

# Toward a Bayesian procedure for using process-based models in plant breeding, with application to ideotype design

M. Van Oijen  · M. Höglind

Received: 27 March 2015 / Accepted: 19 September 2015 / Published online: 29 September 2015  
© The Author(s) 2015. This article is published with open access at Springerlink.com

**Abstract** Process-based grassland models (PBMs) simulate growth and development of vegetation over time. The models tend to have a large number of parameters that represent properties of the plants. To simulate different cultivars of the same species, different parameter values are required. Parameter differences may be interpreted as genetic variation for plant traits. Despite this natural connection between PBMs and plant genetics, there are only few examples of successful use of PBMs in plant breeding. Here we present a new procedure by which PBMs can help design ideotypes, i.e. virtual cultivars that optimally combine properties of existing cultivars. Ideotypes constitute selection targets for breeding. The procedure consists of four steps: (1) Bayesian calibration of model parameters using data from cultivar trials, (2) Estimating genetic variation for parameters from the combination of cultivar-specific calibrated parameter distributions, (3) Identifying parameter combinations

that meet breeding objectives, (4) Translating model results to practice, i.e. interpreting parameters in terms of practical selection criteria. We show an application of the procedure to timothy (*Phleum pratense* L.) as grown in different regions of Norway.

**Keywords** BASGRA · Cold tolerance · Genotype-environment interaction · Plant breeding · Process-based modelling · Yield stability

## Terminology

This paper aims to bridge two different plant disciplines, i.e. process-based modelling and breeding, so terms need to be defined. In modelling, we distinguish between *plant parameters*, which are genotype-specific constants, and *output variables*, which are plant characteristics that are predicted by the model and that vary between environments. *Process-based models* (PBMs) are dynamic models that represent physiological and morphological processes in plants, and their interaction with the environment. *Simulation* is the process of specifying environment and parameter values followed by running the PBM to calculate the output variables. In plant breeding, *traits* are measurable properties of plants that arise from the interaction between the genotype and its environment ( $G \times E$ ). The influence of the environment may be large for some traits (e.g. yield) or small (e.g. flower size). A *quantitative trait* is a property measured on a

---

**Electronic supplementary material** The online version of this article (doi:10.1007/s10681-015-1562-5) contains supplementary material, which is available to authorized users.

---

M. Van Oijen (✉)  
Centre for Ecology & Hydrology, CEH-Edinburgh, Bush  
Estate, Penicuik EH26 0QB, UK  
e-mail: mvano@ceh.ac.uk

M. Höglind  
Norwegian Institute of Bioeconomy Research (NIBIO),  
4353 Klepp St., Norway

continuum scale that depends on many genes. An *ideotype* (Donald 1968) is a collection of traits, not necessarily realised yet by any existing cultivar, that a breeder believes is leading to high performance. *Performance* is generally measured by a small number of *performance traits* targeted by the breeder. An example would be high average yield combined with high yield stability. This paper studies the relationships between plant parameters and output variables on the one hand, and plant traits on the other.

## Introduction

### Plant breeding and process-based modelling

In many studies of plants, an effort is made to relate plant- or vegetation-level characteristics, such as growth rate or yield, to underlying physiological and morphological properties. The objective is to explain or predict how plants respond to their environment, or to determine how plants can be improved by breeding. Nearly five decades ago, Donald (1968) introduced the concept of the ‘ideotype’, defined as a set of plant traits that, if brought together in one new genotype, would lead to high performance. He used the example of breeding for high-yielding wheat cultivars that would pose low demands on external inputs. The wheat ideotype traits identified by Donald (1968) were mainly morphological, such as erect leaves and low stem height, but he suggested that they would contribute to improved crop physiology, e.g. high rate of photosynthesis per m<sup>2</sup> ground area. The idea behind ideotype-design was that it would provide the breeder with an ultimate target for selection, thereby replacing the trial-and-error method of stepwise increasing plant performance.

Ideotype breeding thus focuses on multiple traits simultaneously. The method differs from other multivariate approaches in plant breeding, such as index selection, which tend to focus on the performance traits and not on the underlying plant morphology and physiology. Focusing directly on performance simplifies the problem statistically, as there are fewer traits to consider, but performance traits are difficult to measure reliably in other conditions than full-scale field tests.

Despite Donald’s confidence that “eventually most plant breeding may be based on ideotypes”, the

approach has not been widely used in practice (Zhang et al. 1999). Application is hampered by the fact that we have only limited information about how plant performance depends on the underlying physiology and morphology of the plants, and information about genetic variation for these underlying characteristics is also limited (Marshall 1991).

Process-based models (PBMs) of crops and other managed ecosystems simulate the growth of plants based on a representation of the underlying morphological and physiological processes and the interaction of the plants with their growing environment. Every represented process requires at least one parameter, so PBMs tend to be more parameter-rich than, for example, statistical yield models or formulas for index selection. Parameters representing the environment, such as soil water retention characteristics and fertility, are site-specific parameters that need to be changed whenever the model is applied to a new site. The plant parameters governing physiology and morphology, on the other hand, are generally treated as cultivar- or species-specific constants, sometimes referred to as “genetic coefficients” (Yin et al. 2004). PBMs can speed up assessment of the value of a measured trait, or a specific combination of traits, by predicting the extent to which a measured genetic difference in a crop character will affect plant performance in different environments. Many more years and environments can be simulated than can feasibly be included in trials (Cooper et al. 2014). In view of these characteristics, PBMs have been considered as possible tools for the definition of selection criteria for breeding.

An early example of such use of PBMs was the work by Landivar et al. (1983), who used a cotton model to examine the impact of changing leaf properties, including photosynthesis, on fibre yield, but the work was not coupled to a breeding programme. Van Oijen (1992) used a potato model to identify key components of resistance to late blight, which were later targeted successfully by Colon et al. (1995) in selection for parental line breeding material. This was a simple application, and not an example of model-supported ideotype design, as only five resistance components were investigated and no attempt was made to optimise any of the other plant parameters. In a more comprehensive study, Semenov et al. (2014) used a PBM to define wheat ideotypes for future climate conditions in Europe based on an

optimization procedure including nine plant parameters. The other plant parameters in the model were left out of the study as they were considered less promising for yield improvement based on reasoning not presented in the paper. Further examples of PBM-supported ideotype design, specifically for rice, can be found in the volume edited by Aggarwal et al. (1995)—again without apparent impact on actual breeding practice.

Overall, the role of PBMs in plant breeding seems to have remained small for various reasons (Messina et al. 2009). There are several requirements to be met before a PBM can be used effectively in plant breeding. First of all, the PBM needs to be shown to provide realistic simulations of the crop for the intended growing environments (Tardieu 2003). Secondly, the large number of model parameters all need to be quantified, as does the available genetic variation for the parameters. Parameter optimisation should not target just one trait but be multivariate. Finally, results of the analysis in terms of model parameters need to be translated into practical selection criteria that can be used in breeding programmes. This paper aims to contribute to addressing these problems.

#### Cultivar-specific parameter estimation

The parameterisation of PBMs tends to rely on direct measurement and literature reviews (Breuer et al. 2003; Levy et al. 2004), including meta-analysis (Medlyn et al. 1999). The literature studies often show large variation for each examined parameter, causing large uncertainty in model outputs (Levy et al. 2004).

In recent years, alternative methods for parameter estimation for PBMs have been developed that are based on representing uncertainty about parameter values as probability distributions (Kennedy and O'Hagan 2001; Van Oijen et al. 2005b; Fox et al. 2009; Ogle 2009; Wang et al. 2009). The most generally applicable of these methods is Bayesian calibration (Kennedy and O'Hagan 2001). This is based on Bayes' theorem, according to which a prior distribution for the model's parameters can be updated when new data come in, by multiplying the prior with the likelihood function for the data. Bayesian calibration of PBMs relies on drawing a representative sample of parameter vectors from the parameter distribution, and this is generally carried out using

Markov Chain Monte Carlo techniques (MCMC; Metropolis et al. 1953). Bayesian calibration using MCMC has been applied to models for Norway spruce (Van Oijen et al. 2005b), N<sub>2</sub>O-emitting fields of rapeseed, winter wheat and maize (Lehuger et al. 2009) and the dynamics of soil under grassland during winter (Thorsen et al. 2010). MCMC not only allows the use of complex models, such as PBMs, that are not analytically solvable, but it also allows uncertainties about parameters and measurements to be represented by the most appropriate probability distributions; there is no need to use standard distributions such as the multivariate normal. As far as we are aware, Bayesian calibration has not yet been used to quantify differences in PBM parameter values between cultivars of the same species.

#### Quantifying genetic variation for plant parameters

Yin et al. (1999, 2000, 2004) showed that in some cases, such as for the specific leaf area of barley, it is possible to relate parameters or output variables of PBMs to the genome, usually in the form of polygenic traits corresponding to multiple QTLs (*Quantitative Trait Loci*, parts of DNA that co-vary with a quantitative trait). Similarly, Reymond et al. (2003) were able to relate QTLs to the parameters that govern response of maize leaf expansion to temperature and water status, and Dong et al. (2012) developed a regulatory network model for over 30 genes controlling flowering time in maize. These various studies were largely based on recombinant inbred lines, thus simplifying the genetic differences between genotypes, and the variety of environmental conditions was also limited. Generally G×E interaction will be more complex (Tardieu 2003; Cooper et al. 2014), hampering the calculation of parameter values as functions of QTLs. The effort to explicitly link model parameters to the genome remains an active area of research (Chenu et al. 2009; Hammer et al. 2006; Tardieu 2003; Xu et al. 2011; Zheng et al. 2013).

Because direct measurement of the genetic basis of model parameters is difficult, we may use calibration methods instead, as in the present work. When model calibration leads to highly divergent parameter values for different cultivars, genetic variation for the parameters can be considered to be wide. However, beyond that general statement, there seem to be no formal methods for estimating genetic variation for

plant parameters in PBMs. This stands in contrast to the use of classical mixed models for inferring genetic parameters from selection experiments, which originated in the animal breeding literature of the 1950s (Sorensen 2009), was later introduced in plant breeding (Piepho et al. 2008) and for which both least squares and MCMC estimation methods are in use (Sorensen et al. 1994).

#### Deriving an ideotype in terms of model parameters

Once genetic variation for parameters has been estimated, the next problem is how to identify the ideotype, i.e. the optimal combination of plant parameter values from within the range of genetic variation. When PBMs are used, this will be evaluated on the basis of model outputs that correspond to plant performance traits of interest to the breeder, who must decide how to evaluate trade-offs between different performance traits. A strength of PBMs is that they demonstrate the existence of these trade-offs as the inevitable consequence of limited availability of resources for growth, and feedbacks between different processes (Van Oijen et al. 2004); the trade-offs do not need to be detected by analysing data on large numbers of genotypes.

In terms of modelling, the main problem may be that the dimensionality of trait space is high, even if we restrict ourselves to those traits that are represented by a plant parameter in the PBM. Here we can take advantage of MCMC methods, mentioned above, which are designed to sample efficiently from a high-dimensional probability distribution.

#### Translating results into practical selection criteria

An ideotype that has been determined using a PBM is defined in terms of model parameter values. However, not every model parameter is an easily measured quantity. The next step will therefore be to translate the model parameters into traits that can be measured on real plants. This design of effective selection criteria for use in practical breeding will require the input from breeders who can decide which measurements can be carried out at low cost and at sufficient speed and accuracy. This step goes beyond what a PBM can deliver.

#### Toward a Bayesian procedure for PBM-assisted ideotyping

In the present paper, we propose the following four-step procedure for PBM-assisted ideotype design in plant breeding:

1. Parameterise the model for different genotypes by means of Bayesian Calibration (BC), using information from variety trials as input.
2. Combine the genotype-specific parameter distributions into one distribution representing the overall genetic variation for plant parameters.
3. Limit the genetic variation distribution using the breeding objectives as constraints, and derive the ideotype.
4. Translate the constrained parameter distribution into multivariate selection criteria that can be used by breeders.

This procedure is intended to be generally applicable to any crop for which a sufficiently accurate PBM is available. We shall give an example for timothy (*Phleum pratense* L.), a grass species widely grown at higher latitudes. For this crop, the model BASGRA is available which simulates the year-round growth and survival of timothy plants for a range of climatic conditions (Höglind et al. 2001; Van Oijen et al. 2005a; Thorsen et al. 2010; Höglind et al. in prep.). There has been some study of the variation of BASGRA parameter values (Höglind et al. 2001; Thorsen et al. 2010), but knowledge is still limited and there remains a need for cultivar-specific calibration of the model. Bayesian calibration of BASGRA for two timothy cultivars will be part of the present paper as the first step of the proposed ideotyping procedure.

In the next section, we describe how the four-step procedure that we outlined above was implemented for the timothy test-case, and we provide details of the calibration data and the BASGRA model.

## Materials and methods

### Data

To calibrate the model, data were used from field experiments on timothy (*Phleum pratense* L.) that had been carried out on different locations in Norway (Table 1), covering the major agricultural grassland

**Table 1** Sites with timothy experiments

Location	Experiments	Latitude	Longitude	Elevation (m)	Climatic means, 1995–2012	
					Temperature (°C)	Precipitation (mm year <sup>-1</sup> )
Apelsvoll	3	60°70'N	10°87'E	255	4.8	679
Fureneset	1, 2	61°29'N	5°04'E	12	7.7	2280
Holt	1, 2	69°65'N	18°90'E	12	3.8	966
Kvithamar	1, 2	63°49'N	10°88'E	28	6.0	1007
Løken	–	61°49'N	9°06'E	527	2.7	637
Særheim	4	59°28'N	9°28'E	83	7.8	1430

Data for calibration were collected at all locations except Løken. All locations except Fureneset were included in the subsequent simulation study to identify an ideotype for timothy in Norway

areas in the western coastal regions and the eastern lowlands. Experiments (Exp.) 1–3 included two cultivars, Grindstad and Engmo, whereas Exp. 4 only included Grindstad. More details on Exp. 1 can be found in Höglind et al. (2006), on Exp. 2 in Höglind et al. (2010) and on Exp. 4 in Höglind et al. (2005). Exp. 3 contains previously unpublished material (Sunde 1996). Grindstad and Engmo are the only timothy cultivars for which a wide range of variables have been measured in Norwegian field trials, making them suitable for process-based modelling.

Exp. 1 was carried out at three locations: Fureneset, Holt and Kvithamar (Table 1). From November 2005 to March 2006, on five occasions per location, shoot dry weight, leaf area index (LAI), specific leaf area (SLA), tiller density, content of water soluble carbohydrates (WSC), and frost tolerance (LT50) were determined. In addition, tiller density and DM yield (total dry weight of herbage above a stubble height of 5 cm) was determined in June 2006.

Exp. 2 was carried out at the same locations as Exp. 1 (Table 1). From November 2006 to March 2007, on three occasions per location, shoot biomass, tiller density, WSC, and LT50 were determined. In addition, tiller density was determined at one location in June 2007. The swards were cut once to twice in the growing season 2006, and twice to three times in the growing season 2007, and the DM yield of each cut was determined.

Exp. 3 was carried out at Apelsvoll (Table 1). Sampling was carried out from August 1990 to April 1991, on 13–15 occasions depending on cultivar and plant variable to determine WSC and LT50.

Exp. 4 was carried out at Særheim (Table 1). There were two fields. The first field was established in 1999,

with measurements taken in 2000. The other field was established in 2000, with measurements taken in 2001 and 2002. Two harvesting regimes were compared in each field. One harvesting regime consisted of an early first cut, and the other of a late first cut, each first cut followed by a second cut after 6–8 weeks of regrowth. From April to August each year, with sampling intervals of 7–14 days, shoot biomass, LAI, SLA, tiller density, WSC, leaf appearance rate, number of elongating leaves per tiller, and leaf elongation rate per actively growing leaf were determined.

All experimental locations were equipped with automatic weather stations, located within 500 m from the experimental field, which were connected to the weather database of NIBIO. For the calibration of the model, daily weather data were downloaded from the weather database of NIBIO.

**BASGRA:** a process-based model for managed grasslands

BASGRA is a process-based model (PBM) for grassland that was derived from an earlier model called LINGRA (Schapendonk et al. 1998). General model characteristics remain unchanged: leaf-area determines light interception which drives carbon assimilation through a water-status dependent light-use efficiency, and carbon allocation depends on the balance between carbon sources (assimilation and reserve mobilisation) and sinks (leaf growth, root growth, tillering). To make LINGRA usable for studying climate change impacts, the effects of CO<sub>2</sub> and temperature on the light-use efficiency of the sward were included by Rodriguez et al. (1999). The model was originally used for perennial ryegrass (*Lolium perenne* L.), but to

simulate timothy (*Phleum pratense* L.) as well, tillering was simulated in greater detail, distinguishing elongating from non-elongating tillers (Höglind et al. 2001; Van Oijen et al. 2005a). Algorithms for winter processes, including hardening and frost damage, were developed more recently (Thorsen et al. 2010). The model code was translated from Matlab to FORTRAN by D. Cameron, and the ‘summer’ and ‘winter’ processes were linked together, producing the year-round model now called BASGRA (Höglind et al. in prep.). BASGRA operates on a daily time step, and requires input information on management (cutting regime), environmental variables (radiation, temperature, precipitation, wind speed, [CO<sub>2</sub>]) and environmental constants (soil water retention parameters). The model has 45 plant parameters whose meaning, units and prior probability distributions are given in Online Resource 1. The model is archived online (Van Oijen et al. 2015), from where full code, including for Bayesian calibration, can be downloaded.

#### Multi-site Bayesian calibration of BASGRA

Bayesian calibration (BC) is a procedure for probabilistic parameter estimation of models using measurements of model output variables. BC was used here to parameterise the model BASGRA for two cultivars. BC was implemented in the same way as described by Van Oijen et al. (2005b, 2013) for forest PBMs. We refer to these studies for technical details, and only give a general overview here.

In the present study, Bayesian calibration was carried out separately for timothy cultivars Engmo and Grindstad. Both calibrations were multi-site, i.e. data from sites across Norway (Table 1) were used. Plant parameters were allowed to differ between cultivars, but not between sites. This constraint implies a test of the model: if model performance is poor when parameters are held constant over sites, then those parameters do not represent genetic coefficients and the model is not suitable for ideotype design.

Bayesian calibration of BASGRA consists of three steps: (1) define the prior distribution for the model’s parameters, (2) define the likelihood function for the model’s parameters, (3) sample from the ‘posterior distribution’ given by the normalised product of prior and likelihood. The posterior distribution expresses how the data have reduced our uncertainty about parameter values.

The same prior distribution was used for the parameter sets of both cultivars. Prior parameter ranges for individual parameters were derived from earlier literature study (Höglind et al. 2001; Van Oijen et al. 2005a; Thorsen et al. 2010) where available, and wide ranges of plausible values were assumed otherwise. The marginal prior probability distribution for each individual parameter was defined as a beta distribution over the parameter’s range of plausible values. No information on parameter correlations was available, so the joint prior distribution for the parameters was written as the product of the marginal distributions.

The likelihood function quantified the probability, for any given parameter vector, of the mismatch between the model outputs induced by the parameter vector and the data (5 sites, 11 variables). The measurement error terms in the likelihood function followed the conventional assumption of independent Gaussians with variances that varied between variables and observations.

The sample from the posterior distribution was generated by means of MCMC using the Metropolis algorithm (Metropolis et al. 1953; Van Oijen et al. 2005b). Chain length was 300,000 to ensure convergence for all parameters.

In the summary of the results of the Bayesian calibration for the two cultivars (Tables 2 and 3), we use the Normalised Root Mean Square Error (NRMSE) to quantify the mismatch between the data and the outputs from the calibrated model. The NRMSE is defined as the square root of the mean squared difference between observations and outputs, divided by the mean of the observations.

#### Combining cultivar-specific parameter distributions

Ideally, information on a large number of genotypes is available when estimating genetic variation for plant parameters. However, PBMs require detailed information, which was only available for two timothy cultivars grown in Norway. This may be enough to illustrate the procedure here, but more genotypes would be required for application in practice. The two posterior parameter distributions, for cultivars Engmo and Grindstad, were combined into one distribution representing the estimated overall genetic variation for plant parameters. This was done in two steps. First we

**Table 2** Bayesian calibration for timothy cultivar Engmo

		Apelsvoll	Fureneset	Holt	Kvithamar	Særheim
Biomass (g m <sup>-2</sup> )	Mean of data		210 (5)	38 (8)	89 (3)	35 (8)
	NRMSE		0.23 (–20)	0.90 (4)	0.24 (–57)	0.57 (–54)
Reserves (g g <sup>-1</sup> )	Mean of data	0.22 (13)	0.12 (5)	0.23 (8)	0.34 (3)	0.23 (7)
	NRMSE	0.29 (–37)	0.34 (142)	0.45 (–15)	0.51 (–19)	0.48 (–3)
Yield (g m <sup>-2</sup> )	Mean of data		300 (2)	340 (3)	260 (5)	220 (5)
	NRMSE		0.50 (–5)	0.36 (–50)	0.37 (–40)	0.74 (–14)
LAI (m <sup>2</sup> m <sup>-2</sup> )	Mean of data		1.1 (5)	0.2 (5)		0.2 (5)
	NRMSE		0.18 (–76)	0.45 (–35)		0.60 (–42)
SLA (m <sup>2</sup> g <sup>-1</sup> )	Mean of data		0.027 (5)			
	NRMSE		0.35 (–8)			
Tiller density (m <sup>-2</sup> )	Mean of data		2230 (6)	1420 (9)	2770 (3)	1930 (10)
	NRMSE		0.27 (–36)	0.33 (–45)	0.20 (–21)	0.36 (17)
Generative tiller fraction (–)	Mean of data			0.8 (1)	0.8 (1)	0.8 (1)
	NRMSE			0.08 (–89)	0.04 (–94)	0.02 (–97)
LT50 (°C)	Mean of data	–17.0 (12)	–13.2 (5)	–19.4 (7)	–22.0 (2)	–18.6 (8)
	NRMSE	0.15 (–42)	0.29 (–32)	0.27 (–14)	0.10 (–14)	0.25 (–2)

The mean of the data is given with the number of measurements in brackets. NRMSE is the Normalised Root Mean Square Error for the mismatch between data and model for the mode of the posterior distribution. It is followed in brackets by the percent change in NRMSE from prior to posterior mode (in bold if negative)

calculated the mean and variance of the union of the two distributions. Irrespective of the shape and number of the individual distributions, the mean of the union of multiple distributions is the average of the individual means, whereas the variance of the union is equal to the average of the individual variances plus the variance of the individual means. In the second step, we calculated beta distributions with the given combined mean and variance, under the assumption that prior parameter bounds were conserved. Using beta distributions, which have lower and upper limits, rather than Gaussians prevented spurious inclusion of extreme values in the estimated genetic variation. The result of this procedure was a different beta-distribution for each of the 45 plant parameters of BASGRA. In this calculation, we ignored the correlations between parameters that were present in the two cultivar-specific distributions, because those correlations may represent joint uncertainty rather than genetic linkage.

#### Deriving an ideotype and selection criteria

To derive the ideotype, we evaluated two performance traits: average yield and yield stability. Both traits were calculated using ‘virtual trials’, i.e. BASGRA

simulations at five sites in Norway (Table 1), with a standard sequence at each site of six three-year long grass rotations for the period 1995–2012. These sites were chosen to cover the climatic range of timothy in Norway. Two harvests per year were simulated, so the total number of simulated yield values amounted to 180 (5 sites × 18 years × 2 harvests). Average yield was calculated as the mean of the 180 values, and yield stability as the inverse of the coefficient of variation across the 180 values. Yield stability thus is defined with respect to both spatial and temporal variability. Both values were normalised by dividing them with the values of yield and yield stability for the mode of the genetic variation distribution, i.e. the genotype for which the distribution reaches its peak.

We defined the *performance* of a parameter vector as the minimum value of the two normalised performance traits.

We defined *acceptable genotypes* as those parameter vectors for which performance was greater than one, i.e. both yield and yield stability exceeded that of the mode of the genetic variation distribution. MCMC was used to generate a sample of 500,000 acceptable genotypes, and the ideotype was defined as the highest performing parameter vector in this sample.

The ideotype thus defined consisted of 45 parameter values, which is much more than can be considered in any breeding programme. We therefore determined which of the parameters contributed most to performance by calculating the partial correlation coefficient (PCC) of each parameter with performance. Parameters with highest PCC were considered to be prime candidates for selection criteria.

## Results

### Cultivar-specific Bayesian calibration

A priori, the same probability distribution had been assigned to the parameters of cultivars Engmo and Grindstad. Bayesian calibration introduced differences between the cultivars in the form of divergent marginal posterior distributions for parameters (Fig. 1). For 40 % of the parameters, the difference

**Fig. 1** Distributions for the 45 parameters of the BASGRA model. *Thin black line* prior distribution. *Green* posterior for cv. Engmo. *Blue* posterior for cv. Grindstad. *Red* genetic variation. The abscissa is scaled to the prior, so the mode of the prior is always at parameter value 1.0. The prior limits and mode of the unscaled parameters in their original units are provided in Online Resource 1

between posterior modes was more than one quarter of the overall mean. For some parameters, the posterior means differed, e.g. for parameter LAITIL (tillering rate at low LAI), whereas for other parameters, such as DAYLG1G2 (day length below which no stem elongation takes place), the posterior means were similar but the variance was lower for Grindstad than for Engmo. This reflects the higher information content of the data for Grindstad, for which more observations were available (compare Tables 2 and 3).

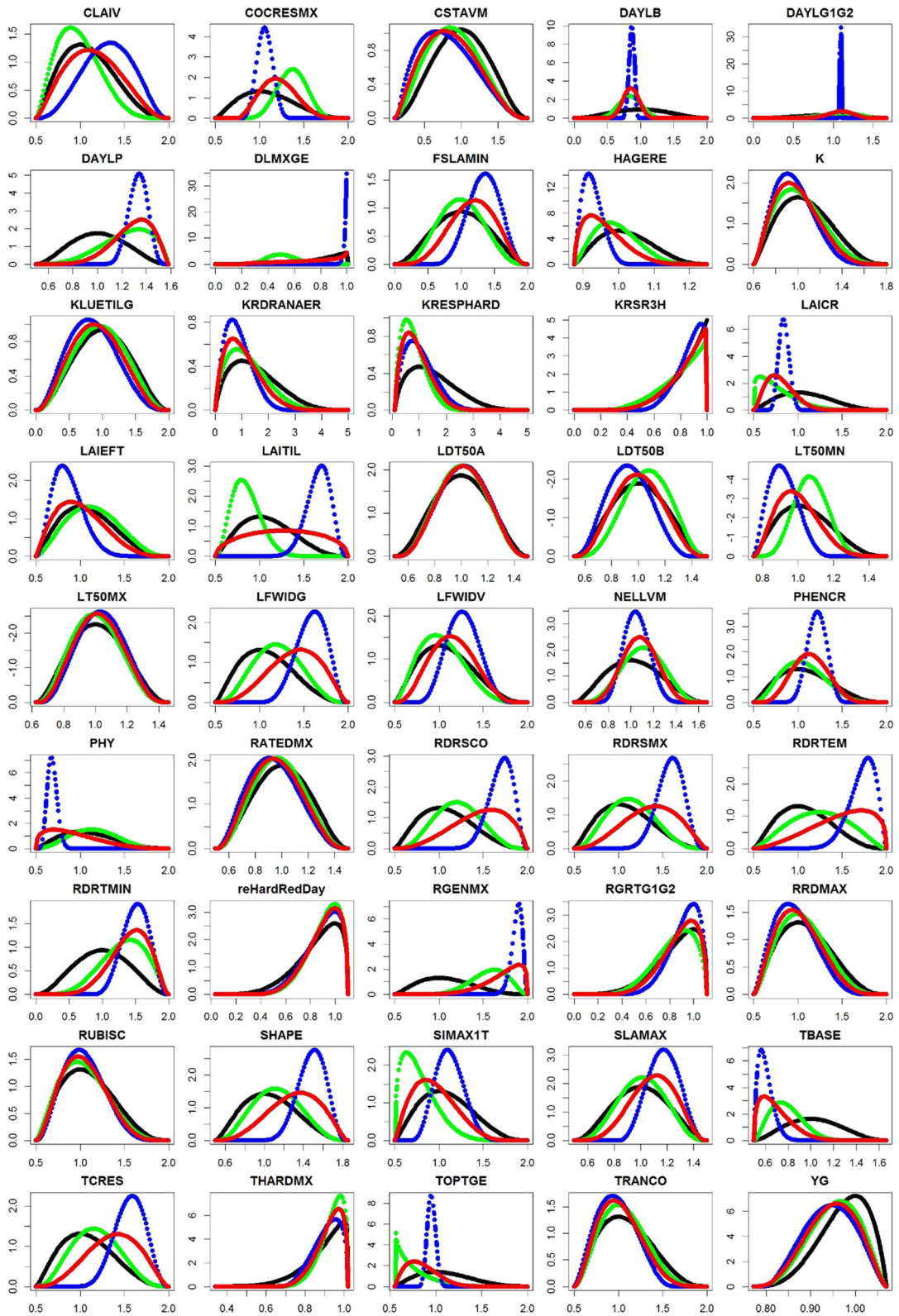
The Bayesian calibration was carried out using data from five different sites on a total of 11 different

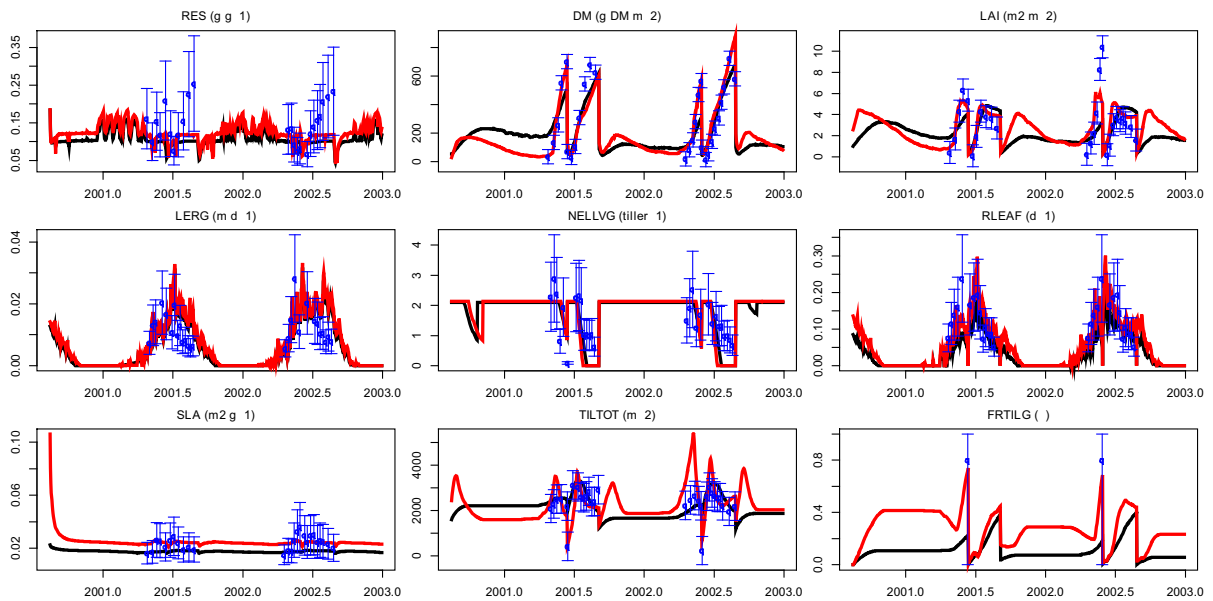
**Table 3** Bayesian calibration for timothy cultivar Grindstad

		Apelsvoll	Fureneset	Holt	Kvithamar	Særheim
Biomass (g m <sup>-2</sup> )	Mean of data		162 (6)	26 (8)	94 (3)	352 (88)
	NRMSE		0.27 (31)	1.01 ( <b>-48</b> )	0.36 ( <b>-22</b> )	0.36 ( <b>-30</b> )
Reserves (g g <sup>-1</sup> )	Mean of data	0.20 (15)	0.12 (5)	0.17 (8)	0.25 (3)	0.14 (79)
	NRMSE	0.45 ( <b>-9</b> )	0.13 (19)	0.49 ( <b>-2</b> )	0.42 ( <b>-17</b> )	0.37 ( <b>-14</b> )
Yield (g m <sup>-2</sup> )	Mean of data		350 (2)	370 (3)	300 (5)	280 (5)
	NRMSE		0.34 ( <b>-33</b> )	0.71 (58)	0.30 ( <b>-33</b> )	0.39 ( <b>-37</b> )
LAI (m <sup>2</sup> m <sup>-2</sup> )	Mean of data		1.3 (5)	0.2 (5)		3.6 (80)
	NRMSE		0.22 ( <b>-55</b> )	0.46 ( <b>-89</b> )		0.40 ( <b>-28</b> )
SLA (m <sup>2</sup> g <sup>-1</sup> )	Mean of data		0.029 (5)	0.033 (5)		0.026 (80)
	NRMSE		0.31 ( <b>-35</b> )	0.34 ( <b>-31</b> )		0.26 ( <b>-36</b> )
Leaf appearance (tiller <sup>-1</sup> d <sup>-1</sup> )	Mean of data					0.13 (80)
	NRMSE					0.45 ( <b>-15</b> )
Leaf elongation (m d <sup>-1</sup> )	Mean of data					0.012 (80)
	NRMSE					0.68 (11)
Elongating leaf density (tiller <sup>-1</sup> )	Mean of data					1.6 (81)
	NRMSE					0.52 (6)
Tiller density (m <sup>-2</sup> )	Mean of data		1870 (6)	1310 (9)	2990 (3)	2680 (102)
	NRMSE		0.32 ( <b>-44</b> )	0.68 ( <b>-59</b> )	0.20 (9)	0.42 (29)
Generative tiller fraction (-)	Mean of data			0.8 (1)	0.8 (1)	0.8 (7)
	NRMSE			0.16 ( <b>-78</b> )	0.14 ( <b>-81</b> )	0.14 ( <b>-79</b> )
LT50 (°C)	Mean of data	-12.8 (13)	-12.4 (5)	-16.1 (8)	-20.0 (3)	-16.6 (8)
	NRMSE	0.35 ( <b>-43</b> )	0.17 ( <b>-65</b> )	0.10 ( <b>-78</b> )	0.18 (3)	0.21 ( <b>-15</b> )

Abbreviations as in Table 2







**Fig. 2** Partial results of the Bayesian calibration for cv. Grindstad, showing prior and posterior time series at location Særheim, with observations in 2001–2002. *Blue* observations  $\pm$  standard deviation. *Black* model outputs for the mode of the prior distribution. *Red* model outputs for the posterior

mode  $\pm$  standard deviation. *RES* reserve content, *DM* dry matter, *LAI* leaf area index, *LERG* leaf elongation rate, *NELLVG* number of elongating leaves, *RLEAF* relative leaf appearance rate, *SLA* specific leaf area, *TILTOT* tiller density, *FRTILG* fraction of tillers that is generative

variables. Figure 2 shows an example of the impact of the calibration on model behaviour for one of the most data-rich sites, the Grindstad experiment at Særheim in the years 2001–2002. When BASGRA is run with the mode of the posterior distribution for the Grindstad parameters (the so-called maximum a posteriori parameter vector, MAP), all nine output variables are consistent with the measurements taken on the site.

Tables 2 and 3 give an overview for all sites of BASGRA behaviour when the model is run with the MAP for Engmo and Grindstad, respectively. The tables confirm that a single parameter vector, the MAP, suffices to simulate nearly all measured variables, including yield, to a NRMSE of less than 0.5. This is in the same order of magnitude as the coefficient of variation of the data. The NRMSE does exceed 0.5 in cases where the observations were restricted to periods of the year with low values of the variable, such as the biomass measurements at Holt, which were carried out during winter time. The two tables also indicate how the NRMSE for the MAP compared to the NRMSE for the mode of the prior distribution for the parameters. In most cases, the NRMSE decreased, but sometimes the prior was

superior. This reflects the fact that the calibration aims to reconcile data and outputs for all sites and variables at the same time, so variables for which there is little information in the data carry little weight in the calibration.

#### Estimating genetic variation

Genetic variation for individual parameters was quantified as beta distributions with the same mean and variance as the union of the cultivar-specific posterior distributions (Fig. 1). The genetic variation mean is thus always exactly midway between those for the individual cultivars. The variance of the genetic variation also tends to be close to the average of the individual variances, but in a small number of cases the variance of the genetic variation distribution is larger than that for the individual cultivars and even larger than for the prior distribution. Such wide distributions for genetic variation are derived when the calibration pulls the two cultivars far apart, such as for DLMXGE (day length below which reserves become a prominent sink) and LAITIL (tillering rate at low LAI).

## Sampling from the genetic variation and identifying the ideotype

Acceptable genotypes were defined as those parameter vectors that led to improved performance compared to the mode of the genetic variation distribution, with performance defined as the minimum of relative average yield and relative yield stability. The sample of 500,000 acceptable genotypes that was generated using MCMC is depicted in Fig. 3. Figure 3a shows that the acceptable genotypes included some where yield or yield stability more than doubled, but never both. Apparently there is a trade-off where very high yields cannot be stable across all the sites and years, and vice versa.

The ideotype, defined as the highest performing acceptable genotype, is flagged in the figure. Its performance was 1.94 because of that level of relative average yield, whereas its yield stability value was 1.95.

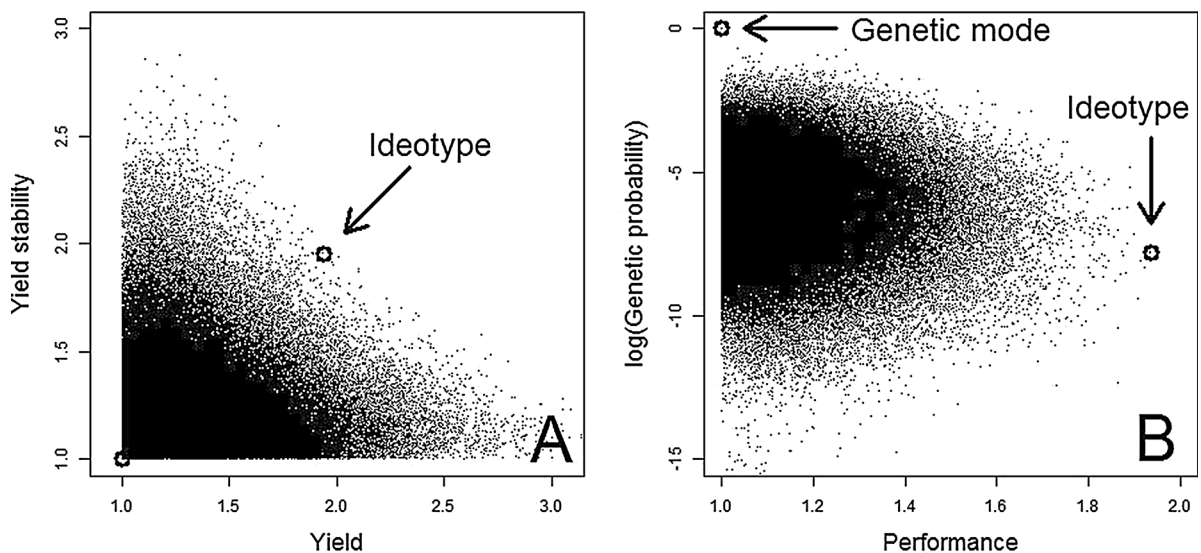
Figure 3b gives an indication of the ‘genetic probability’ of the acceptable genotypes including the ideotype. Genetic probability is calculated as the probability of the parameter vector given the underlying genetic variation distribution. Low values of this probability thus indicate that many plant parameter values are located near the extremes of the genetic variation. The probabilities are normalised with

respect to the mode of the genetic variation distribution. The figure shows that the ideotype has lower genetic probability than most of the other acceptable genotypes.

Figure 4 shows the variation in yields among the five different sites, for each of the six harvests in the three-year rotations that were simulated. The first harvest of the first year tended to be low as swards were being established. The ideotype had higher yields than the genetic variation mode (i.e. the peak of the distribution for genetic variation) for each of the harvests at all sites, but yield increases varied between sites. Harvests that were high using the genetic variation mode, such as the third and fifth harvests at Kvithamar and Særheim, increased less than those that were low to start with. The overall result was a reduction of spatial variation, thus contributing to overall yield stability.

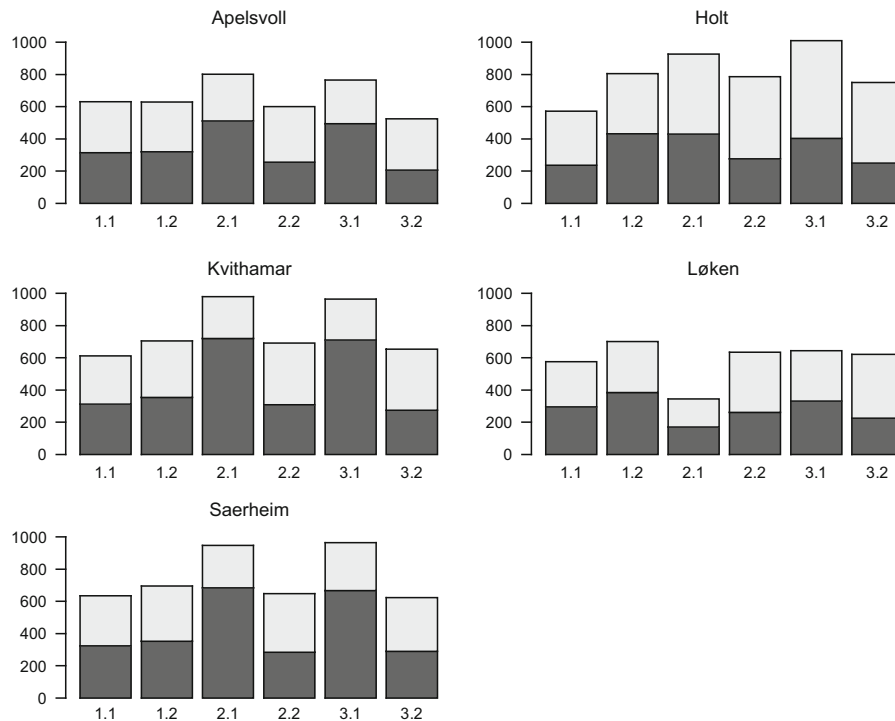
## Partial correlation analysis for acceptable genotypes

For the class of acceptable genotypes, partial correlations (PCC) were calculated between individual parameters and performance (Fig. 5). The absolute PCC-values varied strongly, with the highest values for the parameters PHY, KRDRANAER, LAITIL and LAICR. The KRDRANAER parameter determines



**Fig. 3** A sample of 500,000 acceptable parameter vectors from the genetic variation distribution. *3A* values of the performance traits for each parameter vector. *3B* the logarithm of genetic

probability vs. performance. The ideotype is indicated in both plots, as is the genetic mode (located at point [1,1] in 3A)



**Fig. 4** Yields ( $\text{g m}^{-2}$ ) for each of the six within-rotation harvests, averaged over all three-year long rotations that were simulated (1995–1997, 1998–2000, ..., 2010–2012). 1.1 = first rotation year, first harvest; 1.2 = first rotation year, second

tiller death rate during anaerobic conditions, which may arise in winters during periods of ice encasement. The other three parameters relate to shoot dynamics primarily during the growing season. PHY is the phyllochron, i.e. the thermal time between successive leaf appearances. LAITIL and LAICR govern tillering rate at low and high LAI, respectively.

## Discussion

Bayesian calibration of a PBM for different cultivars

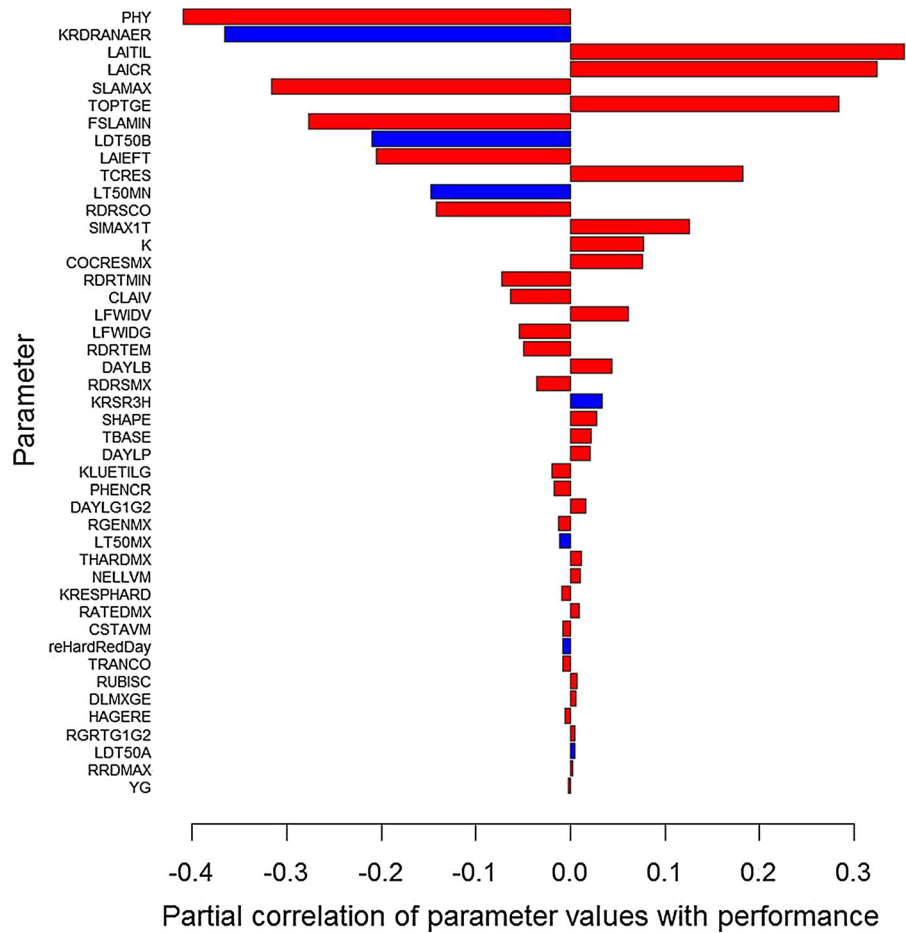
Process-based models (PBMs) offer the possibility to quickly, and at low cost, test different genotypes in a variety of environments, provided the many parameters of such models can be adequately quantified. Here we used Bayesian calibration (BC) to parameterise two cultivars of timothy, using data from five climatically different sites on a total of eleven variables measured in different years. BASGRA, or parts of the

harvest, etc. *Full bars* indicate yields for the ideotype, smaller *black bars* indicate yields for the mode of the genetic variation distribution

model, had been parameterised before (Van Oijen et al. 2005a; Thorsen et al. 2010), but not for cv. Engmo and with much smaller data sets than now available. We carried out the BC using Markov Chain Monte Carlo (MCMC) sampling. Monte Carlo methods such as MCMC have been used fairly often in studies of genetics (Sorensen et al. 1994; Sorensen 2009; Thompson 2014) and metabolism (Jayawardhana et al. 2008), but the models involved in these studies were limited to a small number of parameters. Here all 45 plant parameters of the model were simultaneously estimated. About half the parameters showed clear differences between prior and posterior distribution, for both cultivars, reflecting the rich information content of the data sets that were used (Fig. 1).

BC is increasingly being applied to parameter-rich PBMs (Van Oijen et al. 2005b), but multi-site, multi-output BC as carried out here is still rare. PBMs are more commonly parameterised site-specifically, even when multiple sites are examined in the same study (e.g. Lehuger et al. 2009, but see Reinds et al. 2008 and

**Fig. 5** Partial correlations of individual parameters with performance in the sample of acceptable parameter vectors from the genetic variation distribution. *Red* parameters mainly affecting growing season processes. *Blue* parameters mainly affecting winter survival



Van Oijen et al. 2013 for other examples of multi-site BC of PBMs). The results of the BC in this paper suggest that BASGRA meets the requirements of general model applicability without the need for site-specific parameterisation, at least not for the geographic variation among the five sites in Norway that were considered here. This is evident from the low values of posterior NRMSE for most variables (Tables 2 and 3). Note that the NRMSE-values could have been reduced further if the calibration had been carried out in a less stringent way. This could have been done by reducing the number of simultaneously calibrated variables from eleven to one or two, by calibrating site-specifically or for only one site rather than five, or by using unbounded uniform or Gaussian priors for the parameters rather than bounded beta-distributions. Such more lenient calibration would have increased goodness-of-fit, but would not have shown that the model adequately represents the

underlying processes and that its parameters can be treated as site-invariant—both of which are requirements for the use of the model in ideotype design. Our Bayesian ideotyping procedure thus benefits most from field trials at different sites, on a wide range of genotypes, and with measurements on many more variables than just yield. This will ensure that the majority of model parameters are informed by the data, so that the model can reliably simulate the variety of ways in which performance can be improved.

Genetic variation for plant parameters

We estimated the mean and variance of the genetic variation distribution as those of the union of the cultivar-specific distributions. This simple preliminary approach was used because we had extensive calibration data for only two cultivars. The approach is based on a continuum hypothesis, where we assume

that all parameter vectors covering the posterior distributions for the two cultivars are possible genotypes. More reliable estimation of genetic variation for the plant parameters, and verification of the continuum hypothesis, will require data on a much larger number of genotypes. Such data will also be required to estimate genetic linkages between parameters, which are unidentifiable when only two cultivars are examined.

Once data on more cultivars become available, we should consider replacing the first two steps of our procedure by hierarchical Bayesian calibration (e.g. Banerjee et al. 2012; Condit et al. 2006). In this approach, the plant parameters of all cultivars would be estimated simultaneously, together with hyperparameters representing overall genetic variation and genetic linkage between parameters. A benefit of the approach would be that it would quantify uncertainty about the estimates of genetic variation. Hierarchical approaches that include parameters representing (co)-variance of quantitative traits, are fairly common in evolutionary science, phylogenetics and taxonomy (e.g. Kremer and Le Corre 2012). Hierarchical Bayesian models have also been introduced in breeding science (Meuwissen et al. 2001; Gianola et al. 2009) but only for the linear model of phenotype prediction as the sum of environmental effects and the combined effect of multiple markers, with variance of marker effects represented by unknown parameters that need to be estimated. PBMs used in breeding studies do not assume such separation of additive environmental and genetic effects, and they have, as far as we are aware, not yet been calibrated using hierarchical Bayesian methods. The main obstacle is technical: the computational demand of evaluating PBMs in the high-dimensional hierarchical parameter space covering multiple genotypes. Despite this, the hierarchical approach may become feasible if we restrict the analysis to those parameters that have strong impact on the phenotype. These can be identified from cultivar-specific calibrations as those where the posterior distribution is much narrower than the prior distribution (Fig. 1).

An ideotype for timothy in Norway

To identify the ideotype, we chose as breeding objectives a high average yield and high yield stability across sites in the major Norwegian agricultural

grassland areas. In our study, these objectives implied some degree of cold tolerance as well, because yields were evaluated for 3-year long rotations which included cold winters at some of the sites. Other breeding objectives could have been chosen or different measures for yield stability than the inverse variation coefficient that we used (Annicchiarico 2002). If traits like forage quality, nutrient-use efficiency or disease resistance were to be included as breeding objectives, a more comprehensive model than BASGRA would need to be used. In future work, we plan to extend BASGRA to cover also these aspects of plant performance.

Genotype performance is measured by the degree to which breeding objectives are realised. We defined performance as the minimum of relative yield and relative yield stability, but other definitions such as the mean or a weighted mean could be preferred by breeders. Further, we defined the ideotype as the genotype with maximum performance, irrespective of its genetic probability as measured by proximity to the mode of the genetic variation distribution. With these definitions, the ideotype that was identified had 94 % higher yield than the genetic mode, and 95 % greater yield stability. The analysis suggested a clear trade-off between the two breeding objectives, so breeding exclusively for increasing yields may reduce yield stability. In this analysis, yield stability was defined with respect to both spatial and temporal variability (see “[Deriving an ideotype and selection criteria](#)” section), but the same negative correlation between achievable yield and yield stability was found when only temporal variation was considered.

The performance increase we simulated, with near-doubling of both yield and yield-stability, may be unachievable in practice once additional constraints of genetic linkages between parameters are recognized. This can be examined when information on more cultivars than Engmo and Grindstad becomes available, allowing for the hierarchical approach alluded to above.

Ideotypes should not claim to have universal validity but be aimed at specific breeding objectives (Messina et al. 2009), and any modelling should recognize the population of environments at which the breeding is targeted (Cooper et al. 2014). By evaluating timothy performance on the five selected sites across Norway, we effectively designed a Norway-scale ideotype. That may not represent the best

performing genotype for any given smaller region. For example, an ideotype for the region in South-Western Norway around Særheim may require less cold tolerance than ideotypes for regions with more severe winters. In future work, we plan to compare the Norway-wide ideotype with regional ideotypes.

#### From ideotype to selection criteria

The fourth and final step in our ideotyping procedure, involving the translation of the ideotype into selection criteria of practical use to breeders, was only carried out in a preliminary way. The ideotype consists of specific values for all 45 plant parameters, which is more than can be handled in any breeding programme. Therefore a smaller set of four key parameters contributing most to genotype performance was identified by means of partial correlation analysis. We did not assess the effort involved in measuring the traits that correspond to these parameters, which could in practice hamper their usability in selection (Sinclair 2011), or at least limit their use to the first or last phases in breeding programmes in which the number of genotypes to be evaluated is limited.

Furthermore, the types of selection criteria that can be handled by breeders have changed over time, and genetic tests have become affordable and fast. So modellers may want to translate plant parameters into equivalent genes or collections of genes. This was not attempted, as discussed in the Introduction, because for most parameters the link with DNA sequences or QTLs is still tenuous, with some notable exceptions (Reymond et al. 2003; Yin et al. 2004). Messina et al. (2009) and Cooper et al. (2014) sketch a future of plant breeding in which unravelling the genetic basis of modelled plant properties will support high-throughput phenotyping, which they consider essential for the scaling up of breeding programmes to larger numbers of genotypes. For most plant parameters, however, this will require much progress in elucidating the links between the different levels of organisation that separate genes, cells, plant organs, plant individuals and crops (Tardieu 2003; Messina et al. 2009; Sinclair 2011).

One of the key parameters that was identified, KRDRANAER, relates to winter survival under anaerobic conditions. This result was somewhat surprising because timothy is considered to be more tolerant of severe winter conditions than, for

example, the ryegrasses. The inclusion among the test sites of locations with severe winters may have contributed to the result, notably the northern, coastal location Holt in the region of Troms. In this region, severe winter injury occurred in four of the 18 years studied here according to the official statistics of insurance payment for winter injuries (Landbruksdirektoratet, <https://www.slf.dep.no/no>), with freezing and thawing events leading to ice encasement being a major cause. The other three key parameters influence leaf appearance and tillering rate. These are processes that contribute to tiller recovery and refoliation after the sward has been cut or grazed during summer, or after episodes of winter kill. Notably absent from the shortlist were parameters such as the light extinction coefficient *K* and leaf Rubisco content RUBISC that govern the source-strength of the grassland. The model analysis thus suggests that timothy performance in Norway is more sink- than source-limited.

#### Conclusions and outlook

- A procedure for ideotype breeding involving cultivar-specific Bayesian calibration of process-based models was proposed.
- The method allows for both calibration and subsequent optimisation of all plant parameters simultaneously.
- Results from a preliminary application of the procedure suggest that there is scope for improved yield and yield stability of timothy in Norway, but that there is an inevitable trade-off between the two.
- Parameters that were identified as key contributors to high performance relate to winter survival, leaf appearance and tillering.
- The ideotype identified here was designed for high performance across sites in Norway with widely differing climates. In future work, this nation-wide ideotype will be contrasted with region-specific ones.
- There is a need to assemble detailed data on more genotypes than the two cultivars Engmo and Grindstad. This will facilitate more reliable assessment of genetic variation for the plant parameters than was possible here.

- We plan to develop the procedure further by replacing part of it with hierarchical Bayesian calibration in which genetic variation and genetic linkages appear as hyperparameters.

**Acknowledgments** MvO thanks the staff at NIBIO who arranged for a 1-year sabbatical at their research station in Særheim where this work was carried out. Partial funding also came from the Natural Environment Research Council (NERC) through the MACSUR project. MH's part of the study was funded by the Norwegian Research Council and NIBIO, and is connected to the VarClim project (199664; 'Understanding the genetic and physiological basis for adaptation of Norwegian perennial forage crops to future climates'). We thank two anonymous reviewers whose comments led to considerable improvement of the paper.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Aggarwal PK, Matthews RB, Kropff MJ, Van Laar HH (eds) (1995) Applications of systems approaches in plant breeding. SARP Research Proceedings DLO, Wageningen, p 144
- Annicchiarico P (2002) Genotype  $\times$  Environment Interactions: challenges and opportunities for plant breeding and cultivar recommendations. FAO Plant Production and Protection Paper 174 FAO, Rome
- Banerjee S, Finley AO, Waldmann P, Ericsson T (2012) Hierarchical spatial process models for multiple traits in large genetic trials. *J Am Stat Assoc* 105:506–521
- Breuer L, Eckhardt K, Frede H-G (2003) Plant parameter values for models in temperate climates. *Ecol Model* 169:237–293
- Chenu K, Chapman SC, Tardieu F, McLean G, Welcker C, Hammer GL (2009) Simulating the yield impacts of organ-level quantitative trait loci associated with drought response in maize: a “gene-to-phenotype” modeling approach. *Genetics* 183:1507–1523
- Colon LT, Budding DJ, Keizer LCP, Pieters J (1995) Components of resistance to late blight (*Phytophthora infestans*) in eight South American Solanum species. *Eur J Plant Pathol* 101:441–456
- Condit R, Ashton P, Bunyavejchewin S et al (2006) The importance of demographic niches to tree diversity. *Nature* 313:98–101
- Cooper M, Messina CD, Podlich D, Totir LR, Baumgarten A, Hausmann NJ, Wright D, Graham G (2014) Predicting the future of plant breeding: complementing empirical evaluation with genetic prediction. *Crop Pasture Sci* 65:311–336
- Donald GM (1968) The breeding of crop ideotypes. *Euphytica* 17:385–403
- Dong Z, Danilevskaya O, Abadie T, Messina C, Coles N, Cooper M (2012) A gene regulatory network model for floral transition of the shoot apex in maize and its dynamic modeling. *PLoS ONE* 7:e43450. doi:10.1371/journal.pone.0043450
- Fox A, Williams M, Richardson AD, Cameron D, Gove JH, Quaife T, Ricciuto D, Reichstein M, Tomelleri E, Trudinger CM, Van Wijk MT (2009) The REFLEX project: comparing different algorithms and implementations for the inversion of a terrestrial ecosystem model against eddy covariance data. *Agric For Meteorol* 149:1597–1615
- Gianola D, de los Campos G, Hill WG, Manfredi E, Fernando R (2009) Additive genetic variability and the Bayesian alphabet. *Genetics* 183:347–363
- Hammer G, Cooper M, Tardieu F, Welch S, Walsh B, Van Eeuwijk F, Chapman S, Podlich D (2006) Models for navigating biological complexity in breeding improved crop plants. *Trends Plant Sci* 11:587–593
- Höglind M, Schapendonk AHCM, Van Oijen M (2001) Timothy growth in Scandinavia: combining quantitative information and simulation modelling. *New Phytol* 151:355–367
- Höglind M, Hanslin HM, Van Oijen M (2005) Timothy regrowth, tillering and leaf area dynamics following spring harvest at two growth stages. *Field Crops Res* 93:51–63
- Höglind M, Jørgensen M, Østrem L (2006) Growth and development of frost tolerance in eight contrasting cultivars of timothy and perennial ryegrass during winter in Norway. *Proceedings of NJF Seminar* 384 10–12 August 2006, Akureyri, Iceland: 50–53
- Höglind M, Bakken AK, Jørgensen M, Østrem L (2010) Tolerance to frost and ice encasement in cultivars of timothy and perennial ryegrass during winter. *Grass Forage Sci* 65:431–445
- Jayawardhana B, Kell DB, Rattray M (2008) Bayesian inference of the sites of perturbations in metabolic pathways via Markov chain Monte Carlo. *Bioinformatics* 24:1191–1197
- Kennedy MC, O'Hagan A (2001) Bayesian calibration of computer models. *J Roy Stat Soc B* 63:425–464
- Kremer A, Le Corre V (2012) Decoupling of differentiation between traits and their underlying genes in response to divergent selection. *Heredity* 108:375–385
- Landivar JA, Baker DN, Jenkins JN (1983) Application of GOSSYM to genetic feasibility studies II Analyses of increasing photosynthesis, specific leaf weight and longevity of leaves in cotton. *Crop Sci* 23:504–510
- Lehuger S, Gabrielle B, Van Oijen M, Makowski D, Germon J-C, Morvan T, Hénault C (2009) Bayesian calibration of the nitrous oxide emission module of an agro-ecosystem model. *Agric Ecosyst Environ* 133:208–222
- Levy PE, Wendler R, Van Oijen M, Cannell MGR, Millard P (2004) The effects of nitrogen enrichment on the carbon sink in coniferous forests: uncertainty and sensitivity analyses of three ecosystem models. *Water Air Soil Pollut Focus* 4:67–74
- Marshall DR (1991) Alternative approaches and perspectives in breeding for higher yields. *Field Crops Res* 26:171–190
- Medlyn BE, Badeck F-W, De Pury DGG et al (1999) Effects of elevated [CO<sub>2</sub>] on photosynthesis in European forest species: a meta-analysis of model parameters. *Plant Cell Environ* 22:1475–1495



- Messina C, Hammer G, Dong Z, Podlich D, Cooper M (2009) Modelling crop improvement in a  $G \times E \times M$  framework via gene-trait-phenotype relationships. In: Sadras V, Calderini D (eds) Crop physiology: applications for genetic improvement and agronomy. Elsevier, The Netherlands, pp 235–265
- Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equation of state calculations by fast computing machines. *J Chem Phys* 21:1087–1092
- Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157:1819–1829
- Ogle K (2009) Hierarchical Bayesian statistics: merging experimental and modeling approaches in ecology. *Ecol Appl* 19:577–581
- Piepho HP, Möhring J, Melchinger AE, Büchse A (2008) BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica* 161:209–228
- Reinds GJ, Van Oijen M, Heuvelink GBM, Kros H (2008) Bayesian calibration of the VSD soil acidification model using European forest monitoring data. *Geoderma* 146:475–488
- Reymond M, Muller B, Leonardi A, Charcosset A, Tardieu F (2003) Combining quantitative trait loci analysis and an ecophysiological model to analyze the genetic variability of the responses of maize leaf growth to temperature and water deficit. *Plant Physiol* 131:664–675
- Rodríguez D, Van Oijen M, Schapendonk AHCM (1999) LINGRA\_CC: a sink-source model to simulate the impact of climate change and management on grassland productivity. *New Phytol* 144:359–368
- Schapendonk AHCM, Stol W, Van Kraalingen DWG, Bouman BAM (1998) LINGRA, a sink/source model to simulate grassland productivity in Europe. *Eur J Agron* 9:87–100
- Semenov MA, Stratonovitch P, Alghabari F, Gooding MJ (2014) Adapting wheat in Europe for climate change. *J Cereal Sci* 59:245–256
- Sinclair TR (2011) Challenges in breeding for yield increase for drought. *Trends Plant Sci* 16:289–293
- Sorensen D (2009) Developments in statistical analysis in quantitative genetics. *Genetica* 136:319–332
- Sorensen DA, Wang CS, Jensen J, Gianola D (1994) Bayesian analysis of genetic change due to selection using Gibbs sampling. *Genet Sel Evol* 26:333–360
- Sunde M (1996) Effects of winter climate on growth potential, carbohydrate content and cold hardiness of timothy (*Phleum pratense* L) and red clover (*Trifolium pratense* L). Agricultural University of Norway, Ås, Doctor Scientrum theses 1996: 17
- Tardieu F (2003) Virtual plants: modelling as a tool for the genomics of tolerance to water deficit. *Trends Plant Sci* 8:9–14
- Thompson EA (2014) A journey with statistical genetics. In Lin X et al. (Eds), COPSS 50th anniversary volume: past, present and future of statistical science. CRC Press, Boca Raton, 451–464
- Thorsen SM, Roer A-G, Van Oijen M (2010) Modelling the dynamics of snow cover, soil frost and surface ice in Norwegian grasslands. *Polar Res* 29:110–126
- Van Oijen M (1992) Evaluation of breeding strategies for resistance and tolerance to late blight in potato by means of simulation. *Neth J Plant Pathol* 98:3–11
- Van Oijen M, Cannell MGR, Levy PE (2004) Modelling biogeochemical cycles in forests: state of the art and perspectives. In: Andersson F, Birot Y, Päivinen R (eds) Towards the sustainable use of European forests. EFI, Joensuu, Finland, pp 157–169
- Van Oijen M, Höglind M, Hanslin HM, Caldwell N (2005a) Process-based modelling of timothy regrowth. *Agron J* 97:1295–1303
- Van Oijen M, Rougier J, Smith R (2005b) Bayesian calibration of process-based forest models: bridging the gap between models and data. *Tree Physiol* 25:915–927
- Van Oijen M, Reyer C, Bohn FJ, Cameron DR, Deckmyn G, Flechsig M, Härkönen S, Hartig F, Huth A, Kiviste A, Lasch P, Mäkelä A, Mette T, Minunno F, Rammer W (2013) Bayesian calibration, comparison and averaging of six forest models, using data from Scots pine stands across Europe. *For Ecol Manag* 289:255–268
- Van Oijen M, Höglind M, Cameron DR, Thorsen SM (2015) BASGRA\_2014. [10.5281/zenodo.27867](https://doi.org/10.5281/zenodo.27867)
- Wang Y-P, Trudinger CM, Enting IG (2009) A review of applications of model-data fusion to studies of terrestrial carbon fluxes at different scales. *Agric For Meteorol* 149:1829–1842
- Xu L, Henke M, Zhu J, Kurth W, Buck-Sorlin G (2011) A functional-structural model of rice linking quantitative genetic information with morphological development and physiological processes. *Ann Bot* 107:817–828
- Yin X, Kropff MJ, Stam P (1999) The role of ecophysiological models in QTL analysis: the example of specific leaf area in barley. *Heredity* 82:415–421
- Yin X, Chasalow S, Dourleijn CJ, Stam P, Kropff MJ (2000) Coupling estimated effects of QTLs for physiological traits to a crop growth model: predicting yield variation among recombinant inbred lines in barley. *Heredity* 85:539–549
- Yin X, Struik PC, Kropff MJ (2004) Role of crop physiology in predicting gene-to-phenotype relationships. *Trends Plant Sci* 9:426–432
- Zhang D-Y, Sun G-J, Jiang X-H (1999) Donald's ideotype and growth redundancy: a game theoretical analysis. *Field Crops Res* 61:179–187
- Zheng B, Biddulph B, Li D, Kuchel H, Chapman S (2013) Quantification of the effects of VRN1 and Ppd-D1 to predict spring wheat (*Triticum aestivum*) heading time across diverse environments. *J Exp Bot*. doi:[10.1093/jxb/ert209](https://doi.org/10.1093/jxb/ert209)