canopy, photosynthesis, and accelerates leaf senescence (Bertheloot et al., 2008a,b). Therefore, GPSM may prove to be vital for modeling N distribution and grain quality, because increasing thousand kernel weights caused by low GPSM values can induce dilution effects resulting in low protein concentrations (Kibite and Evans, 1984) and thereby results in reduced baking quality.

For CERES Wheat Otter-Nacke et al. (1986) determined an rRMSE (rRMSE = RMSE/Mean) for simulated GPSM of over 35%, which was later confirmed by Moreno-Sotomayor and Weiss

* Corresponding author. Tel.: +49 431 880 3474; fax: +49 431 880 1396.
E-mail address: ratjen@pflanzenbau.uni-kiel.de (A.M. Ratjen).

0378-4290/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.fcr.2012.04.002

(2004). The predictive quality for a model variable like GPSM generally depends on two main influences. The first is the prediction quality of the simulated input values like biomass, growth, or nutrition status. The second is the algorithm itself. The latter should be a robust description of the empirical relations between explanatory variables and GPSM.

There are some recent examples which indicate that improving models is still worthwhile (e.g. Demotes-Mainard and Jeuffroy, 2001; Moreno-Sotomayor and Weiss, 2004). GPSM was calculated as a function of total shoot biomass or biomass of single organs, as well as biomass growth between different phenological stages. Vos (1981) assumed that GPSM can empirically be correlated to total above ground dry matter at anthesis (DM_{65}), whereas in CERES-Wheat and APSIM (Keating et al., 2003) grain number per plant is calculated as a linear function of stem dry matter at anthesis multiplied with a cultivar specific coefficient (G). However, in the current version of Ceres-Wheat (DSSAT v 4.0.2.0) total shoot dry matter is used instead of stem dry matter for the calculation of GPSM.

Fischer (1985) noted that GPSM could be considered as a linear function of spike dry weight at anthesis. This concept was incorporated into different crop models like ARCWHEAT1, AFRCWHEAT2, and SIRIUS (Moreno-Sotomayor and Weiss, 2004). Moreno-Sotomayor and Weiss (2004) followed this concept by deriving spike biomass as a function of stem growth between flag leaf sheath opening and mid anthesis. Although this approach showed a better prediction for GPSM compared to the old CERES-Wheat approach (even without a cultivar specific calibration) the rRMSE was still higher than 20%. They further pointed out, that the goodness of prediction could be enhanced by assuming grain set reduction due to high temperatures. The STICS model (Brisson and Ripoche, 1998) uses a similar concept as Moreno-Sotomayor and Weiss and calculates GPSM as a linear function of the mean shoot growth rate over a range of 30 days before grain filling.

Fischer (1985) found that the relatively short period between penultimate leaf emergence and mid anthesis is more critical for grain number than total growth period. As the number of grains per spike is proportional to the spike biomass, it increases in response to higher levels of radiation during spike growth. In contrast, high temperature decreases the number of grains per spike through reduced DM accumulation, caused by accelerated development and increased respiration during booting. Hence, the ratio of mean daily incident radiation to mean daily temperature known as the photothermal quotient has also been suggested as a good estimator for the number of competent florets and grains (Midmore et al., 1984; Fischer, 1985). The relationship between GPSM and weather variables, however, can be affected by N supply. Jeuffroy and Bouchard (1999) quantified the influence of N deficiency on the relative grain number (RGN) compared to the N fertilized control. Later, Demotes-Mainard and Jeuffroy (2001) combined nitrogen nutrition index (NNI, derived from critical N content, Justes et al., 1994) at anthesis and photothermal quotient over the 45 days preceding anthesis (Q45), to estimate GPSM using a multiple regression equation.

The objective of this work was to analyze the influence of different explanatory variables on GPSM of winter wheat crops and to compare the predictive quality of models relying on one or combinations of these variables.

2. Material and methods

2.1. Datasets

Model development and testing were based on data from field experiments over three years with destructive sequential harvests near Kiel, Germany and a well documented dataset from field experiments carried out in the Netherlands near Wageningen (Groot, 1987). Some parts of the Wageningen dataset (Eest 84, Bouwing 83 and 84) could not be used because the thousand kernel mass (TKM) was not published and thus GPSM could not be calculated. Crop development was characterized by the decimal code according to Zadoks et al. (1974), which equals the BBCH code (Meier, 2001). All datasets are characterized by experimental location, years and sampling dates (Table 1) description of N application rates, averages of explanatory variables, and measured grain numbers per square meter (GPSM) (Table 2).

2.2. Site conditions

2.2.1. Wageningen dataset (NL 1983–1984)

Two experiments were located in the newly reclaimed polders on the experimental farms 'the Eest' (silty loam) and 'PAGV' (silty loam); characterized by a total annual rainfall average of ~765 mm, from which ~375 mm are received during spring and summer, and ~390 mm during fall and winter (for further information see Groot, 1987; Van Delden et al., 2000; Jacobs et al., 2010).

2.2.2. Hohenschulen dataset (GER 2004–2006)

Experiments were conducted at the Hohenschulen experimental farm of the University of Kiel, located in NW Germany about 15 km west of Kiel (federal state of Schleswig-Holstein); characterized by a total annual rainfall average of ~750 mm, from which ~400 mm are received during April to September (main growing season), and ~350 mm during October–March. The soil type is pseudogleyic luvisol, with a soil texture varying from sandy loam to clayey loam.

2.3. Field trials and experimental design

The experimental design of the Wageningen dataset included three N fertilizer treatments with sowing dates between October 19th and October 25th.

In each experiment of the Hohenschulen dataset, the plot size was 3 m × 12 m. The 2004 experiment was one factorial (cultivar *Ritmo*) with four N levels 0, 120 kg N ha^{-1} (40/40/40), 192 kg N ha^{-1} (30/50/72/40), and 320 kg N ha^{-1} (120/120/80). In 2005 and 2006, four Cultivars were grown (05: *Cubus*, *Ritmo*, *Hybred* and *Dekan*; 06: *Cubus*, *Ritmo*, *Tommi* and *Dekan*) with sowing dates between September 15th and October 6th and four (mineral) N levels 0, 80 kg N ha^{-1} (40/40/0), 160 kg N ha^{-1} (80/40/40) and 240 kg N ha^{-1} (80/80/80).

2.4. Measured variables and sampling dates

In this study, the explanatory variables used were shoot dry matter (DM, g m^{-2}), total shoot N concentration (cN, %), and development stages. Crop development was characterized with the BBCH code. Table 1 lists the sampling dates used for this study.

At each sampling date, 0.25 m^2 of each plot was harvested. In 2004, each data point represents an average of eight replications harvested, in 2005 and 2006, four. To avoid boundary effects, sampling areas did not border each other within a plot. The plants were separated into leaf blades, stem and spike. After drying and weighing separately, the fractions were analyzed for cN using near infrared spectroscopy (NIRS) analysis (NIRS Systems 5000 scanning monochromator: FOSS GmbH, Rellingen, Germany; WINISI software package: Infrasoft International, Port Matilda, PA, USA). Total above-ground dry matter and total N were calculated by summing up the fractions. GPSM from each plot was calculated by dividing grain yield by thousand kernel mass (TKM).

Table 1
Dates and BBCH stages for intermediate destructive samplings at the experimental sites Wageningen (EEST, PAGV) and Hohenschulen (HS).

Site	Sampling dates	BBCH
EEST 1983	09/02/1983[b], 02/03/1983[a], 30/03/1983[a], 20/04/1983[a], 11/05/1983[a], 25/05/1983, 15/06/1983, 06/07/1983, 20/07/1983, 03/08/1983	22-Harvest
PVAG 1983	08/02/1983[b], 01/03/1983[a], 29/03/1983[a], 19/04/1983[a], 10/05/1983[a], 24/05/1983, 14/06/1983, 05/07/1983, 19/07/1983, 02/08/1983	21-Harvest
PVAG 1984	14/02/1984[b], 13/03/1984[a], 03/04/1984[a], 24/04/1984[a], 08/05/1984[a], 28/05/1984, 19/06/1984, 03/07/1984, 17/07/1984, 08/08/1984, 20/08/1984	21-Harvest
HS 2004	17/11/2003, 17/03/2004, 29/03/2004, 14/04/2004, 27/04/2004, 10/05/2004, 25/05/2004, 01/06/2004, 21/06/2004, 14/07/2004, 10/08/2004	22–94
HS 2005	20/04/2005, 02/05/2005, 17/05/2005, 31/05/2005, 13/06/2005, 05/07/2005, 02/08/2005	29–89
HS 2006	11/05/2006, 18/05/2006, 30/05/2006, 18/06/2006, 05/07/2006, 26/07/2006	30–90

[a] Lowest N treatment only.
[b] Assumed as equal for all treatments (no N was applied).

Table 2
N application rates, mean values of explanatory variables (NNI60 = N nutrition status of wheat crops at anthesis, Q45 = the photo thermal quotient 45 days preceding anthesis, DM65 = total shoot dry matter per square meter at anthesis, ΔDM39-65 = accumulated shoot biomass between end of leaf growth and mid anthesis), and measured grain numbers per square meter (GPSM) for the Hohenschulen (HS) and Wageningen (EEST, PAGV) dataset. Corresponding BBCH values are represented as subscripts.

Site	n[a]	Year	Applied N [kg N ha^{-1}]	GPSM [grain m^{-2}]	DM$_{65}$ [g m^{-2}]	ΔDM$_{39–65}$ [g m^{-2}]	NNI$_{60}$	Q45 [MJ (m^2 °Cd)$^{-1}$]
EEST	1	1983	0	14,072	734.33	355.64	0.398	0.618
EEST	1	1983	60	25,537	945.24	487.28	0.734	0.618
EEST	1	1983	160	27,158	1011.37	600.47	0.929	0.618
PAGV	1	1983	80	19,077	1074.75	524.53	0.548	0.623
PAGV	1	1983	140	24,233	1170.83	596.35	0.736	0.623
PAGV	1	1983	240	25,035	1256.57	663.08	0.941	0.623
PAGV	1	1984	80	19,662	1162.51	569.88	0.517	0.534
PAGV	1	1984	180	21,822	1212.67	586.25	0.713	0.534
PAGV	1	1984	240	22,738	1239.04	589.81	0.829	0.534
HS	1	2004	0	16,344	1113.91	546.83	0.473	0.621
HS	1	2004	120	24,363	1416.44	787.11	0.849	0.621
HS	1	2004	192	24,308	1595.43	796.33	1.056	0.621
HS	1	2004	320	23,822	1442.63	666.85	1.125	0.621
HS	4	2005	0	9758	584.55	237.12	0.397	0.628
HS	4	2005	80	17,453	1229.61	508.18	0.616	0.628
HS	4	2005	160	21,618	1349.26	601.38	0.833	0.628
HS	4	2005	240	23,391	1346.12	616.32	1.016	0.628
HS	4	2006	0	11,373	763.34	328.26	0.361	0.721
HS	4	2006	80	18,477	1185.51	578.95	0.628	0.721
HS	4	2006	160	20,131	1211.43	604.85	0.868	0.721
HS	4	2006	240	21,225	1198.51	608.51	1.030	0.721

[a] Number of cultivars.

2.5. Existing models for GPSM estimation used in this study

Four models were tested for estimating GPSM of winter wheat. Model 1 (M1) originally contained a genotype specific parameter (G_x) which was also added to models 2–4 (M2–4), to ensure better comparison. Subscripts of explanatory variables correspond to BBCH values.

M1 (CERES-Wheat v. 4):

GPSM is calculated as a function of total shoot dry matter per square meter at anthesis (CERES-Wheat (DSSAT) version v 4.0.2.0):

$$\mathrm{GPSM} = G_1 \times \mathrm{DM}_{65} \tag{1}$$

M2 (Moreno-Sotomayor and Weiss, 2004):

This approach calculates GPSM as a function of accumulated shoot biomass between end of leaf growth and mid anthesis (BBCH 39–65), while G_2 is the number of grains per $\Delta\mathrm{DM}_{39-65}$:

$$\mathrm{GPSM} = G_2 \times \Delta\mathrm{DM}_{39-65} \tag{2}$$

In the original version of this model, the stem growth rate according to CERES-Wheat v.3 (stem includes spike until the beginning of grain filling phase) was taken instead of shoot growth to model spike biomass. Assuming that leaf growth is completed before this period, shoot growth can be considered as equal to stem growth.

M3 (Demotes-Mainard and Jeuffroy, 2001):

This approach describes GPSM as a multiple regression including the logarithm of NNI at BBCH 60 and the photothermal quotient [MJ (m^2 day degree)$^{-1}$] over the 45 days preceding anthesis (Q45):

$$\mathrm{GPSM} = G_3 \times (-4091.8 + 12160 \times \ln(\mathrm{NNI}_{60}) + 41889 \times Q45) \tag{3}$$

For calculation of Q45, a base temperature of zero and photosynthetically active radiation (PAR) are used. PAR was calculated as 48% of the global radiation, according to Demotes-Mainard and Jeuffroy (2001). NNI was calculated as the ratio between current cN (interpolated between two measurements) and the critical N concentration (N_{crit}) as percentage of shoot dry matter (N_{crit}), determined from the equation proposed by Justes et al. (1994).

For all values of DM greater than 155 g/m^2, N_{crit} (%) was calculated as follows: $N_{\mathrm{crit}} = 5.35 \times (\mathrm{DM}/100)^{-0.442}$. For lower values of DM, N_{crit} was assumed to be 4.4%.

2.6. Interpolation of field data

Logistic growth curves (Verhulst curves) were fitted to the observed data of shoot dry matter to obtain the daily shoot growth rate. For the Wageningen dataset, the first sample date for the lowest N level started in February (BBCH 13-21), whereas the other N treatments started in May (BBCH 37-39). For getting a more robust interpolation of DM during early growth, the first sample dates (before the first N application) of the lowest N treatment were

Table 3
Slope, intercept and r^2 of the linear regression between measured and predicted (considered independent) shoot dry matter (DM) and BBCH values as well as number of observations (n). Additional the RMSE between observed and interpolated DM values is given (3 node until watery ripe stage, BBCH 33–70).

Dataset	Parameter	n	Slope	Intercept	r^2	$RMSE_{33–70}$
Wageningen (1983–1984)	DM [$g\,m^{-2}$]	69	1.01	1.07	0.99	57
	BBCH	67	1.01	−0.54	0.99	–
Hohenschulen (2004–2006)	DM [$g\,m^{-2}$]	251	0.99	11.95	0.98	96
	BBCH	292	0.94	2.7	0.97	–

used as additional data points for the second and third N levels, as well. The daily change of cN was determined by linear interpolation between two subsequent sampling dates.

The explanatory variables (DM_{65}, $\Delta DM_{39–65}$, NNI_{60} and Q45) were derived from interpolated measurements and BBCH development stages. The slope and r^2 of the linear regression between measured and predicted values (Table 3) is near one for DM (logistic growth) and BBCH (discrete value related daily from phenological sub-model). Additionally, the RMSE between measured and interpolated DM values (BBCH 33–70) is shown, but not for BBCH, because these are discrete values.

The phenological model was previously published by Johnen et al. (2006) and considers thermal time, vernalization and day length. The simulated development stages and growth curves were based on thermal time (°Cd) at a base temperature of zero. The derived explanatory variables were calculated daily.

2.7. Parameter estimation, statistical evaluation

Curves and models for the estimation of GPSM were implemented within the HUME modeling environment (Kage and Stützel, 1999) which supports parameter estimation based on the Levenberg–Marquardt algorithm (Marquardt, 1963).

The ratios $GPSM/DM_{65}$, $GPSM/\Delta DM_{39–65}$ and $GPSM/NNI_{60}$ were calculated for detection of cultivar related differences, using one-way analysis of variance (ANOVA) followed by pairwise multiple comparison procedures (Bonferroni's t-test) in sigma plot 11.0 (Sigma, USA), at which $GPSM/DM_{65}$ was logarithmised to provide normal distributed values. Differences were considered significant at $P<0.05$.

There were no significant differences between the cultivars of the Hohenschulen dataset, but the ratio of GPSM and explanatory variables were higher with the cultivar *Arminda* used for Wageningen dataset (Table 4), therefore the genotype specific factor (G_x) was estimated for each dataset and for each model by minimizing the RMSE of GPSM. The exponential parameter of M4 was calibrated using the Hohenschulen dataset only. The Wageningen dataset was used for a *ceteris paribus* comparison, with one fit parameter for each model. Statistical evaluation was done using bias (difference between the means of simulation and observed values), root mean square error (RMSE), and the relative root mean square error (rRMSE), as well as Pearson's correlation (r) and coefficient of determination (r^2).

Table 5
Pearson's correlation (r) between the characteristics of log transformed NNI ($lnNNI_{60}$), shoot dry matter per m^2 (DM_{65}) and shoot growth ($\Delta DM_{39–65}$), as well as grain number per m^2 at harvest (GPSM) for Wageningen ($n=9$) and Hohenschulen ($n=36$).

	DM_{65}	$lnNNI_{60}$	$\Delta DM_{39–65}$	GPSM
Wageningen (1983–1984)				
DM_{65}	1	0.598	0.904	0.465
$lnNNI_{60}$		1	0.598	0.933
$\Delta DM_{39–65}$			1	0.723
GPSM				1
Hohenschulen (2004–2006)				
DM_{65}	1	0.855	0.937	0.925
$lnNNI_{60}$		1	0.858	0.944
$\Delta DM_{39–65}$			1	0.932
GPSM				1

NNI = N nutrition status of wheat crops; DM = total shoot dry matter per square meter; ΔDM = accumulated shoot biomass, corresponding BBCH values are represented as subscripts.

3. Results

3.1. Explanatory variables, GPSM, and their relationship

Regarding the Hohenschulen dataset, DM_{65}, logarithmised NNI_{60} ($lnNNI_{60}$, see M3/Eq. (3)) and $\Delta DM_{39–65}$ are similarly strongly correlated to GPSM (Table 5), while DM_{65} and $\Delta DM_{39–65}$ show a weaker correlation with $lnNNI_{60}$ than with GPSM. Q45 is not shown in this context because of poor variation (Table 2). For the Wageningen dataset, the correlation between GPSM and both DM_{65} and $\Delta DM_{39–65}$ is weaker compared to the Hohenschulen dataset, but the correlation of $lnNNI_{60}$ is comparably strong.

On average, GPSM values are higher in the Wageningen dataset, mainly due to a smaller averaged TKM (33 compared to 45 g, not shown). The mean standard deviation of the observed GPSM values between single replicates of the same treatment is 1470 for the Hohenschulen dataset (unknown for Wageningen dataset).

In Fig. 1, the linear regression between GPSM and different plant parameters is shown. In both datasets, the GPSM vs. NNI_{60} fits very well with a non-linear regression (Fig. 1C and G), whereas the $lnNNI_{60}$ shows a linear relation (Fig. 1D and H).

Regarding the Hohenschulen dataset, the linear regression between GPSM and both DM_{65} and $\Delta DM_{39–65}$ (Fig. 1A and B) fits well with the lowest N treatments, but worse at higher N levels combined with tendencies of bias for specific N levels. They show

Table 4
Cultivar specific ratios between grains per square meter (GPSM) and different crop parameters (NNI_{60} = N nutrition status of wheat crops at anthesis, Q45 = the photo thermal quotient 45 days preceding anthesis, DM_{65} = total shoot dry matter per square meter at anthesis, $\Delta DM_{39–65}$ = accumulated shoot biomass between end of leaf growth and mid anthesis).

	n	Year	$GPSM/DM_{65}$ [grains g^{-1}]	$GPSM/\Delta DM_{39–65}$ [grains g^{-1}]	$GPSM/NNI_{60}$ [grains $(NNI\,m^2)^{-1}$]
Arminda	9	1983–1984	20.52^{A}	40.25^{A}	$32{,}199^{A}$
Cubus	8	2005–2006	15.58^{B}	36.55^{AB}	$25{,}140^{B}$
Dekan	8	2005–2006	15.90^{B}	36.75^{AB}	$25{,}717^{AB}$
Hybred	4	2005	16.77^{AB}	38.00^{AB}	$26{,}186^{AB}$
Ritmo	12	2004–2006	16.44^{AB}	33.35^{B}	$26{,}792^{AB}$
Tommi	4	2006	16.45^{AB}	33.92^{AB}	$25{,}500^{AB}$

Significant differences are indicated by different capital characters, respectively ($P<0.05$). Corresponding BBCH values are represented as subscripts.

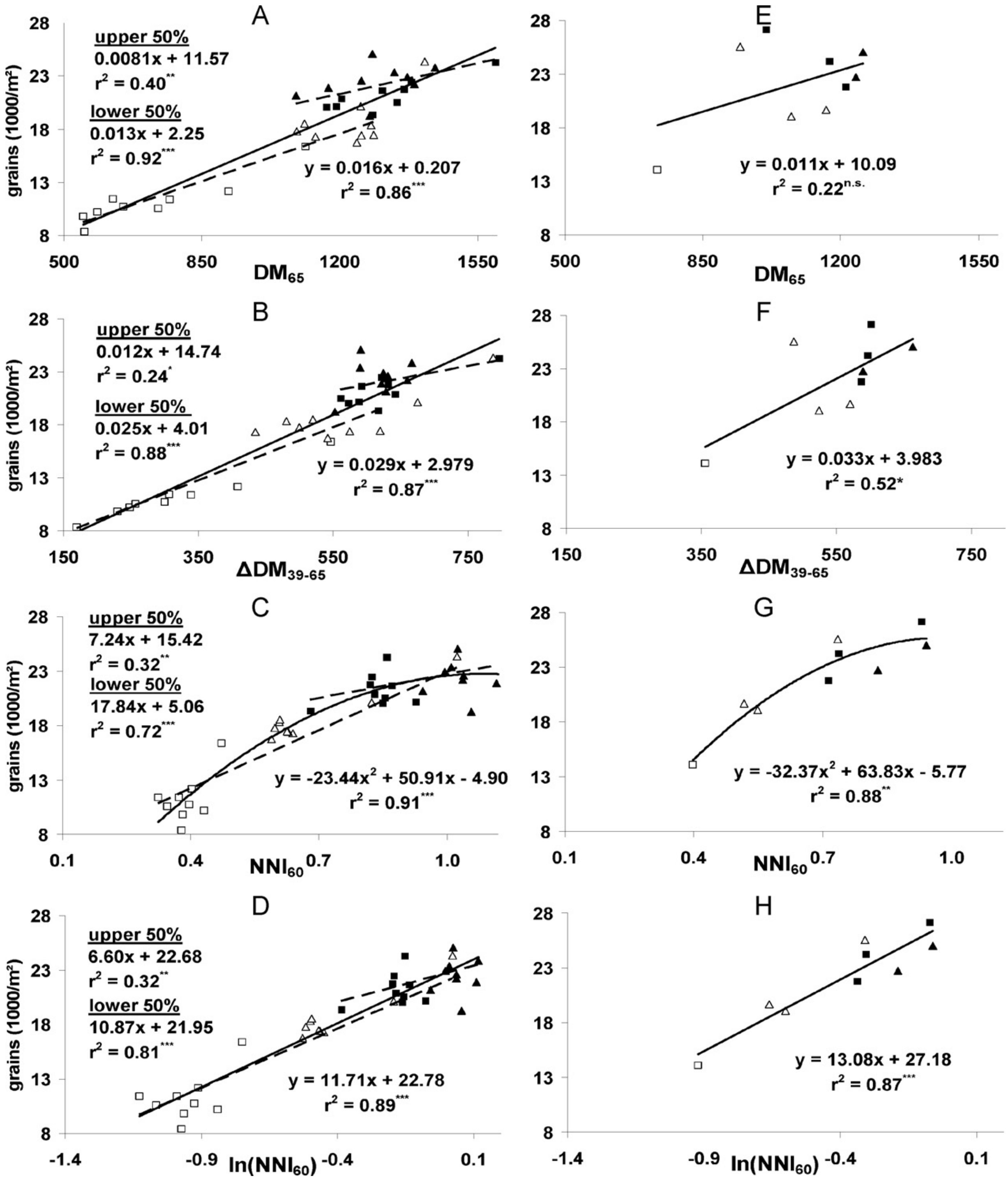

Fig. 1. Relation between different crop parameters (NNI_{60} = N nutrition status of wheat crops at anthesis; Q45 = the photo thermal quotient 45 days preceding anthesis; DM_{65} = total shoot dry matter per square meter at anthesis; ΔDM_{39-65} = accumulated shoot biomass between end of leaf growth and mid anthesis) and GPSM for the Hohenschulen (A)–(D) and Wageningen (E)–(H) dataset (–––), as well as the linear regression for upper and lower 50% quantiles of GPSM (– –, Hohenschulen only). N treatments: (□) 0, (△) 60–120 kg N ha^{-1}, (■) 140–192 kg N ha^{-1} and (▲) 240–320 kg N ha^{-1}. Corresponding BBCH values are represented as subscripts. * Significance at $P < 0.05$., ** Significance at $P < 0.01$. *** Significance at $P < 0.001$.

some clumping at high values of GPSM. In both datasets, GPSM vs. ΔDM_{39-65} shows a GPSM intercept ranging between 3 and 4 thousand grains (Fig. 1B and F). Treatments with no N supply are predicted well by the linear regression of GPSM and both DM_{65} and ΔDM_{39-65}, whereas NNI_{60} is less sensitive (Fig. 1A–D). For the Wageningen dataset, the relation between DM_{65} and GPSM (Fig. 1E) is rather poor.

For the Hohenschulen dataset, the GPSM values were divided in upper and lower 50% quantiles. The r^2 between different explanatory variables and GPSM differs for single quantiles of GPSM. For all explanatory variables, the r^2 in the upper 50% quantile is drastically lower compared to the lower 50% quantile.

Five different cultivars were involved in the Hohenschulen dataset, but there were no significant cultivar specific differences found, in contrast to the cultivar *Arminda* used in the Wageningen dataset (Table 4). Thus, a differentiation between the modern cultivars was not necessary and the *G* coefficients (Table 6) are the result of the general regression for each relationship.

Table 6
Values of the genetic coefficients and statistical parameters of predictive quality for all models (M1–4). The additional exponential fit parameter of M4 (Eq. (4)) was parameterized on the Hohenschulen dataset, whereas the Wageningen dataset was used for the validation under *ceteris paribus* condition with only one fit parameter for each model.

	G_1 [grains g^{-1}]	G_2 [grains g^{-1}]	G_3[a]	G_4[b]
Hohenschulen (n = 36)	16.133	34.205	0.9512	149.153
Wageningen (n = 9)	20.1293	39.919	1.3565	193.476
Hohenschulen	M1	M2	M3	M4
Number of fit parameter	1	1	1	2
SE of estimate[c] (adj.) [grains m^{-2}]	1834	1939	2562	1342
RMSE[d] [grains m^{-2}]	1808	1912	2513	1301
r^2 (adj.)[e]	0.86	0.84	0.72	0.92
Wageningen	M1	M2	M3	M4
Number of fit parameter	1	1	1	1
SE of estimate (adj.) [grains m^{-2}]	3900	2863	2720	1919
RMSE [grains m^{-2}]	3677	2700	2565	1809
r^2 (adj.)	0.07	0.50	0.55	0.77

[a] Non-dimensional factor to adapt the original model output.
[b] Grains per $\ln(DM_{65} \times NNI_{60} \times Q45)^{2.6233}$.
[c] Standard error of estimate.
[d] RMSE: root mean square error.
[e] Goodness of fit (not 1:1 plot).

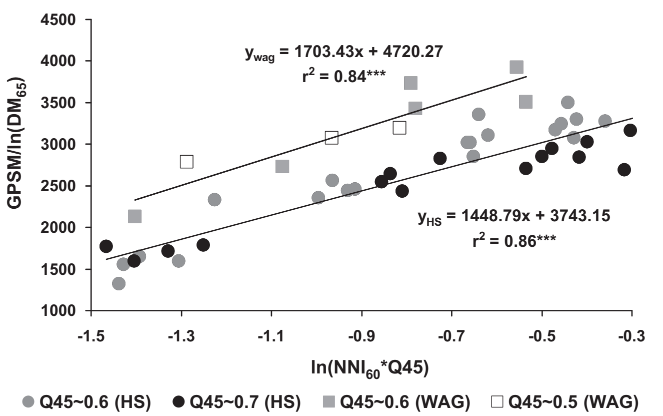

Fig. 2. The ratio of GPSM [grains m^{-2}] divided by the log transformed DM_{65} [g m^{-2}] and the log transformed product of NNI_{60} (N nutrition status of wheat crops at anthesis) and the photo thermal quotient 45 days preceding anthesis (Q45 [MJ (m^2 °Cd)$^{-1}$]), for both datasets (HS = Hohenschulen; WAG = Wageningen). The colors signify the level of Q45: 0.534 (white), 0.618–0.628 (gray), 0.712 (black). *** Significance at P < 0.001.

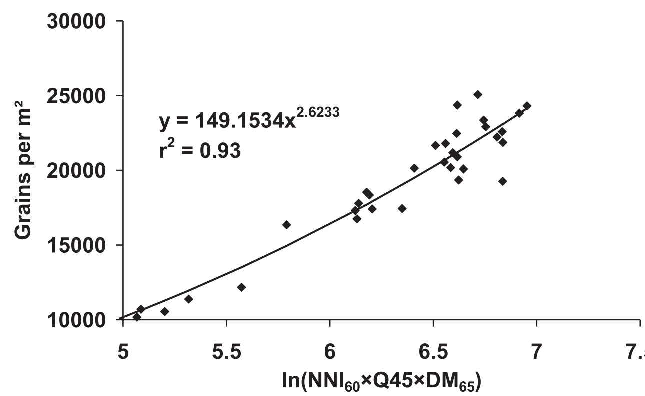

Fig. 3. Relationship between the log transformed product of different plant parameters (NNI_{60} = N nutrition status of wheat crops at anthesis; Q45 = the photo thermal quotient 45 days preceding anthesis; DM_{65} = total shoot dry matter per square meter at anthesis) and grain number per square meter (GPSM) of wheat crops from field experiments with different N fertilization (Hohenschulen dataset).

A non-linear influence of NNI_{60} and Q45 on number of grains per gram shoot was found. Fig. 2 illustrates the relationship between the product of NNI_{60} times Q45 and the number of grains per g shoot. The explanatory variables (DM_{65}, NNI_{60}, Q45) are log transformed and a similar strong and significant relationship was found between the ratio of GPSM/ln(DM_{65}) and the log transformed product of NNI_{60} and Q45 for both datasets.

3.2. The new approach (M4) and its calibration

M4 (newly developed):
The fourth model is a synthesis of M1 and M3 (Eqs. (1) and (3)) and follows the assumption that GPSM is a function of DM_{65}, NNI_{60}, and Q45 (Fig. 2):

$$GPSM = G_4 \times \ln(DM_{65} \times NNI_{60} \times Q45)^a \quad (4)$$

The reason for the log transformation as done in M4 relies on the fact that the response of GPSM to involved plant parameters decreases (Fig. 1), as well as on the aforementioned non-linear influence of NNI_{60} and Q45 on number of grains per gram shoot. However, NNI_{60} is log transformed in M3 as well, in order to linearize the influence of NNI_{60} on GPSM. The relation between $\ln(NNI_{60} \times Q45 \times DM_{65})$ and GPSM is almost linear (Fig. 3). Nonetheless, a power function was used instead of a linear regression, because it should be an a priori requirement that the estimating function is approaching the origin when explanatory variables or the resultant product, respectively, become very low. The contrast here is between explanatory variables that are summed up (Eq. (3)) with considerable intercept and variables that are multiplied (Eqs. (1), (2) and (4)).

Across the Hohenschulen dataset, M4 was adjusted with two fit parameters: 'G_4' and the exponential parameter 'a' (Fig. 3). Only G_4 was fitted to the Wageningen dataset, whereas the 'a' parameter has been taken as estimated from the Hohenschulen dataset to assure a *ceteris paribus* comparison with only one fit parameter for each model.

3.3. Model comparison

The adjusted standard error of estimation and the adjusted r^2 differ between models (Table 6), with M4 showing the best results for both datasets.

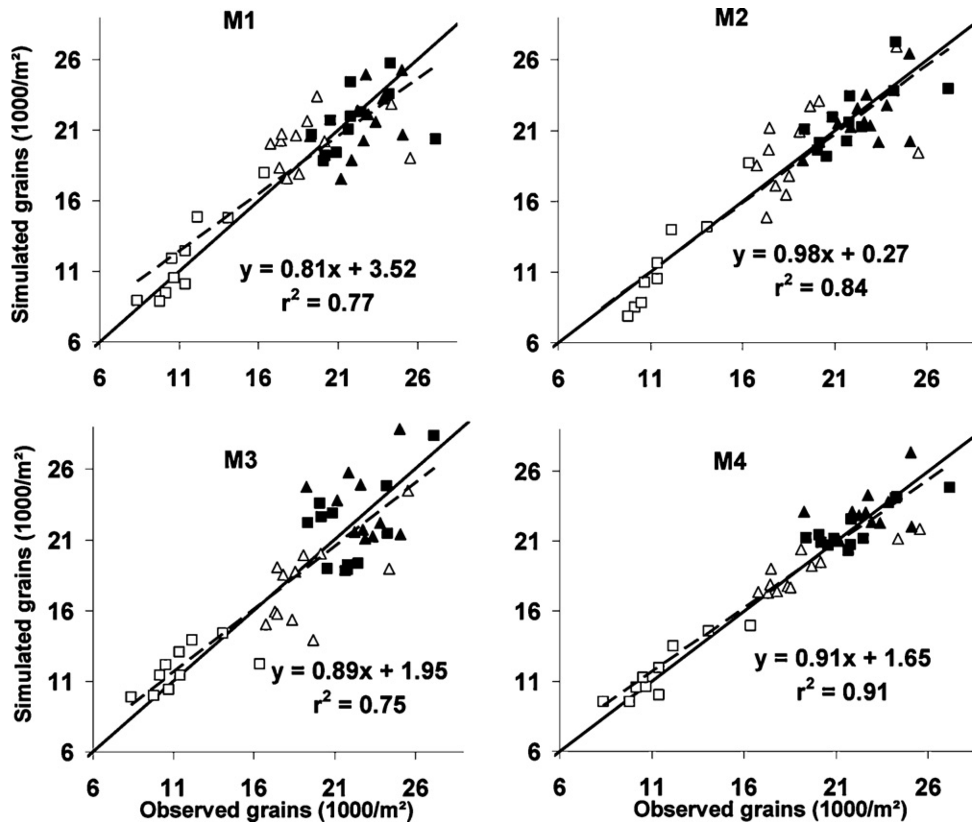

Fig. 4. Simulated vs. observed grain number ($n = 45$) for both data sets (N treatments: (□) 0, (△) 60–120 kg N ha^{-1}, (■) 140–192 kg N ha^{-1} and (▲) 240–320 kg N ha^{-1}) obtained from different model approaches (M1–4) and corresponding linear regression (– –) equations as well as the 1:1 line (–––). For further statistical parameters see Table 6.

M3 shows an almost comparable RMSE in both datasets, while the RMSE of M2, and even more M1, is much higher for the Wageningen dataset, compared with the Hohenschulen results.

Comparing the 1:1 plots of simulated vs. observed GPSM (both datasets), M1 tends to overestimate the second and underestimate the highest N levels (Fig. 4), while the slope differs substantially from one (assuming simulated values as independent). In contrast to M1, the slope of the linear regression is close to one with M2, whereas M3 and M4 are in between. The coefficient of determination is highest for M4 followed by M2, M1, and M3.

The absolute annual bias, defined as the absolute difference of means between simulated and observed grains across years; was lowest for M4 compared with the other models (not shown).

4. Discussion

4.1. Field data interpolation and accuracy of GPSM measurements

The influence of different explanatory variables on GPSM was analyzed and the predictive quality of model approaches relying on one or combinations of these variables was compared.

The evaluation of a model approach implemented in a dynamic crop model can be difficult because the predictive quality of the approach depends on the predictive quality of involved explanatory variables simulated by the crop model it is implemented in. On the other hand, experimental data of explanatory variables are mostly not available in a satisfactorily high resolution and they are often not determined for the specific development stages needed in the algorithms. We present a method comparing the different predictive approaches for GPSM, ensuring a close link to experimental field data, thereby avoiding dependencies on an underlying crop model; to our knowledge, such a strict comparison has not yet been conducted.

Nevertheless, measurement errors of the target value used for the model calibration, as well as the uncertainty of ingoing explanatory variables, potentially limit the precision of any algorithm. The average standard deviation for GPSM between single replicates of the treatments within the Hohenschulen dataset (unknown for the Wageningen dataset) amounts to 1470 grains m^{-2}. In this study, the explanatory variables were derived from interpolated field data, and each data point represents the average of at least four replications. Between measured and interpolated DM (during BBCH 33 and 70), we found an RMSE of 96 g/m^{-2} and 57 g/m^{-2} for the Hohenschulen and Wageningen dataset, respectively. Keeping in mind the average observed number of grains per DM_{65} (Table 4), these measurement errors seem to be sufficiently small to allow a model comparison.

4.2. Adjustment of genotype specific coefficients

Four different models were tested, assuming GPSM as a function of DM_{65} (Eq. (1)), ΔDM_{39-65} (Eq. (2)), a combination of Q45 and NNI_{60} (Eq. (3)), or as a product of Q45, NNI_{60} and DM_{65} (Eq. (4)). Genotype specific coefficients (G) were fitted to the data and served for the model adjustment to dataset. However, the ratio of GSPM and the explanatory variables did not vary greatly between the modern cultivars used for the Hohenschulen dataset, but differences were found between these compared to the thirty years older cv. *Arminda* used in the Wageningen dataset (Table 4) – also evident by much lower TKM values. Although the Wageningen dataset is small, the results suggest that a genotype specific adjustment can be useful especially if approaches calibrated in the past should be transferred to modern cultivars.

The original concept of M2 provides no genotype specific adjustment, assuming 100 grains per g spike biomass, where spike biomass is 50% of the accumulated stem growth between BBCH 39 and 65 (Moreno-Sotomayor and Weiss, 2004). This leads to 50 grains g^{-1} (stem growth), assuming the spike as a part of the stem. Subsequently, a grain set reduction was calculated according to the influence of high temperatures (Moreno-Sotomayor and Weiss, 2004). In our study, we assumed that stem growth during this phase equals shoot growth (ΔDM_{39-65}), since leaf growth is already completed. However, the G_2 coefficients (grains g^{-1}) of both datasets are smaller (Table 6) compared to the assumptions of Moreno-Sotomayor and Weiss.

Regarding M3, the genotype specific coefficient (G_3) is serving as a non-dimensional factor to modify the original model output. The obtained explanatory variables and GPSM values of the Hohenschulen dataset are within a similar range as those published by Demotes-Mainard and Jeuffroy (2001), from where M3 originated. Consequently, the genotype specific adjustment was small and G_3 is near one for the Hohenschulen dataset, but higher for the Wageningen dataset (0.95/1.36). Analogous to this, G_1 and G_4 are higher for the Wageningen dataset, showing that there is a noticeable genetic distance between *Arminda* and the modern cultivars used in the Hohenschulen dataset.

4.3. Interaction of explanatory parameters and physiology aspects

Our results confirm the possibility to estimate GPSM using the input parameters Q45 and NNI_{60} only (Demotes-Mainard and Jeuffroy, 2001), but for the Hohenschulen dataset the overall RMSE of M3 was fairly high compared to the other approaches (Table 6). It has to be noted, that there are substantial differences between the parameter DM_{65} and the parameters Q45 and NNI_{60}. The latter are ratios and therefore not influenced by factors like sowing density, drought stress or fraction of intercepted radiation. Although the correlation between a term including these variables exclusively (Eq. (3)) and GPSM might be high, both parameters are of qualitative nature and cannot give any authentic information about GPSM until spike or shoot biomass is incorporated.

Physiologically, it makes sense to include biomass or growth into the estimation of GPSM (M1, M2, M4), because shoot biomass is linked to spike biomass and number of spikes per m^2. Strong N deficiency leads to reduced growth and therefore N deficiency is indirectly included in models using biomass as input exclusively. Thus, a simple and physiologically reasonable model like M1 or M2 seems to be more appropriate at first sight, compared to M4 and M3 which are more empirical. Nevertheless, total dry weight at anthesis, as an index of resource accumulation, does not always correlate well with GPSM as reported by Fischer (2008). This is confirmed by our findings, since the correlation between NNI_{60} and GPSM is greater than between NNI_{60} and DM_{65} (Table 5).

The relationship between NNI_{60} and GSPM was non-linear in both datasets but became linear after log transformation of NNI_{60} (Fig. 1C, D, G and H). Thus, it makes sense to log transform NNI_{60} as an explanatory variable within a multiple linear regression (Eq. (3)). Although the relationship between GPSM and DM_{65} seems to be more linear (than for NNI_{60}), a much weaker slope was found for the upper 50% quantile compared to the lower 50% quantile (Fig. 1A). Thus, the assumption of a linear relationship between GPSM and these variables is problematic especially under varying N supply.

Although DM_{65}, NNI_{60} and Q45 are correlated (Table 5), M4 uses the product of these for the GPSM calculation. The underlying idea is that the GPSM/DM_{65} ratio is influenced by NNI_{60} and Q45 (as discussed). Following the regression of Fig. 2, the ratio GPSM/ln(DM_{65}) is a function of NNI_{60} times Q45, which is significant for both datasets:

$$\frac{\mathrm{GPSM}}{\ln(\mathrm{DM}_{65})} = f_1 + f_2 \times \ln(\mathrm{NNI}_{60} \times \mathrm{Q45}), \tag{5}$$

were f_1 and f_2 are the fit parameters of the linear regression. Further, GPSM/DM_{65} can be derived from this relationship:

$$\frac{\mathrm{GPSM}}{\mathrm{DM}_{65}} = [f_1 + f_2 \times \ln(\mathrm{NNI}_{60} \times \mathrm{Q45})] \times \frac{\ln(\mathrm{DM}_{65})}{\mathrm{DM}_{65}}. \tag{6}$$

The ratio ln(DM_{65})/DM_{65} on the right-hand side of Eq. (6) signifies that GPSM/DM_{65} and its response to NNI_{60} and Q45 is smaller for high values of DM_{65}. This might be due to the fact that the potential grain number an area can bear is limited physiologically.

However, the underlying mechanism of the variation in grain number per shoot weight can be explained by the fact that wheat adjusts the number of its offspring to environmental conditions, in particular radiation, temperature, N availability and drought, through large changes in the survival rate of floret primordials and competent florets (e.g. Ferrante et al., 2010). In this context, Fischer and Stockman (1980) found that pre-anthesis shading periods reduce kernel number per spike associated with reductions in the number of competent florets, while the potential number of tillers (which might later be able to bear grains) obtained under otherwise unstressed conditions is higher at cooler temperatures and greater light intensities (Huibert et al., 1998). Ferrante et al. (2010) found a diminished rate of floret abortion during the late part of stem elongation in response to enhanced N availability for durum wheat, which confirms earlier results obtained by Sibony and Pinthus (1988) in bread wheat. The relationship of light intensity and temperature during the phase of spike growth is reflected by the photothermal quotient, which is closely related to GPSM (e.g. Fischer, 1985; Demotes-Mainard and Jeuffroy, 2001; Bassu et al., 2010). The presented datasets contain five different years (Table 2), where Q45 ranges from 0.5 to 0.7 [MJ (m^2 °Cd)$^{-1}$]. Since radiation is the main driving force of growth and cooler temperatures reduce respiration losses, Q45 is linked to the potential growth rate during spike growth. Therefore, Q45 and DM_{65} are correlated but if GPSM could be sufficiently explained by Q45 or DM_{65} only, multiplying both parameters should increase the mean annual bias, defined as the absolute difference of means between simulated and observed grains for a specific year. However, the opposite is true: on average, the absolute annual bias is 632 for M4 compared with 1038–2219 [grains m^{-2}] obtained from the other approaches (not shown). The physiological explanation might be that Q45 as a representative of growing conditions during spike growth relates to the fraction of spike weight per shoot weight.

However, the variation of Q45 within the used datasets is too small to make a final conclusion. Nonetheless, in both datasets, at least Q45 does not disturb the role of ln(NNI_{60}) as a “modifier” of GPSM/ln(DM_{65}) (Fig. 2), which confirms the concept of M4 (Eq. (4)).

5. Conclusion

Our data shows that the N nutrition status of wheat crops at anthesis expressed by NNI_{60} modifies the number of grains per unit shoot biomass. In addition, there are indications that the photothermal quotient 45 days preceding anthesis (Q45) has an influence on this ratio as well. As a consequence, the consideration of Q45 and NNI_{60} could enhance the GPSM predictive quality for both datasets, compared to models using biomass or growth exclusively. Since M4 relies only on climatic variables, aerial growth, and N accumulation needed for the NNI_{60} calculation as simulated by most wheat crop models, this new approach could be useful for estimating GPSM in wheat crop models.

Acknowledgments

We gratefully thank F. Meier-Schatz and T. Johnen for providing their experimental data. Special thanks to Andreas Pacholski, Monika Sajduk and the unknown reviewer for their critical editorial remarks.

References

Bassu, S., Giunta, F., Motzo, R., 2010. Effects of sowing date and cultivar on spike weight and kernel number in durum wheat. Crop Pasture Sci. 61, 287–295.

Bertheloot, J., Andrieu, B., Fournier, C., Martre, P., 2008a. A process-based model to simulate nitrogen distribution within wheat (*Triticum aestivum* L.) during grain filling. Funct. Plant Biol. 35, 781–796.

Bertheloot, J., Martre, P., Andrieu, B., 2008b. Dynamics of light and nitrogen distribution during grain filling within wheat canopy. Plant Physiol. 148, 1707–1720.

Brancourt-Hulmel, M., Doussinault, G., Lecomte, C., Berard, P., Le Buanec, B., Trottet, M., 2003. Genetic improvement of agronomic traits of winter wheat cultivars released in France from 1946 to 1992. Crop Sci. 43, 37–45.

Brisson, N., Ripoche, M.B., 1998. STICS: a generic model for the simulation of crops and their water and nitrogen balances. I. Theory and parameterization applied to wheat and corn. Agronomie 18, 311–346.

Demotes-Mainard, S., Jeuffroy, M.H., 2001. Incorporating radiation and nitrogen nutrition into a model of kernel number in wheat. Crop Sci. 41, 415–423.

Duggan, B.L., Domitruk, D.R., Fowler, D.B., 2000. Yield component variation in winter wheat grown under drought stress. Can. J. Plant Sci. 80, 739–745.

Ferrante, A., Savin, R., Slafer, G.A., 2010. Floret development of durum wheat in response to nitrogen availability. J. Exp. Bot. 61 (15), 4351–4359.

Fischer, R.A., Stockman, Y.M., 1980. Kernel number per spike in wheat (*Triticum aestivum* L.): responses to preanthesis shading. Aust. J. Plant Physiol. 7, 169–180.

Fischer, R.A., 1985. Number of kernels in wheat crops and the influence of solar radiation and temperature. J. Agric. Sci. 105, 447–461.

Fischer, R.A., 1993. Irrigated spring wheat and timing and amount of nitrogen fertilizer. II. Physiology of grain yield response. Field Crops Res. 33, 57–80.

Fischer, R.A., 2008. The importance of grain or kernel number in wheat: a reply to Sinclair and Jamieson. Field Crops Res. 105, 15–21.

Gonzalez, F.G., Slafer, G.A., Miralles, D.J., 2003. Grain and floret number in response to photoperiod during stem elongation in fully and slightly vernalized wheats. Field Crops Res. 81, 17–27.

Groot, J.J.R., 1987. Simulation of nitrogen balance in a system of winter wheat and soil. Simulation Report CABO-TT 69.

Huibert, J., Jan, B.O.S., Neuteboom, H., 1998. Morphological analysis of leaf and tiller number dynamics of wheat (*Triticum aestivum* L.): responses to temperature and light intensity. Ann. Bot. 81, 131–139.

Jacobs, A.F.G., Heusinkveld, B.G., Holtslag, A.A.M., 2010. Eighty years of meteorological observations at Wageningen, the Netherlands: precipitation and evaporation. Int. J. Climatol. 30, 1315–1321.

Jeuffroy, M.H., Bouchard, C., 1999. Intensity and duration of nitrogen deficiency on wheat grain number. Crop Sci. 39, 1385–1393.

Johnen, T., Böttcher, U., Jörg, E., Kage, H., 2006. Parametrisierung und Evaluierung zweier Phänologiemodelle für Winterweizen anhand umfangreicher, mehrjähriger, mehrortiger Versuchsdaten. Mitt. Gesell. Pfl. bauwiss. 18, 244–245.

Justes, E., Mary, B., Meynard, J.M., Machet, J.M., Thélier-Huche, L., 1994. Determination of a critical nitrogen dilution curve for winter-wheat crops. Ann. Bot. 74, 397–407.

Kage, H., Stützel, H., 1999. HUME: an object oriented component library for generic modular modelling of dynamic systems. In: Donatelli, C.S.M., Villalobos, F., Villar, J.M. (Eds.), Modelling Cropping Systems. European Society of Agronomy, Lleida, pp. 299–300.

Keating, B.A., Carberry, P.C., Hammer, G.L., Probert, M.E., Robertson, M.J., Holzworth, D., Huth, N.I., Hargreaves, J.N.G., Meinke, H., Hochman, Z., McLean, G., Verburg, K., Snow, V., Dimes, J.P., Silburn, M., Wang, E., Brown, S., Bristow, K.L., Asseng, S., Chapman, S., McCown, R.L., Freebairn, D.M., Smith, C.J., 2003. An overview of APSIM, a model designed for farming systems simulation. Eur. J. Agron. 18, 267–288.

Kibite, S., Evans, L.E., 1984. Causes of negative correlations between grain-yield and grain protein-concentration in common wheat. Euphytica 33, 801–810.

Marquardt, D.W., 1963. An algorithm for least-squares estimation of nonlinear parameters. J. Soc. Ind. Appl. Math. 11, 431–441.

Meier, U., 2001. Entwicklungsstadien mono- und dikotyler Pflanzen. Biologische Bundesanstalt für Land und Forstwirtschaft. http://www.jki.bund.de/fileadmin/dam_uploads/_veroeff/bbch/BBCH-Skala_englisch.pdf.

Midmore, D.J., Cartwright, P.M., Fischer, R.A., 1984. Wheat in tropical environments, 2. Crop growth and grain yield. Field Res. 8, 207–227.

Moreno-Sotomayor, A., Weiss, A., 2004. Improvements in the simulation of kernel number and grain yield in CERES-Wheat. Field Crops Res. 88, 157–169.

Otter-Nacke, S., Goodwin, D.C., Ritchie, J.T., 1986. Testing and validating the CERES-Wheat model in diverse environments. In: AGRISTARS Program, USDA-ARS, PB87-181483, 147 p., cited in: Moreno-Sotomayor, A., Weiss, A., 2004. Improvements in the simulation of kernel number and grain yield in CERES-Wheat. Field Crops Res. 88, 157–169.

Reynolds, M.P., Pellegrineschi, A., Skovmand, B., 2005. Sink-limitation to yield and biomass: a summary of some investigations in spring wheat. Ann. Appl. Biol. 146, 39–49.

Ritchie, J.T., Otter, S., 1985. Description and Performance of CERES-wheat: A User Oriented Wheat Yield Model. ARS Wheat Yield Project. USDA-ARS, Springfield, VA, pp. 159–175.

Savin, R., Slafer, G.A., 1991. Shading effects on the yield of an Argentinian wheat cultivar. J. Agric. Sci. 116, 1–7.

Sayre, K.D., Rajaman, S., Fischer, R.A., 1997. Yield potential progress in short bread wheats in northwest Mexico. Crop Sci. 37, 36–42.

Sibony, M., Pinthus, M.J., 1988. Floret initiation and development in spring wheat (*Triticum aestivum* L.). Ann. Bot. 61, 473–479.

Sinclair, T.R., Jamieson, P.D., 2006. Grain number, wheat yield, and bottling beer: an analysis. Field Crops Res. 98, 60–67.

Sinclair, T.R., Jamieson, P.D., 2008. Yield and grain number of wheat: a correlation or causal relationship? Authors' response to "The importance of grain or kernel number in wheat: a reply to Sinclair and Jamieson" by R.A. Fischer. Field Crops Res. 105, 22–26.

Van Delden, A., Pecios, A., Haverkort, A.J., 2000. Temperature response of early foliar expansion of potato and wheat. Ann. Bot. 86, 355–369.

Vos, J., 1981. Effect of temperature and nitrogen supply on post anthesis growth of winter wheat: measurements and simulations. Agric. Res. Rep. 811, Pudoc, Wageningen, 164 p. cited in Groot, J.J.R., 1987: Simulation of nitrogen balance in a system of winter wheat and soil. Simulation Report CABO-TT: 69.

Zadoks, J.C., Chang, T.T., Konzak, C.F., 1974. A decimal code for the growth stages of cereals. Weed Res. 14, 415–421.