NERC Open Research Archive
sharing our research

Centre for Ecology & Hydrology
NATURAL ENVIRONMENT RESEARCH COUNCIL

Article (refereed) - postprint

Persson, T.; Höglind, M.; Van Oijen, M.; Korhonen, P.; Palosuo, T.; Jégo, G.; Virkajärvi, P.; Bélanger, G.; Gustavsson, A.-M. 2019. **Simulation of timothy nutritive value: a comparison of three process-based models**.

© 2018 Elsevier B.V.
This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/

This version available http://nora.nerc.ac.uk/521729/

NERC has developed NORA to enable users to access research outputs wholly or partially funded by NERC. Copyright and other rights for material on this site are retained by the rights owners. Users should read the terms and conditions of use of this material at http://nora.nerc.ac.uk/policies.html#access

NOTICE: this is the authors' version of a work that was accepted for publication in *Field Crops Research*. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in ***Field Crops Research*** **(2019), 231. 81-92. https://doi.org/10.1016/j.fcr.2018.11.008**

www.elsevier.com/

Contact CEH NORA team at
noraceh@ceh.ac.uk

The NERC and CEH trademarks and logos ('the Trademarks') are registered trademarks of NERC in the UK and other countries, and may not be used without the prior written consent of the Trademark owner.

[This paper has been published in: Field Crops Research 231 (2019): 81-92]

Simulation of timothy nutritive value: A comparison of three process-based models

Persson, T.[1*], Höglind, M.[1], Van Oijen, M.[2], Korhonen P.[3], Palosuo, T.[3], Jégo, G.[4], Virkajärvi, P.[3], Bélanger, G.[4], Gustavsson, A.-M.[5]

[1]Norwegian Institute of Bioeconomy Research (NIBIO), Særheim, Norway

[2]Centre for Ecology and Hydrology, UK

[3]Natural Resources Institute Finland (Luke), Helsinki, Finland

[4]Agriculture and Agri-Food Canada (AAFC), Québec, Canada

[5]Swedish University of Agricultural Sciences (SLU), Umeå, Sweden

[*]Corresponding author

E-mail: tomas.persson@nibio.no

Abstract

Different forage grass models are used to simulate forage yield and nutritive attributes, but these models are seldom compared, particularly those for timothy (*Phleum pratense* L.), a widely grown forage grass species in agricultural regions with a cold temperate climate. We compared the models BASGRA, CATIMO and STICS for their predictions of timothy crude protein (CP) concentration, neutral detergent fibre (NDF) concentration and NDF digestibility (dNDF), three important forage nutritive attributes. Data on CP and NDF concentrations, and dNDF and the associated weather and soil data for seven cultivars, taken from eight field experiments in Canada, Finland, Norway, and Sweden, were divided into calibration and validation datasets. Model parameters were estimated for each cultivar separately (cultivar-specific calibration) and for all cultivars together (generic calibration), using different methods for the three models. Normalized root mean square error (RMSE) in prediction of CP concentration varied between 16 and 26 % for BASGRA, 45 and 101 % for CATIMO and 27 and 45 % for STICS across the two calibration methods and the calibration and validation datasets. Normalised RMSE in prediction of NDF concentration varied between 8 and 13 % for BASGRA, 14 and 21 % for CATIMO and 8 and 12 % for STICS, while for dNDF it varied between 7 and 22 % for BASGRA, 7 and 38 % for CATIMO and 5 and 6 % for STICS. Cultivar-specific calibration improved the performance of CATIMO and STICS, but not BASGRA, compared with generic calibration. The prediction accuracy for NDF concentration and dNDF with the three models was within the same range or better than that for forage dry matter (DM) yield of timothy. Overall, the three models performed well in predicting some nutritive attributes and yield in Northern Europe and Canada, but improvements are required, particularly to increase the prediction accuracy of CP concentration.

Key words: BASGRA, CATIMO, crude protein, fibre, forage grass, grassland, NDF, dNDF, STICS

1. Introduction

Forage grasses serve as the main source of energy and nutrients for ruminant livestock, including dairy cows, beef cattle, sheep and goats, in many regions of the world (Fulkerson et al., 2007; Thornton, 2010). Hence, management for optimal energy and nutrient content in forage crops is crucial to these animals. Feed evaluation for ruminants usually takes into account the digestibility and protein concentration of the forage (Bruinenberg et al., 2002). Because the cell contents are almost totally digestible, the concentration of cell walls is crucial to the nutritive value of forages (Buxton, 1996). Typically measured indicators of forage nutritive value are the concentration and digestibility of neutral detergent fibre (NDF) and the crude protein (CP) concentration. The development of stem and inflorescence on reproductive tillers generally lowers the nutritive value of the forage, because these plant parts are less digestible than leaves (Chapman et al., 2014; Elgersma and Søegaard, 2018). However, as the forage grass sward grows and develops more reproductive tillers, the total aboveground biomass also increases causing a negative relationship between nutritive value and dry matter yield (Wilkinson and Rinne, 2018).

In Northern Europe and Canada, perennial forage grasses grown for intensive dairy production are usually cut and harvested 2-3 times per growing season and conserved as silage (Höglind et al., 2005; Casler and Kallenbach, 2007; Jing et al., 2012). Timothy (*Phleum pratense* L.) is one of the most widely used forage grass species in cold-temperate regions of the world (Wilkinson and Rinne, 2018), where it is grown under a wide range of climate, soil and management conditions. This species exhibits slower development, expressed as growing degree-days from the start of the growing season until the start of anthesis, than many other cold temperate forage grasses (Pontes et al., 2007). Comparisons with other forage grasses in the same environment also show that timothy has high (Pontes et al., 2007) to intermediate (Jensen et al., 2016) CP concentration and digestibility. The DM yield and nutritive value of the timothy vary with growing conditions and management practices, such as cutting and fertilisation regimes (Bélanger et al., 2001). In addition, the relationship between development, growth and nutritive value varies between timothy cultivars (Jokela et al., 2015). Length of the growing season, temperature and precipitation patterns during the growing season and conditions in the previous winter are particularly important for growth, yield development and management of this species.

Process-based simulation models for forage grass (e.g. Bonesmo and Belanger, 2002a; Wu et al., 2007; Köchy, 2008; Chang et al., 2013; Jégo et al., 2013; Vital et al., 2013; Höglind et al., 2016) seek to represent the physiological processes behind sward growth and development. However, the representation of processes such as water and nutrient uptake, carbon (C) assimilation and carbohydrate allocation and transfer between plant compartments varies between models (Kipling et al., 2016; Sándor et al., 2017). Previous studies showed different responses in gross primary production (Sándor et al., 2016), biomass (Hurtado-Uria et al., 2013; Sándor et al., 2017; Ehrhardt et al., 2018) and N_2O emissions (Ehrhardt et al., 2018) for different grassland models when compared under various environmental conditions. As for timothy, Korhonen et al. (2018) compared three models for their ability to predict DM yield in Northern Europe and Canada. However, to our knowledge, there are no other published comparisons of the ability of forage grass models to predict nutritive value.

The underlying processes explaining the yield and nutritive value in forage grasses are arguably more complex than those explaining only DM production. In particular, as pointed out by Virkajärvi et al. (manuscript under preparation), models of forage grasses differ considerably in how they handle plant processes related to plant N requirements and cell wall formation and content. A comparison of the ability of forage grass models to predict nutritive value in field experiments could provide knowledge about the utility of these models under different weather, soil, cutting and fertiliser management conditions. Such knowledge could help select prediction models for different conditions, in quantifying uncertainty in model predictions under different conditions, and in identifying potential improvements in the representation of physiological processes in different models.

In this study, the ability of three simulation models (BASGRA, CATIMO and STICS) to predict three key nutritive attributes [CP concentration, NDF concentration and the digestibility of NDF (dNDF)] in timothy in a wide range of environments representing the main regions where timothy is grown in the northern hemisphere was compared. In addition, we tested two different calibration strategies: generic and cultivar-specific.

2. Materials and methods

2.1. Model descriptions

The BASGRA, CATIMO, and STICS models simulate the growth and the development of the primary growth of timothy and its first regrowth as a function of the weather, soil conditions, and management practices, with a daily time step. In all three models, accumulation of biomass is based on the concept of radiation use efficiency where intercepted solar radiation is converted into biomass.

2.1.1. BASGRA

The Basic Grassland (BASGRA) model (Höglind et al., 2016) is a further development of the LINGRA model, which was initially developed to simulate perennial ryegrass (Schapendonk et al., 1998) and later adapted to timothy (Höglind et al., 2001). In BASGRA, the plant is divided into stem, leaf, stubble, root and reserve compartments. The model is based on the source-sink concept. The source tissue, with net export of photosynthetic assimilates, consists of developed leaves, other photosynthetic tissues and carbohydrate reserves. The sink tissue, with net import of photosynthetic assimilates, comprises newly developed plant parts and roots. Sward development is driven by air temperature and day length. Carbohydrate reserves are used for producing new leaf tissue at the start of the growing season or after defoliation when there is little source tissue. Equations to simulate soil and plant N and forage nutritive value have recently been developed (Höglind et al., manuscript under preparation). The soil is described as one single homogeneous layer. Plant-available water in the soil is set as the difference between the water content at field capacity and the water content at wilting point. The soil water content is affected by infiltration, soil surface evaporation and run-off, water uptake by plants and percolation of water above field capacity, simulated using the tipping-bucket method. Soil C is divided into three pools with different residence times, i.e. litter, soil organic matter with a fast decomposition rate and soil organic matter with a slow decomposition rate. Soil N is divided into four pools: similar litter and soil organic matter pools as used for C, plus a pool of mineral N. Nitrogen can flow between these pools through decomposition, mineralisation and immobilisation processes, which are all affected by soil temperature. Nitrogen is added to the litter pool by shoot senescence, while root senescence adds N to the fast-decomposing soil organic pool. Decomposition of organic N adds to the soil mineral N pool, which is depleted by leaching, emissions of nitrous oxide and nitrogen oxide, and plant N uptake. Nitrogen leaching is driven by the rate of water drainage which, in turn, is affected by soil hydraulic properties and infiltration, transpiration and evaporation. Nitrogen emissions increase with availability of mineral N. The soil N functions are obtained from the forest model BASFOR (Van Oijen et al., 2005).

Sub-optimal plant N status affects the shoot C sink strength and thus shoot growth. Tillering rate also depends on the plant N status. Plant N availability is the sum of soil N that is available for plant uptake and plant N that is available for remobilisation within the above-ground plant parts. The latter is the amount of N above an optimal N concentration profile that follows the light extinction profile from the top to the bottom of the canopy, which is allocated to growing plant tissue. Consequently, the optimal N concentration decreases as more light is extinguished through the canopy as it grows. The nitrogen-carbon ratio in the roots is constant. The plant CP concentration is the N concentration multiplied by 6.25. The fraction of cell walls in the biomass, as expressed by the NDF concentration, is allowed to differ between leaves and stems, and increases with phenological stage (Bélanger and McQueen, 1999; Nordheim-Viken et al., 2009), but is not directly affected by temperature or N concentration. The digestibility of the cell wall (dNDF) of both leaves and stems decreases with phenological stage (Bélanger and McQueen, 1999; Nordheim-Viken et al., 2009). The digestible fraction of the cell wall is assumed to be the same in all plant components. In stubble, the cell wall fraction is set at 100%, whereas there is no cell wall fraction in the reserves. The digestibility of the cell content is set at 100%.

2.1.2. CATIMO

The Canadian Timothy Model (CATIMO) was developed to simulate the growth of timothy, including N processes (Bonesmo and Bélanger, 2002a) and fibre concentration and digestibility (Bonesmo and Bélanger, 2002b). The model allocates biomass into green leaves, stems and roots. Similarly to BASGRA, a portion of the biomass that is allocated to the above-ground compartments is allocated to reserves, which is used to form new tissue after defoliation or winter. The light-driven biomass growth is decreased under sub-optimal soil water, plant N, photosynthetically active radiation (PAR) and temperature conditions. The potential radiation use efficiency, which determines growth when there are no limiting factors, is constant throughout the growing season. The soil hydraulic properties and C and N content are simulated in one homogeneous layer. The N stress is estimated from an index of N nutrition that is calculated as the ratio of N concentration to the critical N concentration for a given biomass (Bélanger and Gastal, 2000). Plant N uptake is determined by crop demand and soil N supply. The soil N supply is estimated from soil mineral N content and N mineralisation. The N demand is defined as the difference between actual and maximum N concentration, with the latter decreasing with increasing sward biomass using an N dilution curve. The plant CP concentration is calculated by multiplying the N concentration by 6.25. For simulation of forage digestibility, the sward is considered to consist of green leaves, dead leaves and stems including leaf sheaths, each with their own NDF concentration and digestibility. The dry matter (DM) digestibility of the sward is calculated by combining the DM digestibility of green leaves, dead leaves, and stems with their respective weight. The DM digestibility of the cellular content of green leaves and stems is set at 0.98 g g^{-1} DM. Dead leaves are assumed to have a NDF concentration of 1.0 g g^{-1} DM, with a DM digestibility of 0.70 g g^{-1} DM. The NDF concentration of green leaves and stems is obtained by integrating the proportion of the respective daily growth rates partitioned to cell wall, the daily rates of conversion of cellular contents into cell wall and the daily death rate of leaves. The dNDF of green leaves and stems is determined from an initial maximum value and a daily rate of decrease related to daily mean temperature. Both temperature and N stress are taken into account in simulation of the NDF concentration and dNDF of green leaves and stems.

2.1.3. STICS

The multidisciplinary simulator for standard crops (Simulateur mulTIdisciplinaire pour les Cultures Standard, STICS) is a model for simulation of agricultural crops and cropping systems (Brisson et al., 1998, 2008). It has an add-on module for timothy, including N and nutritive value-related functions (Jégo et al., 2013). The potential radiation use efficiency, setting the growth under non-limiting conditions, varies between juvenile, vegetative and reproductive phenological phases. The model simulates soil water dynamics and C and N processes in a multi-layer profile. Plant N demand is driven by the N dilution curve concept for calculating the N requirements of the plants (Bélanger and Gastal, 2000). In the STICS model, the optimal crop N uptake is described using the relationship between the critical N concentration and total biomass. The critical N concentration (Nc, % N per DM unit) is defined as the lowest plant N concentration required for maximum growth. As most crops can take up more N than is needed for optimum growth, a maximum N concentration curve is also required in STICS, but no additional biomass growth occurs for N uptake between the critical and maximum N concentrations. The effective total N uptake rate is limited either by the crop N demand or by the soil N availability. Plant metabolism is affected when the total N concentration is below the critical concentration for a given biomass defined by the critical N concentration curve. Functions to calculate NDF concentration and digestibility are from CATIMO (described briefly above).

2.2. Crop data

Data on timothy from experimental sites at Fredericton (45°55′N; 66°32′W; 35 m asl), Lacombe (52°28′N; 113°44′W; 860 m asl) and Québec (46°47′N; 71°07′W; 75 m asl) in Canada; Maaninka (63°09′N; 27°17′E; 90 m asl), Rovaniemi (66°35′N; 26°01′E; 106 m asl) and Ruukki (64°40′N; 25°06′ E; 48 m asl) in Finland; Særheim (58°46′N; 5°39′E; 90 m asl) in Norway; and Umeå (63°45′N; 20°17′E; 12 m asl) in Sweden were used for model calibration and validation (Fig. 1). Data were from the spring growth before and during the first cut and the summer regrowth after the first cut until the second cut. They covered forage dry matter (DM) yield, DM yield of stems and leaves, leaf area index, tiller density, water-soluble carbohydrate concentration and nutritive attributes (CP concentration, NDF concentration, dNDF, ash concentration, digestible DM and digestible organic matter). Data were not available for all experimental sites and years (see Table 1). The dataset was divided into combinations of sites, years, cultivars and management regimes, with each unique combination called a "treatment" according to Korhonen et al. (2018). In total, there were 101 treatments. Thirty-three of the treatments were previously used in Korhonen et al. (2018) whereas the remaining 68 treatments have not been used in any other previous modelling study.

The methods used to measure nutritive value varied slightly between the locations. Nitrogen concentration was analysed using a standard Kjeldahl method at Rovaniemi (Nissinen et al., 2010) and Umeå (Gustavsson and Martinsson, 2004), near-infrared reflectance spectroscopy

(NIRS) at Maaninka, Ruukki and Særhiem (Marum, 1990). The NDF concentration at Fredericton, Lacombe, and Québec was determined using a combination of chemical and NIRS methods. At Fredericton, the NDF analyses were based on methods by Van Soest et al. (1991) without using sodium sulphite, while at Lacombe and Québec the analyses were carried out using an Ankom Fiber Analyzer. At Maaninka, Ruukki and Særheim, the NDF concentration was analysed using NIRS and at Umeå using an ANKOM filter bag technique. The NDF digestibility at Fredericton and Québec was analysed from rumen contents using a method described by Van Soest et al. (1966).

Table 1.

Figure 1.

2.3 Weather and soil data

Daily weather data on minimum and maximum temperature, precipitation, global solar radiation, wind speed and relative air humidity were obtained from weather stations near the experimental sites. The data for Fredericton, Québec and Lacombe were obtained from Environment Canada (http://climat.meteo.gc.ca/historical_data/search_historic_data_e.html), those for Maaninka, Rovaniemi and Ruukki from the Finnish Meteorological Institute, those for Særheim from the Agrometeorology Norway network (http://lmt.nibio.no/) and those for Umeå from the Swedish Meteorological and Hydrological Institute (SMHI) (www.smhi.se). The soil input data comprised texture fractions, bulk density, soil organic material and pH. Soil hydraulic characteristics, including water content at permanent wilting point, field capacity and saturation, which are input to all three models, were either measured or estimated based on available data on soil properties at each site.

2.4. Model calibration and validation

The dataset was divided into calibration and validation data by randomly selecting one treatment for model evaluation from each treatment type from sites with more than two treatments or years except for Rovaniemi, for which no nutritive value data were used in this study (Table 2). Differences in nutritive attribute data availability between sites, geographical location and differences in climate and soil conditions and management practices among sites were taken into account in this division.

Table 2.

Two types of calibrations were conducted for each model. In one, parameters were calibrated using data for each cultivar separately (cultivar-specific calibration). In the other, a common set of parameter values representing all cultivars was obtained by using the data for all cultivars together (generic calibration). The division between calibration and validation datasets was the same for the two calibration types. In the two calibrations, each model was calibrated using model-specific methods. BASGRA and CATIMO were calibrated using Bayesian techniques (Van Oijen et al., 2005). For BASGRA, a prior probability distribution was first defined for each parameter to be calibrated, which was then updated using the observed data, which included nutritive value data as well as observations of biomass, and biomass-related data such as leaf area index, specific leaf area and tiller density. For the BASGRA calibration, beta prior distributions were used for all calibration parameters (Table S1). The prior parameter distribution for most parameters was set using information from a previous calibration for the cultivar Grindstad in the LINGRA model (the predecessor of BASGRA), in which timothy data from the Nordic region of Europe were used (Persson et al., 2014). For those parameters relating to nutritive value that were introduced into BASGRA later, the prior probability distribution was set within a wide, yet plausible, range with the help of literature information and preliminary calibrations. The BASGRA calibration was carried out by sampling from the posterior distribution using the Metropolis algorithm and a chain length of 350 000. A likelihood function by Sivia (2006), which is more robust to outliers than the Gaussian distribution, was used in the calibration. For CATIMO, the prior probability distribution of parameters was obtained from a previous calibration (Korhonen et al., 2018) for the same cultivars as used in this study except for cv. Nuutti, for which the prior probability distribution was set based on cv. Tammisto II. The posterior sampling chain length for the Metropolis algorithm was 150 000 iterations for Grindstad, Champ, Climax, Jonatan and generic, 200 000 iterations for Nuutti and 250 000 iterations for Tammisto II and Iki. For both BASGRA and CATIMO, the maximum *a posteriori* (MAP) vectors from the calibration were used to evaluate the models, not the whole posterior distribution, since uncertainty quantification was not within the scope of this study. In the STICS and CATIMO calibrations, only the parameters involved in calculation of NDF concentration and dNDF were calibrated. For STICS, parameters calibrated in Korhonen et al. (2018) were used to simulate plant growth and N uptake except for cv. Nuutti, for which a new calibration was required since this cultivar was not included in the previous study. The parameters of the maximum and critical N dilution curves used in this study were those defined by Jégo et al. (2013). These parameters were not calibrated, because in STICS they are supposed to be common to all cultivars of the same species and because it is not recommended to calibrate them directly in the model. If it is considered necessary to define new parameters, then this should be done in a separate study following the methodology proposed by Justes et al. (1994). All parameters used to calculate NDF concentration were calibrated simultaneously by minimising the sum of squared differences between measured and simulated NDF values. Two optimisation algorithms available in the Flexible Modelling Environment (FME) package in R were used. The two-step method was used, to avoid the problem of local minima. For both steps, calibrated values were constrained in a calibration range. In the first step, the pseudo algorithm, a pseudo-random search algorithm, was used with the maximum number of iterations (1000). A second algorithm (L-BFGS-B; constrained quasi-Newton method) was then used to refine the calibration. The dNDF parameters were calibrated using the same method.

2.5. Statistical analysis

The prediction accuracy of the observed CP concentration, NDF concentration and dNDF was evaluated with the root mean square error (RMSE):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_i - O_i)^2}{n}} \qquad (1)$$

where n is the number of observations and P_i and O_i are the predicted and observed values for each data pair. The closer the RMSE is to 0, the better the agreement. The RMSE was divided by the mean of the observed values (normalised RMSE) to allow comparison of the prediction accuracy among different nutritive attributes. In addition, predictions were evaluated with the relative mean bias error (rMBE):

$$rMBE = \frac{\sum_{i=1}^{n} P_i - O_i}{\sum_{i=1}^{n} O_i} \qquad (2)$$

The rMBE provides a measure of the relative magnitude of over- or under-estimation of the nutritive attributes. Willmott’s index of agreement (d-index) was also used to evaluate the model predictions:

$$d = 1 - \left[\frac{\sum_{i=1}^{n}(P_i - O_i)^2}{\sum_{i=1}^{n}\left(|P_i'| + |O_i'|\right)^2}\right] \qquad (3)$$

where P_i' and O_i' are the means of the predicted and observed values and the closer d is to 1, the better the agreement between observed and simulated values. According to Willmott (1981), d-index should be used to show the agreement between observations and predictions in a dimensionless way, as a complement to the RMSE. Observed and simulated pairs of nutritive attributes were also plotted against the amount of N applied per cut, mean annual temperature and accumulated annual precipitation, to identify any trends in prediction accuracy across the environmental variability within the calibration and validation datasets.

3. Results

3.1. Cultivar-specific calibration and validation

Predictions of CP concentration with BASGRA had a lower normalised RMSE (19 %) than those predicted by CATIMO and STICS (50 % and 40 %, respectively) in the cultivar-specific calibration (Fig. 2; Table 3). Both STICS and BASGRA had a lower normalised RMSE (24 % and 26 %, respectively) than CATIMO (45 %) in the cultivar-specific validation (Table 4). BASGRA under-estimated observed CP concentrations (rMBE = -6 %) in the cultivar-specific calibration (Table 3; Fig. 2), due to under-estimation of high CP concentrations, whereas CATIMO and STICS over-estimated the observed CP concentrations (rMBE = +19 % and +29 % respectively) (Table 3), mostly because of over-estimation of high CP concentrations. In the cultivar-specific calibration with data from several locations and cultivars, BASGRA and STICS predicted the

NDF concentration with lower normalised RMSE (13 % and 8 %, respectively) and greater d-index (0.59 and 0.75, respectively) than CATIMO (21 % and 0.43, respectively) (Fig. 3; Table 3). For the cultivar-specific validation, however, there were no clear differences between the three models in their ability to predict NDF concentration (Fig. 3; Table 4). The NDF concentration was slightly under-estimated by all three models in the cultivar-specific calibration (Table 3) and validation (Table 4). This under-estimation tended to be greater for BASGRA (rMBE = -9.0 %) than for CATIMO and STICS (rMBE = -0.2 % and -0.4 %, respectively) in the cultivar-specific calibration. CATIMO and STICS predicted dNDF with lower normalised RMSE than BASGRA in the cultivar-specific calibration (10 %, 6 % and 22 %, respectively) (Table 3) and the cultivar-specific validation (7 %, 5 % and 10 %, respectively) (Table 4). STICS under-estimated and CATIMO over-estimated dNDF in both the cultivar-specific calibration and validation, while BASGRA slightly over-estimated dNDF in the cultivar-specific calibration and under-estimated it in the cultivar-specific validation. However, the over-estimation in the cultivar-specific calibration with BASGRA was greatly influenced by a large error in one single measurement (Fig. 4), so it can be assumed that BASGRA under-estimated dNDF in both calibration and validation.

Figure 2.

Figure 3.

Figure 4.

Table 3.

Table 4.

When plotted against the amount of N applied per cut, there was a slight increase in both observed and simulated CP concentrations with increasing N level. However, CATIMO and STICS tended to over-estimate CP concentration. For CATIMO, this trend was more noticeable at high than at low N fertiliser levels (Fig. 5; Fig. 6). There were no clear trends in the predictability of NDF concentration and dNDF across N fertiliser levels for any of the three models (not shown). Moreover, it was not possible to discern any trends in predictability among climate conditions when the three observed and simulated nutritive attributes were plotted against mean annual air temperature and mean annual accumulated precipitation (data not shown).

Figure 5.

Figure 6.

There was no clear trend between N fertilizer level and DM yield, possibly because many of the measurements of dry matter and nutritive value were taken in between normal cutting times. To further analyse the underlying mechanisms for the differences in the prediction accuracy of CP concentration, simulated CP concentrations with the three models were plotted against simulated dry matter yield and, while observed CP concentrations were plotted against observed dry matter yield for the cultivar specific and generic calibration (Fig. 7). The plotted relationships indicate that CATIMO, and in some cases the other two models, simulated higher CP concentration than what was observed at a similar dry matter yield, especially at low dry matter yields.

3.2. Generic calibration and validation

The prediction accuracy across the three models in the generic calibration and validation followed the same pattern as the cultivar-specific calibration and validation. The prediction accuracy of CP concentration in the generic calibration was higher for BASGRA (normalised RMSE = 16 %, d-index = 0.89) and STICS (normalised RMSE = 38 %, d-index = 0.92) than for CATIMO (normalised RMSE = 101 %, d-index = 0.36) (Table 5). Similar differences in prediction accuracy between the three models were obtained with the validation dataset (Table 6). In the calibration (Table 5) and validation datasets (Table 6), the observed CP concentration was slightly under-estimated by BASGRA, over-estimated by STICS, and greatly over-estimated by CATIMO, based on rMBE. The prediction accuracy for NDF concentration and dNDF was also higher for BASGRA and STICS than for CATIMO in the calibration (Tables 5) and validation (Table 6), as indicated by lower normalised RMSE. However, prediction of NDF concentration had a lower d-index with STICS than with BASGRA (Tables 5 and 6). The NDF concentration was slightly under-estimated by CATIMO in the calibration and validation datasets, whereas it was slightly under-estimated in the calibration dataset and over-estimated in the validation dataset by BASGRA. CATIMO and to a lesser degree BASGRA under-estimated dNDF in the calibration and validation datasets. STICS slightly over-estimated NDF concentration and dNDF in both datasets (Tables 5 and 6). The generic calibration of BASGRA resulted in slightly better predictions of the three nutritive attributes than the cultivar-specific calibration across locations and cultivars, as indicated by lower normalised RMSE (Table 5). CATIMO predicted CP concentration and dNDF less accurately in the generic calibration than in the cultivar-specific calibration, whereas it predicted NDF concentration better in the generic calibration (Table 5). STICS predicted NDF concentration better in the cultivar-specific calibration than in the generic calibration, whereas the predictions of dNDF and CP concentration differed only slightly between the cultivar-specific and generic calibrations (Table 5).

The trends in prediction of CP concentration across N fertiliser levels for the three models were similar to those in the cultivar-specific calibration and validation, but with a tendency for larger over-estimations by CATIMO under conditions with high N-fertiliser levels (Fig. 6). Similarly to the cultivar-specific calibration, simulated NDF concentration and dNDF did not show any trends across N-fertiliser levels for any of the three models. Moreover, there were no discernible trends in predictability of the three nutritive attributes across differences in mean annual air temperature and accumulated precipitation (not shown).

Table 5.

Table 6.

Figure 7.

4. Discussion

4. 1. Differences in prediction accuracy among nutritive attributes

This study examined how models with different structures and calibration procedures affect the prediction of dNDF and concentrations of CP and NDF, in timothy under a broad range of environmental conditions in the northern hemisphere. The predictions of NDF concentration and dNDF were generally better than those of CP concentration, as indicated by lower normalised RMSE and relative MBE and higher d-index in the calibrations of the three models. This indicates that fibre concentration and digestibility can be predicted with higher accuracy than N or CP concentration. These patterns in prediction accuracy of nutritive attributes that were previously reported from evaluations of CATIMO (Bonesmo et al., 2005; Jing et al., 2013) and STICS (Jégo et al., 2013) against data from field experiments in Canada are confirmed and extended to BASGRA. Even though the timothy CP concentration was less accurately simulated than the NDF concentration and dNDF in the studies cited above, it was generally predicted with better accuracy than in our study. In both CATIMO and STICS, crop N demand is based on critical and maximum N dilution curves. The parameters of those curves established for Canadian cultivars, which were not calibrated in our study, might not be adequate for European cultivars. Our results indicate that existing forage grass models are more efficient at predicting NDF concentration and dNDF than CP concentration.

4. 2. Differences in predictability between cultivar-specific and generic and calibration

The variability in prediction accuracy between cultivar-specific and generic calibrations provides information on the required calibration of forage grass models used to predict nutritive value. The fact that CATIMO and STICS tended to have better prediction accuracy with the cultivar-specific calibration than with the generic calibration suggests that separate calibrations for different cultivars could improve their predictive capacity. The overall slightly better prediction accuracy of BASGRA in the generic calibration than in the cultivar-specific calibration is, however, surprising. One reason could be that the larger dataset in the generic calibration than in the cultivar-specific calibration limited the influence of outliers and resulted in more accurate predictions. Van Oijen et al. (2013) found that a generic calibration of models for Scots pine trees did not result in less accurate growth predictions than calibrations using country-specific data. It should be noted, however, that the cultivar-specific datasets in our study were obtained from experiments under different environmental conditions. Hence, differences in prediction accuracy between the cultivar-specific and generic calibrations could be at least partly the result of non-cultivar differences between experimental sites, including differences in climate, soils and crop management. However, this was not confirmed by the analyses of observed and simulated nutritive attributes against N fertiliser levels, mean annual air temperature and accumulated annual precipitation, which revealed little information about the impact of environmental variability on model prediction ability. Nevertheless, a previous study in which LINGRA, the predecessor of BASGRA, simulated only one timothy cultivar (cv. Grindstad) in a number of field experiments in northern Europe showed better prediction of aboveground DM biomass when the model was calibrated specifically for one experimental site than when it was calibrated using data from several sites (Persson et al., 2014). To single out the effects of cultivars on calibration accuracy without any possible confounding effects from weather, soil or other environmental factors, comparisons of cultivar-specific and generic calibrations could be performed against data from one single site should there be any such datasets available. Moreover, further knowledge on cultivar-specific traits that are important to the prediction accuracy for nutritive attributes could possibly be obtained by grouping cultivars with similar traits together in the same calibration.

4. 3. Comparisons with dry matter yield predictability

The prediction accuracy of nutritive attributes was generally within the same range or better than the prediction accuracy of the forage DM yield for the same three models and partly the same experimental data (Korhonen et al., 2018). The normalised RMSE for the forage DM yield predictions reported from the study by Korhonen et al. (2018), which varied between 24 and 93 % across calibrations and validations, was generally greater than that for the nutritive attributes in both generic calibrations and validations. Sixty-eight out of the 101 treatments that were used in our study were not included in the study of Korhonen et al. (2018). The calibration techniques applied for CATIMO and STICS meant that dry matter related parameters calibrated in the study of Korhonen et al. (2018) for the other cultivars and the generic calibration did not change. The new Grindstad treatments added here can hence be regarded as an additional validation of the Grindstad and generic calibrations. The normalized RMSE for the calibration treatments of the Nuutti (CATIMO 62 %, STICS 27 %) and the newly added Grindstad treatments (CATIMO 66 %, STICS 25 %) from Maaninka and Ruukki 2015 and 2016, and the normalised RMSE of the generic calibration as evaluated against the same treatments (CATIMO 129 %, STICS 26 %) were mostly within the same range as the normalised

RMSE of the DM yield predictions in Korhonen et al. (2018). For comparing models, we calculated the RMSE of the DM predictions for the same treatment also for the BASGRA calibrations in which, unlike the CATIMO and STICS calibrations, the values of all parameters changed during the cultivar-specific calibration of this model. The normalised RMSE for the calibration treatments of the Nuutti and Grindstad cultivars from Maaninka 2015 and 2016 was 15 and 20 % respectively. For the generic calibration of BASGRA, the normalised RMSE for the same treatments was 32 %. In total, these results illustrate that regardless of the calibration technique the inclusion of nutritive value here was not at the expense of the predictability of the DM yield.

In previous evaluations of STICS (Jégo et al., 2013) and CATIMO (Bonesmo et al., 2005; Jing et al., 2013) for Canadian locations and timothy cultivars, the normalised RMSE for forage DM yield predictions was between 70 and 120 % greater than for NDF concentration, and between 220 and 440 % greater than for dNDF. Our results confirm that nutritive value predictions can be as accurate as DM yield predictions in forage grasses.

4. 4. Possible explanations for the differences in prediction accuracy

Crude protein concentration in plants results from rather complex soil and plant N processes, which are all affected by soil conditions, weather and crop management. Besides possible errors in the input data, errors in the descriptions of processes could have affected the CP concentration predictions. That those parameters, which were related to CP concentration were calibrated in BASGRA, but not in CATIMO and STICS, may have been a reason for the difference in prediction accuracy of this attribute among the three models. The higher simulated CP concentration at low simulated dry matter yield than the observed CP concentration at similar observed dry matter yield in CATIMO and to a lesser extent in STICS (Figure 7) indicates that the assumption of N dilution with biomass that was taken from previous model development against field trial data in Canada was not applicable to the cultivars and environmental conditions in northern Europe that were investigated here, at least not at low dry matter yield. Further experimental studies are needed to demonstrate whether there are differences in N demand and uptake between timothy cultivars. However, the greater difference in CP prediction accuracy between CATIMO and STICS than between the latter and BASGRA indicates that there are other underlying reasons than the representation of plant N dilution with biomass or light extinction for the differences in CP prediction accuracy. One reason could be differences in leaf/stem ratio prediction accuracy, which would affect growth and hence N uptake and concentration. Unfortunately, there were insufficient data available to thoroughly analyse correlations between leaf/stem ratio and predictions of CP concentration. Forage NDF concentration increases and dNDF decreases with phenological development, and these variables also directly affected by temperature in CATIMO and STICS. Although severe N stress affects NDF concentration and dNDF in CATIMO and STICS, there is no effect of soil and plant N on NDF concentration and dNDF under normal N conditions in any of the three models which, given the complexity of soil-plant N processes, could explain why they are better predicted than CP concentration. Differences between models in calculation of leaf/stem ratios could also explain some of the variation in predicted NDF concentration and dNDF among the three models. However, the effect of the leaf/stem ratio on CP concentration is probably larger, due to the complex interaction between N distribution in the plant and growth.

4. 5. Uncertainty in input data

Because the methods used for analysis of the three nutritive attributes were not always the same at all sites, there is some uncertainty in the values (Huhtanen et al., 2006). Of the three nutritive attributes included in our study, dNDF most likely has the largest uncertainty associated with the analysis methods and CP concentration the smallest. Different dNDF values for the same forage sample analysed *in vitro* in different laboratories may stem from differences in the pore size of the nylon bags in which the samples are incubated and from differences in the incubation time and the rumen liquid used. Similarly, differences in methodology between laboratories, such as the use of different extraction chemicals, may result in laboratory differences in NDF concentration estimates (Tavares da Silva et al., 2018). It should also be noted that NIRS often has poorer prediction accuracy for NDF concentration and dNDF than for CP concentration, although high accuracy can also be achieved for the former attributes if the method is carefully calibrated with an adequate number of representative reference samples and suitable reference analysis methods (Huhtanen et al., 2006). Nevertheless, the better prediction accuracy of NDF concentration and dNDF than of CP concentration indicates that other reasons than the uncertainty in nutritive value measurements were more important to the prediction accuracy. Errors related to the weather input data, mainly due to the distance between weather stations and observations in the field, could also have affected our results. For most sites, there were no direct measurements of soil hydraulic properties available as input to the soil modules of the grass models and instead these variables were estimated from data on soil texture fractions. For BASGRA and STICS, the soil water contents at wilting point and at field capacity were therefore treated as parameters. However, that was not the case for CATIMO. Possible within-field variation in soil texture may also have caused differences between the actual soil properties and those that were input to the simulations.

4. 6. Suggested further studies

The low prediction accuracy of CP concentration, the importance of CP concentration for nutritive value and the general importance of N for crop performance and for its environmental impact emphasise the need for improved descriptions of soil and plant N in the three models. Moreover, studies with synchronised calibration procedures could help assess whether the differences in prediction accuracy between the models were due to differences in calibration methods or the model structure. Such information could increase the performance of models and thus their prospects of being applied in analysis of forage grass performance under various existing or hypothetical environmental conditions. Further calibrations with experimental data from other areas, such as Russia, northern Japan and mountainous regions at lower latitudes in Europe, could provide more information about the general applicability of the models. To place the performance of the three models in a broader context, validation of the performance of other grassland models in predicting forage nutritive value could also be valuable.

5. Conclusions

Three models with different structures (BASGRA, CATIMO and STICS) predicted NDF concentration and digestibility in timothy with similar accuracy to previous predictions of forage DM yield of timothy across a wide range of climate and soil conditions in Canada and northern Europe. However, prediction of CP concentration was rather poor compared with the other nutritive attributes. Cultivar-specific calibrations improved the performance of CATIMO and STICS, but not of BASGRA, compared with calibrations where data on all cultivars were used together.

Acknowledgements

This research was conducted in the context of MACSUR (Modelling European Agriculture with Climate Change for Food Security), a project launched by the Joint Research Programming Initiative (JPI) on Agriculture, Food Security and Climate Change. The authors gratefully acknowledge financial support from the following sources: Academy of Finland (project PLUMES, decision nos. 277403 and 292836), the European Agricultural Fund for Rural Development (project NuRa) and the Research Council of Norway.

References

Bélanger, G., and Gastal, F. 2000, Nitrogen utilization by forage grasses. Can. J. Plant Sci. 80:11-20.

Bélanger, G., McQueen, R.E., 1999. Leaf and stem nutritive value of timothy grown with varying N nutrition in spring and summer. Can. J. Plant Sci. 79, 223-229.

Bélanger, G., Michaud, R., Jefferson, P.G., Tremblay, G.F., Brégard, A., 2001. Improving the nutritive value of timothy through management and breeding. Can. J. Plant Sci. 81, 577-585.

Bonesmo, H., Bélanger, G., 2002a. Timothy yield and nutritive value by the CATIMO model: I. Growth and nitrogen. Agron. J. 94, 337-345.

Bonesmo, H., Bélanger, G., 2002b. Timothy yield and nutritive value by the CATIMO model: II. Digestibility and fiber. Agron. J. 94, 345-350.

Bonesmo, H., Bélanger, G., Charmley, E., Drapeau, R., McKenzie, D.B., Michaud, R., Tremblay, G.F., 2005. Timothy yield and nutritive value by the CATIMO model: III. Validation for eastern Canada. Agron. J. 97, 32-40.

Brisson, N., Launay, M., Mary, B., Beaudoin, N., 2008, Conceptual basis, formalisations and parameterization of the STICS crop model. Editions QUAE (Versailles).

Brisson, N., Mary, B., Ripoche, D., Jeuffroy, M.H., Ruget, F., Nicoullaud, B., Gate, P., Devienne-Barret, F., Antonioletti, R., Durr, C., Richard, G., Beaudoin, N., Recous, S., Tayot, X., Plenet, D., Cellier, P., Machet, J.M., Meynard, J.M., and Delécolle, R. 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.

Bruinenberg, M.H., Valk, H., Korevaar, H., Struik, P.C., 2002. Factors affecting digestibility of temperate forages from seminatural grasslands: a review. Grass Forage Sci. 57, 292-301.

Buxton, D.R., 1996. Quality-related characteristics of forages as influenced by plant environment and agronomic factors. Anim. Feed Sci. Tech. 59, 37-49.

Casler, M.D., Kallenbach, R.L., 2007. Cool-season grasses for humid areas. Forages. The Science of Grassland Agriculture 2, 211-220.

Chang, J., Viovy, N., Vuichard, N., Ciais, P., Wang, T., Cozic, A., Lardy, R., Graux, A.-I., Klumpp, K., Martin, R., 2013. Incorporating grassland management in ORCHIDEE: model description and evaluation at 11 eddy-covariance sites in Europe. Geosci. Model Dev. 6, 2165-2181.

Chapman, D.F., Lee, J.M., Waghorn, G.C., 2014. Interaction between plant physiology and pasture feeding value: a review. Crop Pasture Sci. 65, 721-734.

Ehrhardt, F., Soussana, J.F., Bellocchi, G., Grace, P., McAuliffe, R., Recous, S., Sándor, R., Smith, P., Snow, V., de Antoni Migliorati, M., 2018. Assessing uncertainties in crop and pasture ensemble model simulations of productivity and N2O emissions. Global Change Biol. 24, e603-e616.

Elgersma, A., Søegaard, K., 2018. Changes in nutritive value and herbage yield during extended growth intervals in grass–legume mixtures: effects of species, maturity at harvest, and relationships between productivity and components of feed quality. Grass Forage Sci. 73, 78-93.

Fulkerson, W.J., Neal, J.S., Clark, C.F., Horadagoda, A., Nandra, K.S., Barchia, I., 2007. Nutritive value of forage species grown in the warm temperate climate of Australia for dairy cows: Grasses and legumes. Livest. Sci. 107, 253-264.

Gustavsson, A.-M., Martinsson, K., 2004. Seasonal variation in biochemical composition of cell walls, digestibility, morphology, growth and phenology in timothy. Eur. J. Agron. 20, 293-312.

Huhtanen, P., Nousiainen, J., Rinne, M., 2006. Recent developments in forage evaluation with special reference to practical applications. Agr. Food Sci. 15, 293-323.

Hurtado-Uria, C., Hennessy, D., Shalloo, L., Schulte, R.P.O., Delaby, L., O'Connor, D., 2013. Evaluation of three grass growth models to predict grass growth in Ireland. J. Agr. Sci-Cambridge 151, 91-104.

Höglind, M., Schapendonk, A., Van Oijen, M., 2001. Timothy growth in Scandinavia: combining quantitative information and simulation modelling. New Phytol. 151, 355-367.

Höglind, M., Hanslin, H.M., 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., Van Oijen, M., Cameron, D., Persson, T., 2016. Process-based simulation of growth and overwintering of grassland using the BASGRA model. Ecol. Model. 335, 1-15.

Jégo, G., Bélanger, G., Tremblay, G.F., Jing, Q., Baron, V.S., 2013. Calibration and performance evaluation of the STICS crop model for simulating timothy growth and nutritive value. Field Crops Res. 151, 65-77.

Jensen, K.B., Robins, J.G., Rigby, C., Waldron, B.L., 2016. Comparative trends in forage nutritional quality across the growing season in 13 grasses. Can. J. Plant Sci. 97, 72-82.

Jing, Q., Bélanger, G., Baron, V., Bonesmo, H., Virkajärvi, P., Young, D., 2012. Regrowth simulation of the perennial grass timothy. Ecol. Model. 232, 64-77.

Jing, Q., Bélanger, G., Baron, V., Bonesmo, H., Virkajärvi, P. 2013. Simulating the nutritive value of timothy summer regrowth. Agron. J. 105, 563-572.

Jokela, V., Trevaskis, B., Seppänen, M.M., 2015. Genetic variation in the flowering and yield formation of timothy (Phleum pratense L.) accessions after different photoperiod and vernalization treatments. Front. Plant Sci. 6.

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

Kipling, R.P., Virkajärvi, P., Breitsameter, L., Curnel, Y., De Swaef, T., Gustavsson, A.-M., Hennart, S., Höglind, M., Järvenranta, K., Minet, J., Nendel, C., Persson, T., Picon-Cochard, C., Rolinski, S., Sandars, D.L., Scollan, N.D., Sebek, L., Seddaiu, G., Topp, C.F.E., Twardy, S., Van Middelkoop, J., Wu, L., Bellocchi, G., 2016. Key challenges and priorities for modelling European grasslands under climate change. Sci. Total Environ. 566–567, 851-864.

Köchy, M., 2008. Effects of simulated daily precipitation patterns on annual plant populations depend on life stage and climatic region. BMC Ecol. 8, 4.

Köppen, W., 1936: Das Geographische System der Klimate. In: W. Köppen and R. Geiger eds. Handbuch der Klimatologie, vol I, Part C. Gebrüder Borntraeger, Berlin. Germany.

Korhonen, P., Palosuo, T. Persson, T., Höglind, M., Jégo, G., Van Oijen, M., Gustavsson, A-M., Bélanger, G., Perttu Virkajärvi, P., 2018. Modelling grass yields in northern climates - a comparison of three growth models for timothy. Field Crops Res. 224,37-47.

Marum, P., 1990. Bestemmelse av kvalitet i forvekster ved hjelp av NIRS og metodens muligheter i engvekstforedling. Norsk Landbruksforsking (Norway). Norwegian Agricultural Research. no. 9/90 (suppl.).(NLH accession no. NOR).

Nissinen, O., Kalliainen, P., Jauhiainen, L., 2010. Development of yield and nutritive value of timothy in primary growth and regrowth in northern growing conditions. Agr. Food Sci. 19, 252-268.

Nordheim-Viken, H., Volden, H., Jorgensen, M., 2009. Effects of maturity stage, temperature and photoperiod on growth and nutritive value of timothy (Phleum pratense L.). Anim. Feed Sci. Tech. 152, 204-218.

Persson, T., Höglind, M., Gustavsson, A.-M., Halling, M., Jauhiainen, L., Niemeläinen, O., Thorvaldsson, G., Virkajärvi, P., 2014. Evaluation of the LINGRA timothy model under Nordic conditions. Field Crops Res. 161, 87-97.

Pontes, L.S., Carrère, P., Andueza, D., Louault, F., Soussana, J.F., 2007. Seasonal productivity and nutritive value of temperate grasses found in semi-natural pastures in Europe: responses to cutting frequency and N supply. Grass Forage Sci. 62, 485-496.

Sándor, R., Barcza, Z., Acutis, M., Doro, L., Hidy, D., Köchy, M., Minet, J., Lellei-Kovács, E., Ma, S., Perego, A., 2017. Multi-model simulation of soil temperature, soil water content and biomass in Euro-Mediterranean grasslands: Uncertainties and ensemble performance. Eur. J. Agron. 88, 22-40.

Sándor, R., Barcza, Z., Hidy, D., Lellei-Kovács, E., Ma, S., Bellocchi, G., 2016. Modelling of grassland fluxes in Europe: evaluation of two biogeochemical models. Agr. Ecosyst. Environ. 215, 1-19.

Schapendonk, A., Stol, W., van Kraalingen, D.W.G., Bouman, B.A.M., 1998. LINGRA, a sink/source model to simulate grassland productivity in Europe. Eur. J. Agron. 9, 87-100.

Sivia, D.S., 2006. Data Analysis: A Bayesian Tutorial. Oxford University Press, Oxford, UK.

Tavares da Silva, R.S., Fernandes A.M., dos Santos Gomez, R., Ramos Bendia, L.C., da Costa e Silva, L., Mendonca Viera, R.A. 2018. On the specificity of different methods for neutral fiber and related problems. Anim. Feed Sci. Tech. 240, 128-144.

Thornton, P.K., 2010. Livestock production: recent trends, future prospects. Philos. T. Roy. Soc. B. 365, 2853-2867.

Van Oijen, M., Reyer, C., Bohn, F., Cameron, D., Deckmyn, G., Flechsig, M., Härkönen, S., Hartig, F., Huth, A., Kiviste, A., 2013. Bayesian calibration, comparison and averaging of six forest models, using data from Scots pine stands across Europe. Forest Ecol. Manag. 289, 255-268.

Van Oijen, M., Rougier, J., Smith, R., 2005. Bayesian calibration of process-based forest models: bridging the gap between models and data. Tree Physiol. 25, 915-927.

Van Soest, P.J., Robertson, J.B., Lewis, B.A., 1991. Methods for dietary fiber, neutraldetergent fiber, and non starch polysaccharides in relation to animal nutrition. J. Dairy Sci. 74, 3583–3597.

Van Soest, P.J., Wine, R.H., Moore, L.A., 1966. Estimation of the true digestibility of forages by the in vitro digestion of cell walls. Proc. Int. Grassland Congress, 10th, Helsinki, Finland, 7–16 July 1966. Finnish Grassl. Assoc., Helsinki (1966), pp. 438-441.

Vital, J.-A., Gaurut, M., Lardy, R., Viovy, N., Soussana, J.-F., Bellocchi, G., Martin, R., 2013. High-performance computing for climate change impact studies with the Pasture Simulation model. Comput.Electron. Agr. 98, 131-135.

Wilkinson, J.M., Rinne, M., 2018. Highlights of progress in silage conservation and future perspectives. Grass Forage Sci. 73, 40-52.

Willmott, C.J., 1981. On the validation of models. Phys. Geogr. 2, 184-194.

Wu, L., McGechan, M., McRoberts, N., Baddeley, J., Watson, C., 2007. SPACSYS: integration of a 3D root architecture component to carbon, nitrogen and water cycling—model description. Ecol. Model. 200, 343-359.

Table 1. Cultivars, nutritive value (crude protein (CP) concentration, neutral detergent fibre (NDF) concentration, digestibility of NDF (dNDF)) and N fertilizer treatments at the experimental sites

Location	Mean annual temp (°C) /annual acc. precipitation (mm)	Köppen climate classification[1]	Soil type	Dataset year	Cultivar	Nutritive values	N fert. regime (kg ha^{-1} yr^{-1})
Fredericton (45°55′N; 66°32′W; 35 m asl)	5.7/1108	Dfb (Warm-summer humid continental climate)	Loam	1991-1993	Champ	NDF, dNDF	0, 70, 140, 168, 200, 210
Lacombe (52°28′N; 113°44′W; 860 m asl)	3.5/429	Dfb (Warm-summer continental climate)	Silty clay loam	2004-2005	Climax	NDF	100
Québec (46°47′N; 71°07′W; 75 m asl)	5.3/1009	Dfb (Warm-summer humid continental climate)	Loam	1999-2001	Champ	NDF, dNDF	0, 60, 120

Maaninka (63°09′N; 27°17′E; 90 m asl)	4.2/560	Dfc (Continental Subarctic Climate).	Silt loam / loam[2]	2006-2007, 2015-2016	Tammisto II, Nuutti, Grindstad	NDF, CP	0, 150, 180, 190, 200, 250, 300, 350, 400, 450
Rovaniemi (66°35′N; 26°01′E; 106 m asl)	1.0/610	Dfc (Continental Subarctic Climate).	Silt loam	1999-2001	Iki	-	160, 200
Ruukki (64°40′N; 25°06′ E; 48 m asl)	2.6/513	Dfc (Continental Subarctic Climate).	Silt loam	2015-2016	Nuutti, Grindstad	NDF, CP	0, 150, 200, 250, 300, 350, 400, 450
Særheim	8.0/1392	Cfc (Cool oceanic climate)	Sandy loam	2000-2002	Grindstad	NDF, CP	220

(58°46′N; 5°39′E; 90 m asl)							
Umeå (63°45′N; 20°17′E; 12 m asl)	3.3/595	Dfc (Continental subarctic climate).	Silt loam	1995-1996	Jonatan	NDF, CP	180

[1] Köppen 1936

[2] Treatments 1-3, 33-37: Silt loam soil. Treatments 38-68: Loam soil

Table 2. Division between calibration and validation data within the dataset

Treatment number	Location	Cultivar	N fertiliser application (kg ha^{-1} yr^{-1})	Calibration/validation
1-2	Maaninka	Tammisto II	180	Calibration
3-8	Rovaniemi	Iki	160, 200	Calibration
9, 11-12, 14	Særheim	Grindstad	220	Calibration
10, 13	Særheim	Grindstad	220	Validation
15-21,23	Québec	Champ	0,60,120	Calibration
22	Québec	Champ	60	Validation
24-25	Lacombe	Climax	100	Calibration
26-27	Umeå	Jonatan	180	Calibration
28, 30-32	Fredericton	Champ	0, 70, 140, 168, 210	Calibration
29, 33	Fredericton	Champ	200, 210	Validation
34-35, 37, 46-50, 52-53, 62-68	Maaninka	Nuutti	0, 150, 190, 200, 250, 300, 350, 400, 450	Calibration
36, 51, 69	Maaninka	Nuutti	190, 350, 450	Validation
38-40, 42-45, 54-59, 61	Maaninka	Grindstad	0, 150, 200, 250, 300,	Calibration

			350, 400, 450	
41,60	Maaninka	Grindstad	250, 400	Validation
71-77, 86-91, 93	Ruukki	Grindstad	0, 150, 200, 250, 300, 350, 400, 450	Calibration
70, 92	Ruukki	Grindstad	0, 400	Validation
78-82, 84-85, 94-97, 99-101	Ruukki	Nuutti	0, 150, 200, 250, 300, 350, 400, 450	Calibration
83, 98	Ruukki	Nuutti	300, 350	Validation

Table 3. Statistics on the cultivar-specific calibration: Observed and simulated means, root mean squared error (RMSE), normalised RMSE, relative mean bias error (rMBE) and Willmott's index of agreement (d-index) for crude protein (CP) concentration, neutral detergent fibre (NDF) concentration and digestibility of NDF (dNDF)

	Number of observations	Mean of observation	Mean of simulation	RMSE	Normalised RMSE (%)	rMBE (%)	d-index
			CP concentration (g g^{-1} DM)				
BASGRA	173	0.14	0.13	0.027	19	-6.0	0.82
CATIMO			0.17	0.070	50	19.0	0.57
STICS			0.18	0.055	40	29.0	0.88
			NDF concentration (g g^{-1} DM)				
BASGRA	252	0.56	0.51	0.072	13	-9.0	0.59
CATIMO			0.57	0.120	21	-0.2	0.43
STICS			0.56	0.045	8	-0.4	0.75
			dNDF (g g^{-1} NDF)				
BASGRA	28	0.78	0.71	0.170	22	0.7	0.72
CATIMO			0.82	0.077	10	5.0	0.64
STICS			0.78	0.046	6	-3.0	0.82

Table 4. Statistics of the cultivar-specific validation: Observed and simulated means, root mean squared error (RMSE), normalised RMSE, relative mean bias error (rMBE) and Willmott's index of agreement (d-index) for crude protein (CP) concentration, neutral detergent fibre (NDF) concentration and digestibility of NDF (dNDF)

	Number of observations	Mean of observation	Mean of simulation	RMSE	Normalized RMSE (%)	rMBE (%)	d-index
CP concentration (g g^{-1} DM)							
BASGRA	48	0.14	0.13	0.037	26	-5.0	0. 72
CATIMO			0.17	0.063	45	18.0	0.67
STICS			0.16	0.034	24	11.0	0.93
NDF concentration (g g^{-1} DM)							
BASGRA	62	0.55	0.52	0.063	11	-0.1	0.63
CATIMO			0.55	0.077	14	-0.7	0.64
STICS			0.55	0.047	9	-0 5	0.71
dNDF (g g^{-1} NDF)							
BASGRA	14	0.78	0.70	0.081	10	-0.1	0.70
CATIMO			0.81	0.053	7	0.5	0.81
STICS			0.75	0.041	5	-3.0	0.87

Table 5. Statistics of the generic calibration: Observed and simulated means, root mean squared error (RMSE), normalised RMSE, relative mean bias error (rMBE) and Willmott's index of agreement (d-index) for crude protein (CP) concentration, neutral detergent fibre (NDF) concentration and digestibility of NDF (dNDF)

	Number of observations	Mean of observation	Mean of simulation	RMSE	Normalised RMSE (%)	rMBE (%)	d-index
CP concentration (g g^{-1} DM)							
BASGRA	173	0.14	0.15	0.022	16	-0.8	0.89
CATIMO			0.26	0.14	101	87	0.36
STICS			0.17	0.052	38	25	0.92
NDF concentration (g g^{-1} DM)							
BASGRA	252	0.56	0.56	0.050	8.8	-0.4	0.72
CATIMO			0.55	0.095	17	-3	0.49
STICS			0.56	0.066	12	0.8	0.46
dNDF (g g^{-1} NDF)							
BASGRA	28	0.78	0.75	0.072	9.3	-4	0.59
CATIMO			0.51	0.29	38	-34	0.34
STICS			0.79	0.050	6.4	0.23	0.82

Table 6. Statistics of the generic validation: Observed and simulated means, root mean squared error (RMSE), normalised RMSE, relative mean bias error (rMBE) and Willmott's index of agreement (d-index) for crude protein (CP) concentration, neutral detergent fibre (NDF) concentration and digestibility of NDF (dNDF)

	Number of observations	Mean of observation	Mean of simulation	RMSE	Normalized RMSE (%)	rMBE (%)	d-index
			CP concentration (g g^{-1} DM)				
BASGRA	48	0.14	0.15	0.025	18.0	-0.2	0.91
CATIMO			0.26	0.130	92.0	86.0	0.47
STICS			0.16	0.032	23.0	11.0	0.95
			NDF concentration (g g^{-1} DM)				
BASGRA	62	0.56	0.56	0.043	7.8	2.0	0.77
CATIMO			0.55	0.095	17.1	-2.0	0.51
STICS			0.56	0.069	12.4	0.9	0.49
			dNDF (g g^{-1} NDF)				
BASGRA	14	0.78	0.75	0.050	6.5	-4.0	0.99
CATIMO			0.51	0.290	37.0	-36.0	0.34
STICS			0.79	0.047	6.0	0.2	0.99

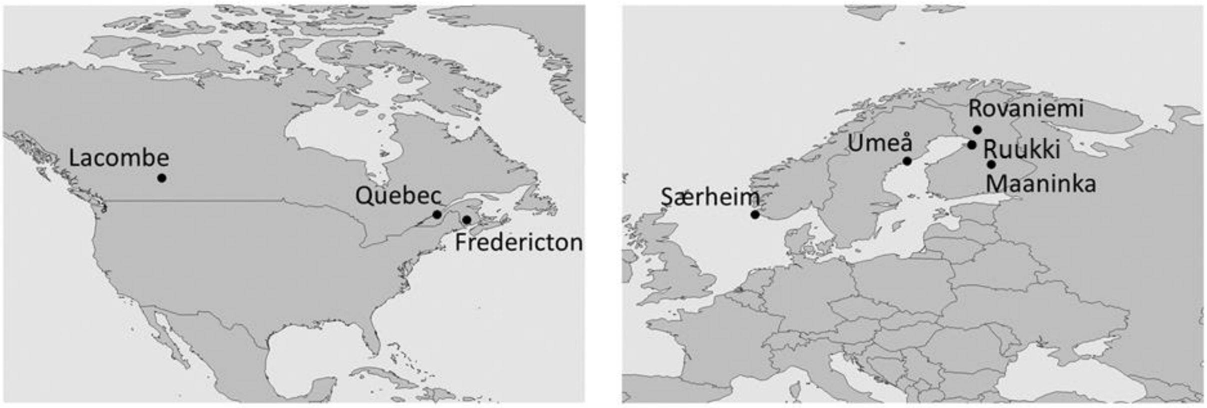

Figure 1. Geographical location of the eight experimental sites in Canada and Northern Europe.

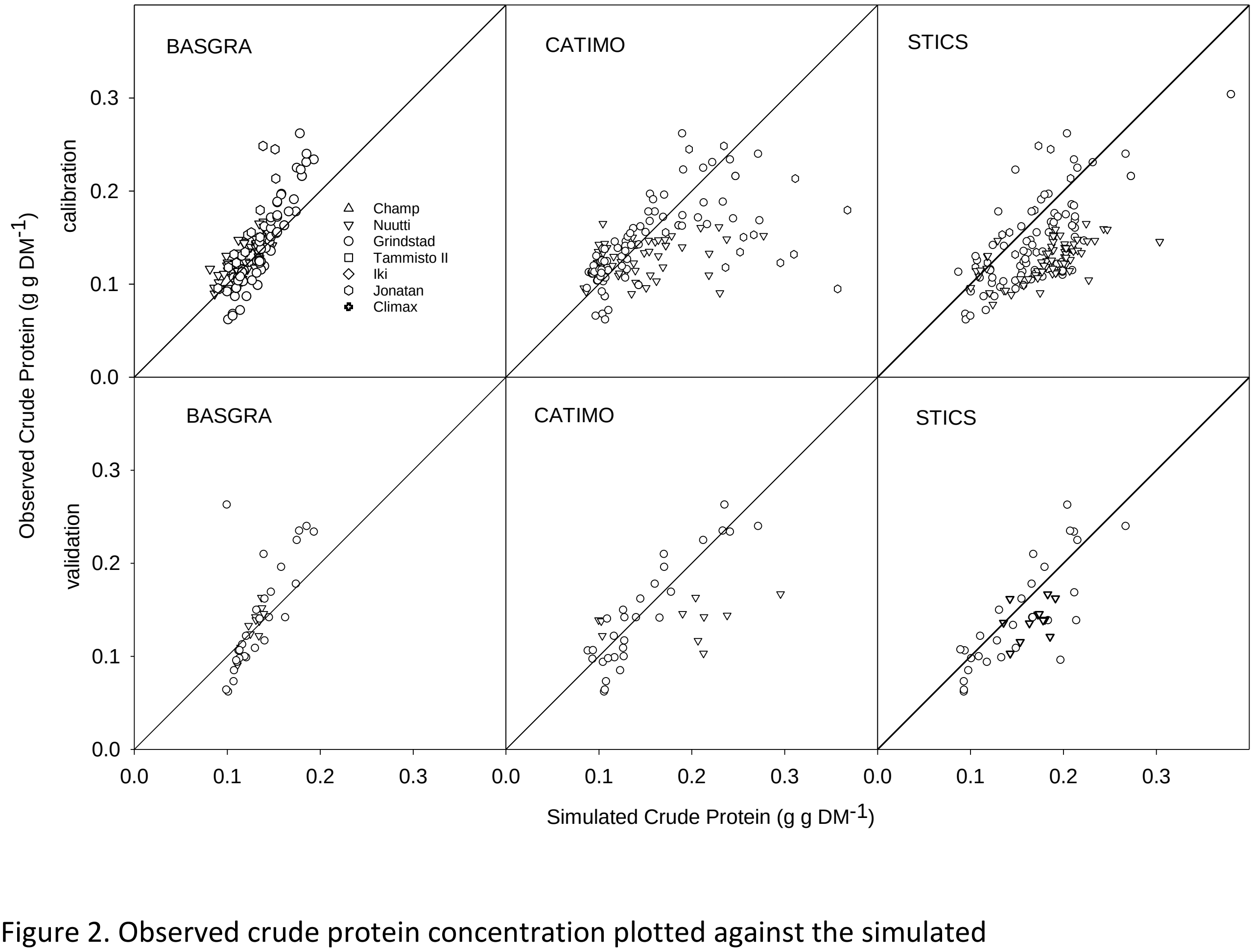

Figure 2. Observed crude protein concentration plotted against the simulated concentrations produced by BASGRA, CATIMO and STICS in the cultivar-specific calibration (upper row) and validation (lower row). Statistics on the cultivar-specific calibration and validation are presented in Tables 3 and 4.

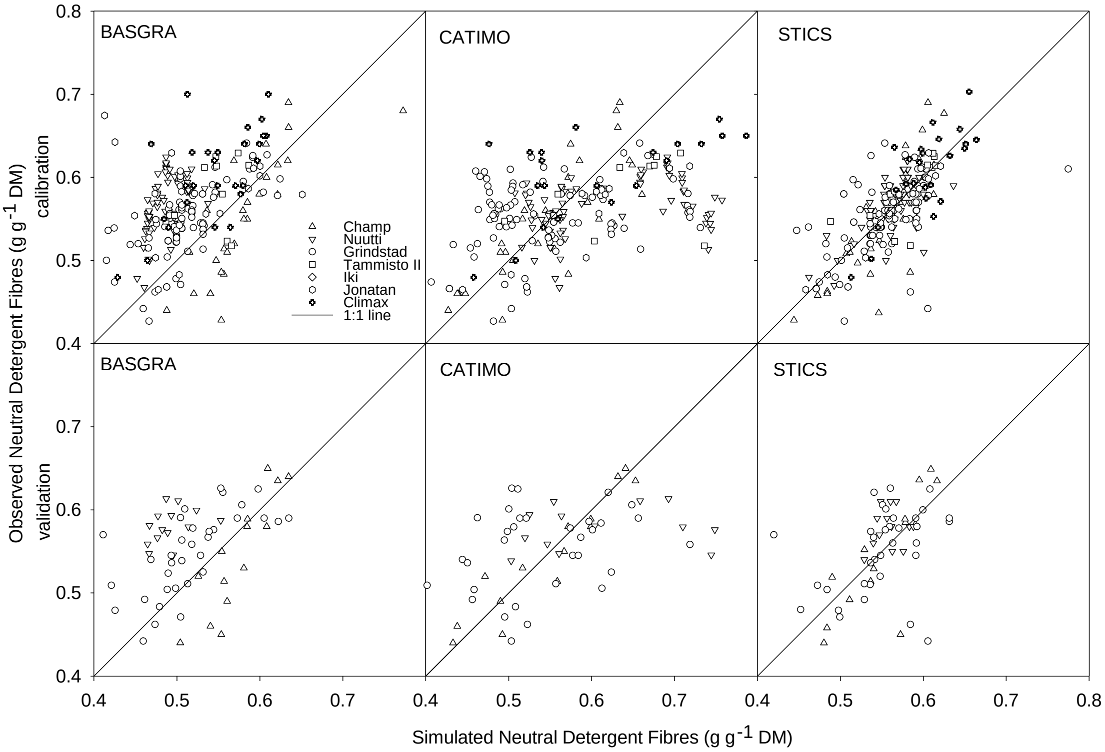

Figure 3. Observed neutral detergent fibre concentration plotted against the simulated concentrations produced by BASGRA, CATIMO and STICS in the cultivar-specific calibration (upper row) and validation (lower row). Statistics on the cultivar-specific calibration and validation are presented in Tables 3 and 4.

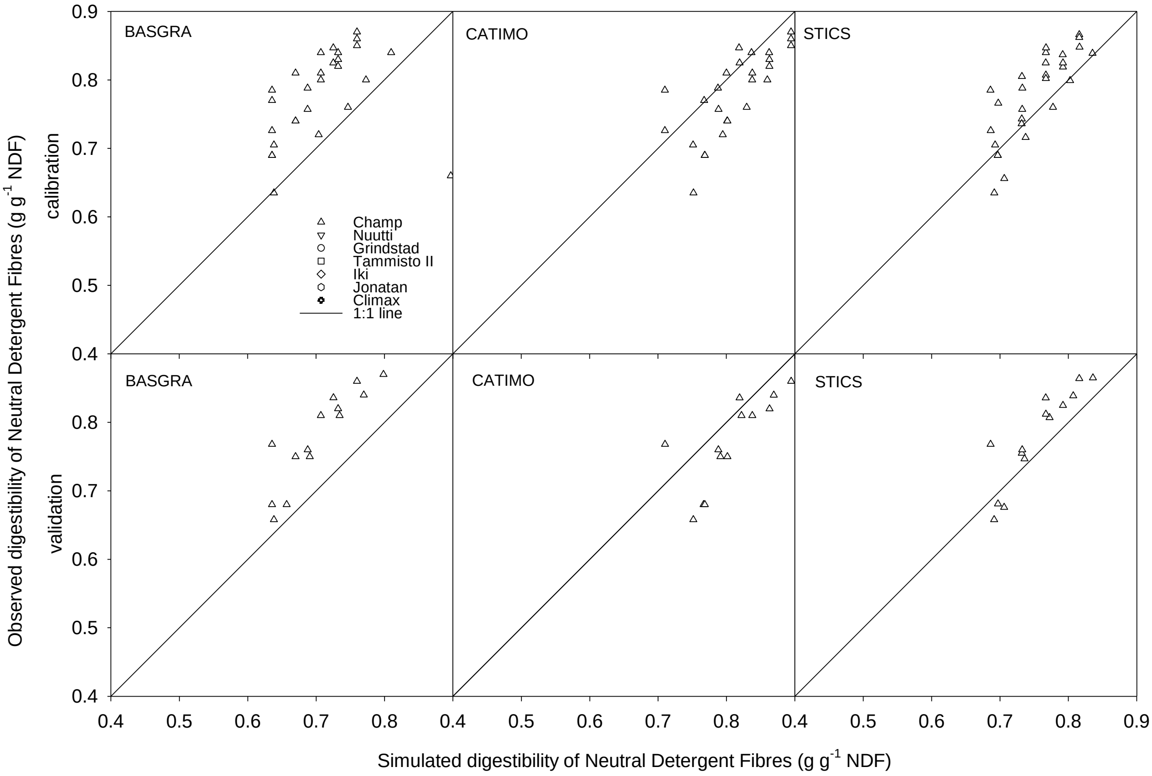

Figure 4. Observed digestibility of neutral detergent fibre (dNDF) plotted against the simulated values produced by BASGRA, CATIMO and STICS in the cultivar-specific calibration (upper row) and validation (lower row). Statistics on the cultivar-specific calibration and validation are presented in Tables 3 and 4.

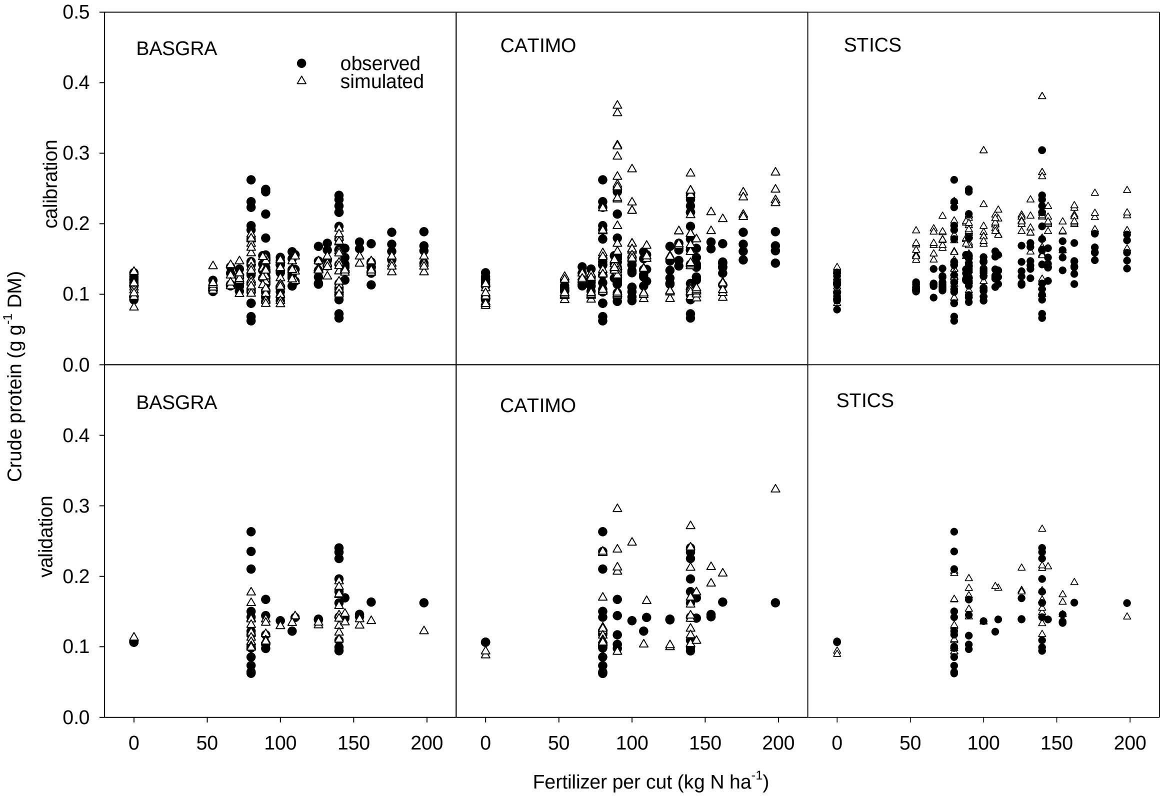

Figure 5. Observed crude protein concentration and concentration simulated by BASGRA, CATIMO, and STICS with cultivar-specific parameters as a function of amount of N applied per cut. Upper row: calibration dataset, lower row: validation dataset.

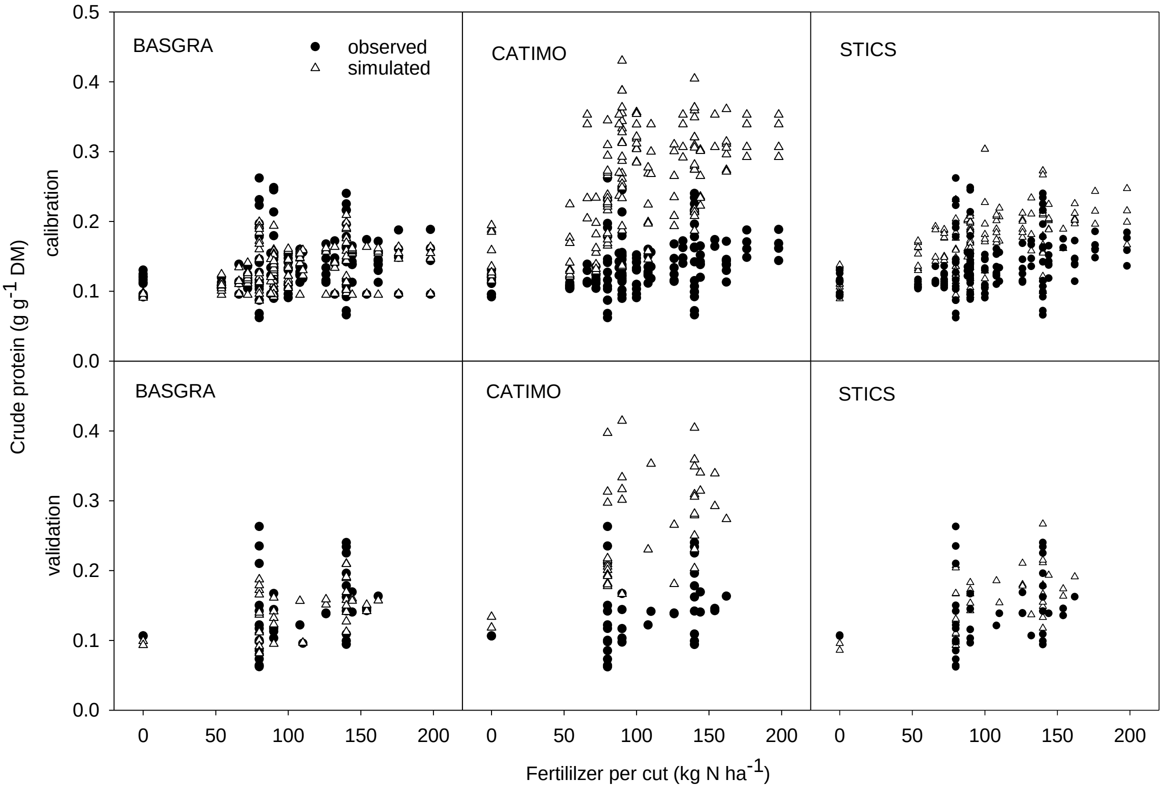

Figure 6. Observed crude protein concentration and concentration simulated by BASGRA, CATIMO, and STICS with generic parameters as a function of the amount of N applied per cut. Upper row: calibration dataset, lower row: validation dataset.

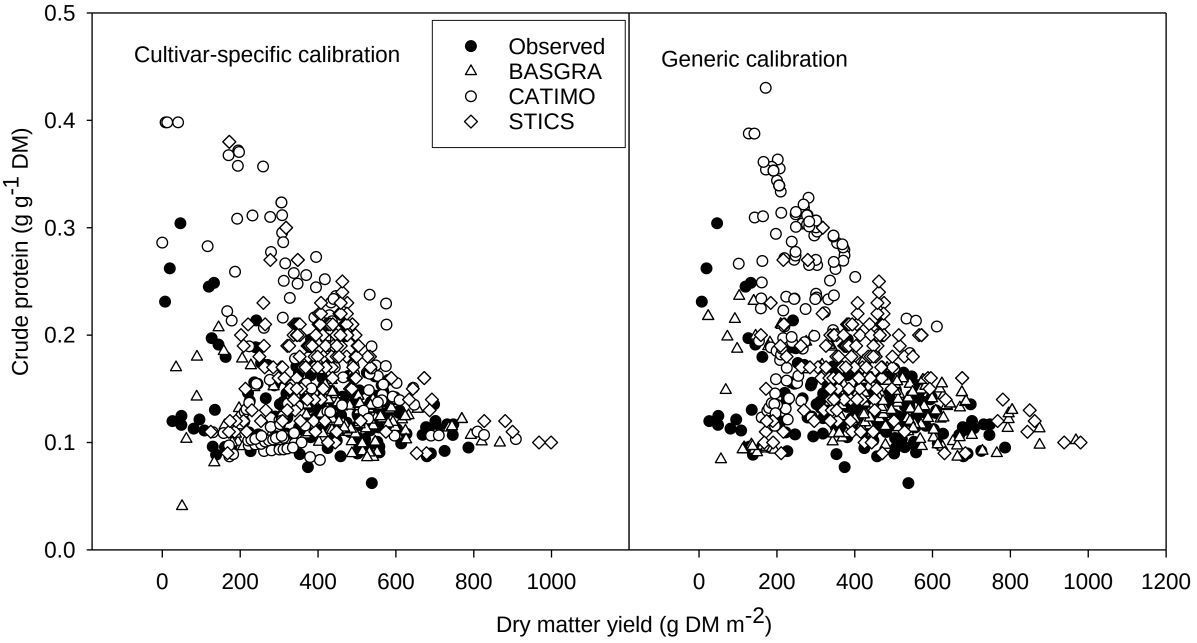

Figure 7. Observed crude protein concentration vs observed dry matter yield, and simulated crude protein vs simulated dry matter yield for the cultivar specific (left) and generic (right) calibrations.