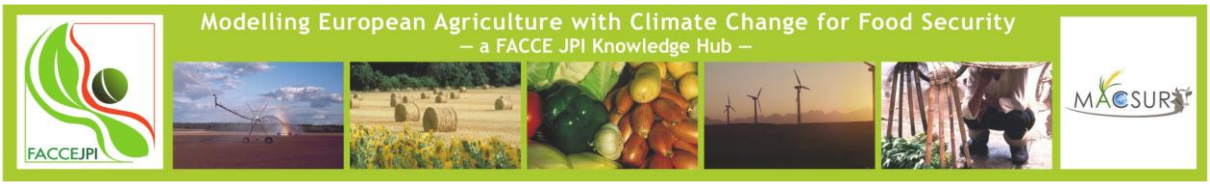

FACCE-MACSUR

D-L2.4 Model inter-comparison

Gianni Bellocchi*, Renáta Sándor and the partners of MACSUR LiveM WP2

French National Institute for Agricultural Research, 5 chemin de Beaulieu, 63039 Clermont-Ferrand, France

*gianni.bellocchi@clermont.inra.fr

Instrument:	Joint Programming Initiative
Topic:	Agriculture, Food Security, and Climate Change
Project:	Modelling European Agriculture with Climate Change for Food Security (FACCE-MACSUR)
Start date of project:	1 June 2012
Duration:	36 months
Theme, Work Package:	LiveM 2
Deliverable reference num.:	D-L2.4
Deliverable lead partner:	INRA
Due date of deliverable:	month 36
Submission date:	2015-06-09
Confidential until:	

Revision	Changes	Date
1.0	First Release	2015-06-09
1.01	Section on model uncertainty updated	2015-06-09

Abstract/Executive summary

This deliverable focuses on some illustrative results obtained with different grassland-specific, grassland adapted crop and dynamic vegetation models selected out of the first list of models compiled in D-L2.1.1 to simulate biomass and flux data from grassland sites in Europe and peri-Mediterranean regions (D-L2.1.1 and D-L2.1.2). Results from uncalibrated simulations were documented in the D-L2.3 report as a blind exercise. Some model improvements are emphasized in this report due to the higher information level of the model calibrations. The complete set of results will include simulations from uncalibrated and calibrated models.

Table of Contents

Introduction

Grassland ecosystem models have become important tools for extrapolating local observations and understanding to much larger regions and testing hypotheses on grassland ecosystem functioning (Chang et al., 2013; Graux et al., 2013; Vital et al., 2013; Ma et al., 2015). Applications range from field to global scales, targeting the evaluation of management and policy options. With increasing spatial extent of the area under investigation, input data for mechanistic grassland models are scarce and uncertain and data to test relevant state variables are insufficiently available. There is also an increasing demand by both model users and decision makers for analysis of the robustness of models and the uncertainties of model results in climate change impact studies. However, grassland models have seldom been inter-compared to assess the uncertainties in estimates that can also be large. Supporting climate change impact studies includes an up-to-date geographical coverage of climate, soil and vegetation data, and access to secondary information (e.g. soil information obtained via transfer functions from the primary data), as associated with changes in management options. A protocol based on the principles laid down by the Agricultural Model Intercomparison and Improvement Project (AgMIP, http://www.agmip.org) was developed in LiveM WP2 (M L2.2: Protocol for model evaluation), which includes: evaluation of uncalibrated (blind) and calibrated model simulations against observations, and sensitivity tests of models to changes of CO_2, temperature and precipitation. This deliverable illustrates the results obtained with seven grassland models, which are described in an inventory of modelling approaches made available through the MACSUR consortium (D L2.1.1) and for which the calibration process was completed (other modelling teams did not undertake this step of the analysis). They were applied in nine sites across Europe and peri-Mediterranean regions (D L2.1.1 and D L2.1.2).

Grassland datasets

Long-term (five to 31 years of data) grassland sites were identified (D L2.1.1, D L2.1.2), covering a gradient of geographic and climatic conditions (Figure 1, left) and a variety of management practices. Four of them (Laqueuille, France, Klumpp et al., 2011; Monte Bondone, Italy, Wohlfahrt et al., 2008; Grillenburg, Germany, Prescher et al., 2010; Oensingen, Switzerland, Amman et al., 2007), equipped with eddy covariance systems to determine the net ecosystem exchange of CO_2, are semi-natural grasslands in place for a long time including vegetation types representative of the zone (with the exception of Oensingen, established in 2001). Other sites (Kempten, Germany, Schröpel and Diepolder, 2003; Lelystad, The Netherlands, Schils and Snijders, 2004; Matta, Israel, Golodets et al., 2013; Rothamsted, United Kingdom, Silverston et al., 2006; Sassari, Italy, Cavallero et al., 1992) from experimental research focus on biomass production. The limits of the De Martonne-Gottmann index (*b*, De Martonne, 1942) discriminate between aridity conditions (Figure 1, right): $b<5$: extreme aridity; $5 \le b \le 14$: aridity; $15 \le b \le 19$: semi-aridity; $20 \le b \le 29$: sub-humidity; $30 \le b \le 59$: humidity; $b>59$: strong humidity (Diodato and Ceccarelli, 2004).

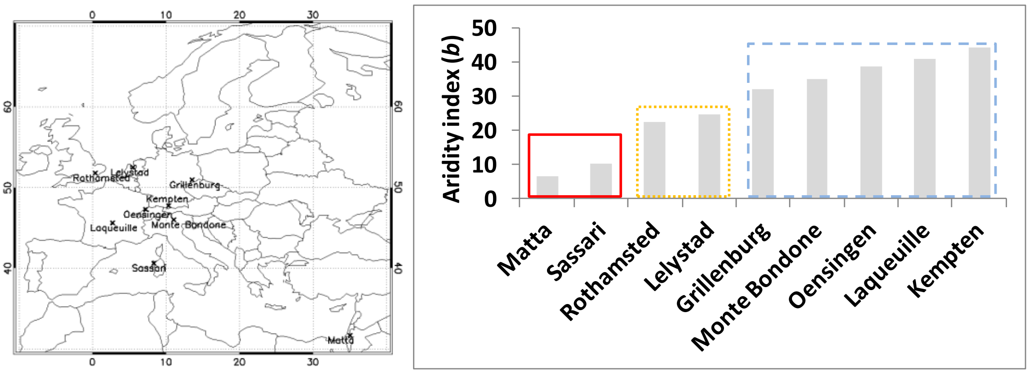

Figure 1. Geographic location (left) and classification (right) of grassland sites with respect to De Martonne-Gottmann aridity index (*b*). The solid box, dotted box and hatched box represent arid, sub-humid and humid sites, respectively.

Grassland models

Nine models were selected for the inter-comparison (D L2.1.1). Three of them are **grassland-specific models**. **AnnuGrow** (Köchy, 2008) quantifies the effect of daily rainfall distributions and compares it to the effect of a change in mean annual amount on vegetation. **PaSim** (Ma et al., 2015) simulates water, carbon and nitrogen cycles in grassland plots at sub-daily time step via modules of climate, soil biology and physics, vegetation and management (including grazing animals). **SPACSYS** (Wu et al., 2007) is a multi-dimensional, field-scale, daily time-step model of carbon and nitrogen cycles between plants, soils and microbes, with fine representation of the root system. The following are **crop models** with grassland options. **ARMOSA** (Perego et al., 2013) estimates nitrogen dynamics in soil-crop-atmosphere continuum and evaluates the impact of management on shallow and groundwater quality via modules of energy, water, carbon and nitrogen balances, and plant development and growth. **EPIC**, originally developed to estimate soil productivity as affected by erosion (Williams et al., 2008), is designed to allow simulation of a large variety of crops and grasses with unique parameter values. **STICS** (Brisson et al., 2003) is a generic, daily-step, patch-scaled model covering many crops and conditions of climate, soil and management, being set to simulate either sown or established mowed grasslands. Three **dynamic vegetation models** include grasslands as biome type. **Biome-BGC MuSo** (Hidy et al., 2012) implements a multilayer soil module, improved grassland phenology and management routines into the Biome-BGC, originally developed to simulate undisturbed ecosystems, with allometric relationships used to initialize carbon and nitrogen pools. **CARAIB** (Warnant et al., 1994), a process-based vegetation model of carbon assimilation in the biosphere, implements a range of plant functional types including C_3 and C_4 grasses. Based on the LPJ-Dynamic Global Vegetation Model, **LPJmL** simulates vegetation composition and distribution as well as stocks and land-atmosphere exchange flows of carbon and water (Waha et al., 2012) using generic crop functional types to represent plant prototypes.

Simulations

For the uncalibrated (blind) simulations, we used the nine models above. For blind tests, only weather, soil and management inputs were provided. In this way, blind simulations were run at each site (for the years of available data) using only the above described input information, with no parameter adjustment. After the blind simulations were completed, additional plant and soil information from a sub-set of flux-tower site data was supplied to each modelling

group, i.e. the first half of the whole series of available data or the first half plus one in the case of an uneven number of years. The information provided were daily time series of GPP, soil water content, soil temperature, and actual evapotranspiration. For the same output variables, calibrated simulation results were evaluated against observations from the validation sub-set of years. Biomass data (only available at some dates) were not used for calibration; the full set of biomass data (from both flux tower and other sites) was held back for validation purpose. It was requested that each modelling group adjusts model parameters (especially vegetation parameters) to improve the simulations based on the observed data, using whatever techniques they normally use and documenting the changes.
Seven modelling teams completed the full assessment of that step (two modelling team did not, and the respective models were not further considered in successive analyses). Simulation results from the blind tests over the calibration time period were compared with the measured data over the same period. Seven models were applied for the calibration and validation runs for evaluating fluxes ─ such as Gross Primary Production (GPP), Net ecosystem exchange (NEE), Ecosystem Respiration (RECO), Actual Evapotranspiration (ET), Soil Temperature (ST) and Soil Water Content (SWC) ─ and yield biomass productions across the mentioned long-term grassland sites. In our study, the whole simulation period was split into two parts: the first part was used for calibration tests and the second part for the validation. Four time series were assessed: uncalibration (U1, U2), calibration (C) and validation (V) years (U1 and C refer to the first half of the whole series of available data, or the first half plus one in the case of an uneven number of years, which was used for calibration; U2 and V refer to the years which were excluded from calibration). For flux data, the model simulations were based on daily output values which were further aggregated or averaged to different time resolutions (daily, weekly, monthly and yearly). Most of the simulated sites were mown grasslands thus harvested aboveground biomass values were assessed. The French site of Laqueuille was the only grazed one, for which standing aboveground biomass data were used.
The agreement between simulation and observations was evaluated by the inspection of time series graphs and, numerically, through a set of performance metrics.

Illustrative results

Performances of individual models are discussed on a sample of model outputs and sites. In order to assess the utility of using multi-model ensemble for the simulation of grassland functioning, performance of the multi-model simulation range and median is also assessed against measurement data. We used the median in order to avoid the possible negative effect of outliers in the multi-model ensemble construction. The models are used anonymously in the paper, in order to avoid identification of models providing a specific output. This is done intentionally to ensure that no criticism can be raised against any of the models used, as this aspect is outside the scope of the present work. So, in order not to provide space for raising criticism against any of the used models, the grassland models are only distinguished by using numbers from 1 to 9.

Yield biomass
Calibrated simulations of yearly aggregated harvested biomass at Rothamsted (United Kingdom), a multi-year experimental site (with cuts in June and November and fertilized with 48 kg NH_4-N ha^{-1} yr^{-1} in April), show that some models (Model-3, Model-6) approach the observations with far less bias than others (Figure 2a). Box-plots of harvested/standing aboveground biomass obtained with seven models (Figure 2b) indicate the improvement in model performance (e.g. less number of outliers) obtained with calibration at all investigated sites. Figure 2c illustrates the considerable improvement obtained with the calibration on the estimation of the biomass values at Laqueuille-extensive (LAQ2). The ensemble multi-model median considerably improved after calibration against flux data (adjusted R^2 values from

~0.19 without calibration to ~0.91 after calibration). This means that better simulations can be obtained with some kind of calibration, even without using for it site-specific yield data.

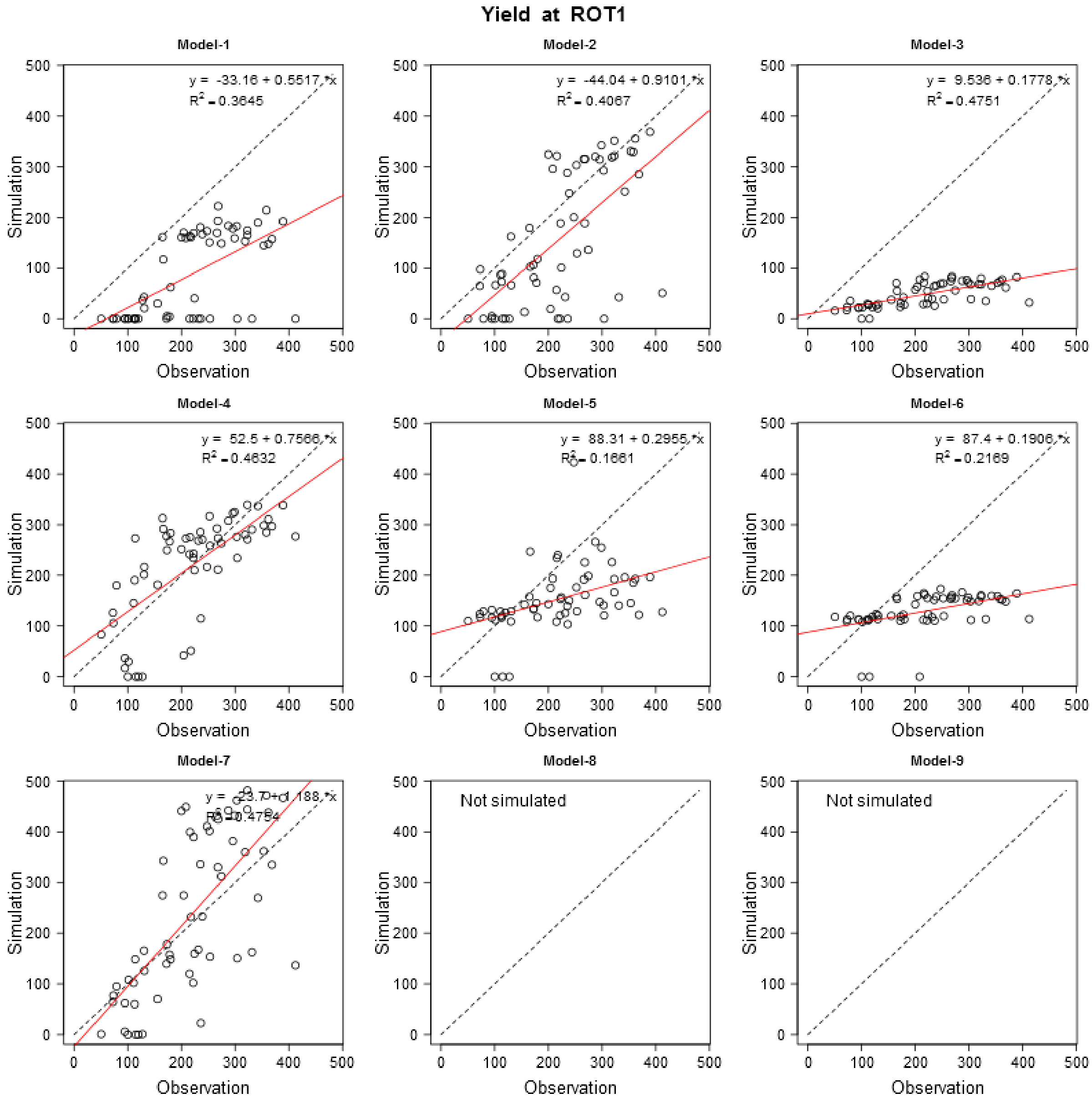

Figure 2a. Calibrated tests: simulated (seven models) versus observed harvested above ground biomass (g DM m^{-2}) at Rothamsted (1981-2011), United Kingdom (solid line: linear regression between simulations and observations; hatched line: 1:1 line).

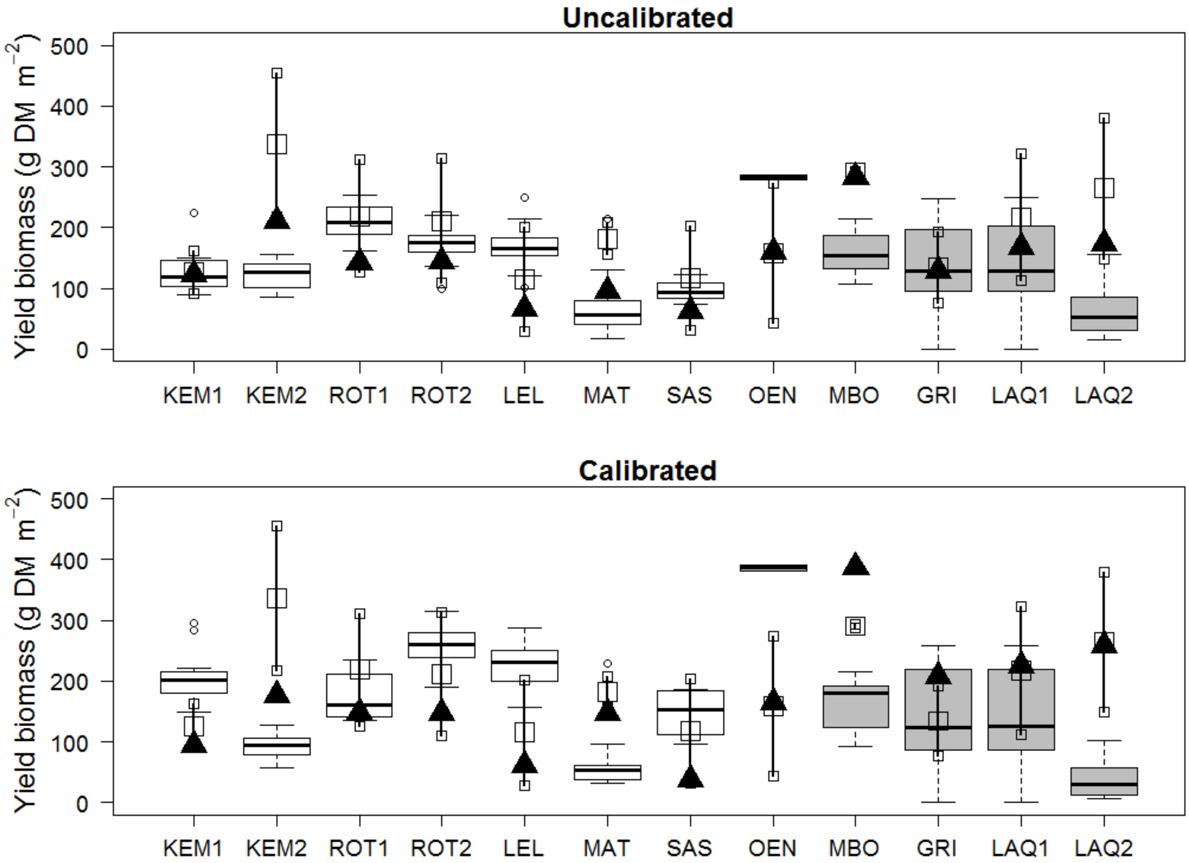

Figure 2b. Observed and simulated yield biomass per cutting event (g DM m^{-2}) for the whole simulated period at 12 location-management combinations with both uncalibrated (nine models) and calibrated (seven) runs. Locations are: Kempten, Germany (KEM1, cut four times per year; KEM2, cut two times per year); Rothamsted, United Kingdom (ROT1, NH_4–N fertilization; ROT2, NO_3–N fertilization); Lelystad, the Netherlands (LEL); Matta, Israel (MAT); Sassari, Italy (SAS); Oensingen, Switzerland (OEN); Monte Bondone, Italy (MBO); Grillenburg, Germany (GRI); Laqueuille, France (LAQ1, intensively fertilized; LAQ2, non-fertilized). Open squares are mean observed yields plus or minus one standard deviation. Filled triangles are the mean of simulated yields for each location. Boxes are delimiting the 25th and 75th percentiles with the median inside. Whiskers are 10th and 90th percentiles. Hollow circles indicate outliers.

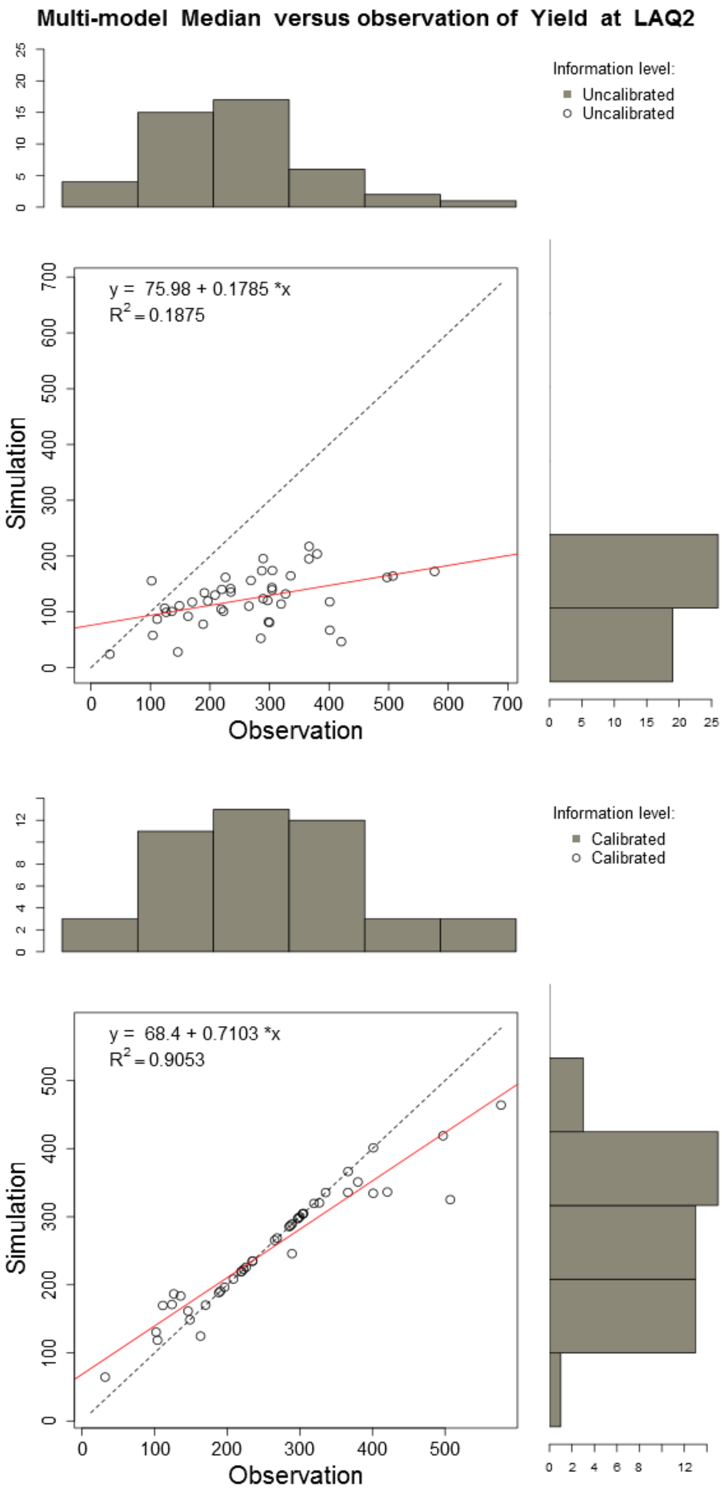

Figure 2c. Uncalibrated (top) and calibrated (bottom) standing aboveground biomass (g DM m^{-2}) multi-model median simulations compared with observed values at Laqueuille-extensive site (France) in the period 2004-2010. Solid red line: linear regression between simulations and observations; hatched line: 1:1 line; grey bars: frequencies of the values.

Gross primary productivity (GPP)

Another example (Figure 3) refers to gross primary production (weekly values). Calibration and validation simulations (accomplished by five models only) were compared to observations at the Swiss site of Oensingen, where the grassland was mown 3-4 times each year and highly fertilized (more than 200 kg N ha^{-1} yr^{-1} split into four events).

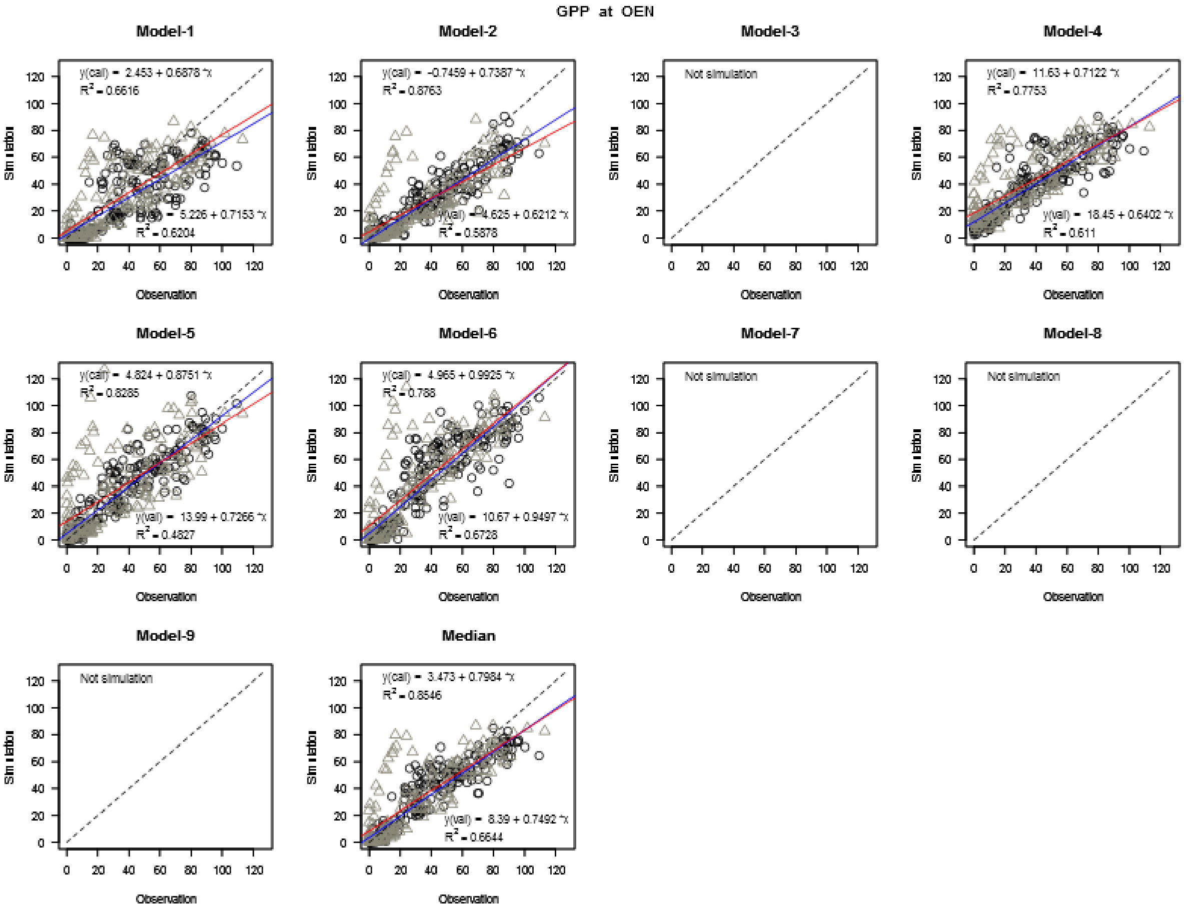

Figure 3. Calibration and validation tests for individual models and median of the model ensemble: simulated versus observed gross primary production (GPP, g C m^{-2} $week^{-1}$) at Oensingen (2002-2009), Switzerland (solid line: linear regression between simulations and observations; hatched line: 1:1 line). Only five models provided GPP. Black open circles are the calibration simulations while the grey triangles indicate the validation tests. Red line is the calibration regression line whilst the blue line is the validation regression line.

Compared with the blind parameterizations, regression lines (Figure 3) indicate that we had some improvements after the calibration was performed, which roughly matched GPP observations for all models (slope and intercept near 1 and 0, respectively; adjusted R^2 were ~0.6 but the R^2 of the multi-model median was ~0.86 after the calibration). The uncertainty envelope obtained with the ensemble of model estimates (Figure 4) shows that the influence of extreme events such as the hot and dry summer 2003 can lead to an amplification of uncertainties.

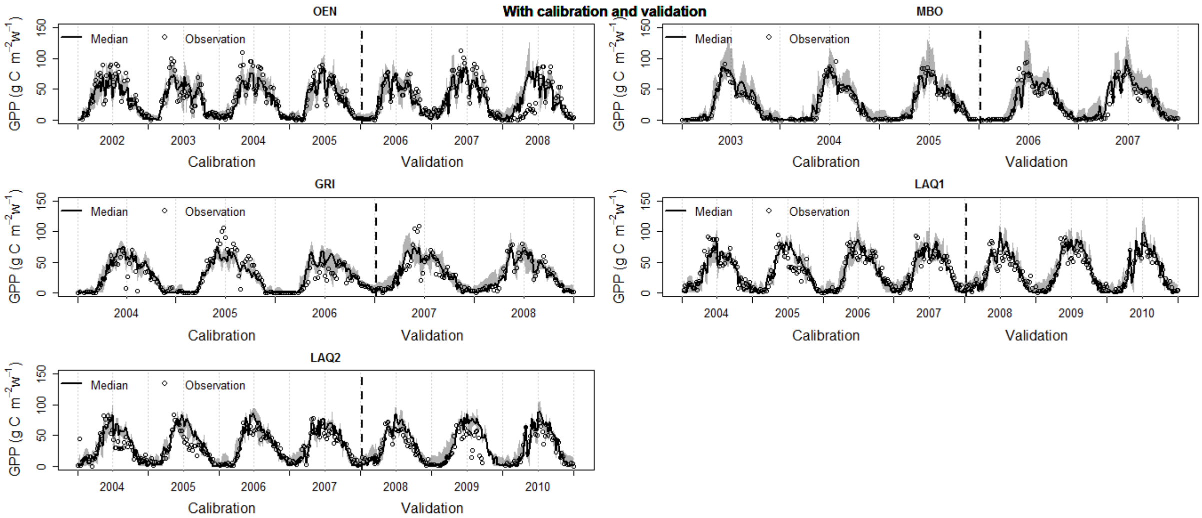

Figure 4. Calibration and validation tests: fluctuations of simulated (median of five models) and observed gross primary production (GPP, g C m^{-2} $week^{-1}$) at the flux sites.

Ecosystem respiration (RECO)

Figure 5 shows the ensemble model uncertainty of the weekly based RECO values with different information levels such as blind (Uncalibrated-1 and -2), calibration and validation tests. By comparing the related time periods (uncalibrated-1 with calibration; uncalibrated-2 with validation), we obtained some improvement at Oensingen (OEN), Monte Bondone (MBO), Grillenburg (GRI) and Laqueuille-intensive (LAQ1). However, the models roughly estimated the RECO values at Laqueuille-extensive (LAQ2), which is grazed and non-fertilized. On the whole, the simulated mean values better approached the measured values after the calibration and validation. Therefore, the calibration work is a necessary and not negligible task in order to achieve reasonable RECO results.

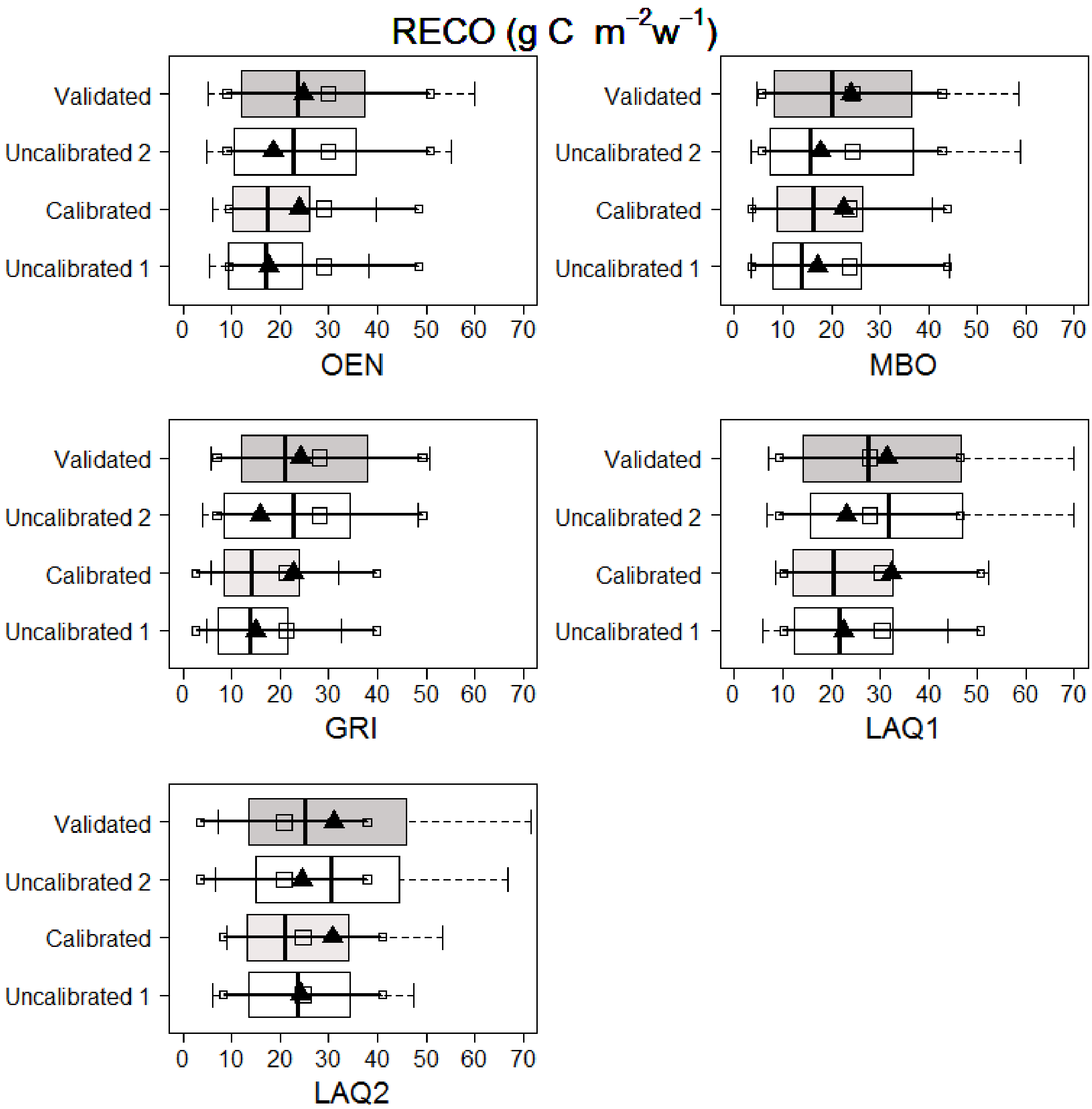

Figure 5. Observed and simulated average weekly ecosystem respiration (RECO, g C m^{-2}) for the whole simulated period at flux sites using nine models for uncalibrated runs and seven models for calibration and validation runs. Open squares are mean observed RECO values plus/minus the standard deviation. Filled triangles are the mean of simulated RECO values for each location. Boxes are delimiting the 25th and 75th percentiles with the median inside. Whiskers are 10th and 90th percentiles. Hollow circles indicate outliers.

Net ecosystem exchange (NEE)

For NEE, Figure 6 indicates that the mean of the multi-model median simulations approached the observed values better than individual model performances. There are some improvements after calibration and validation in the case of RRMSE and R^2 calculations.

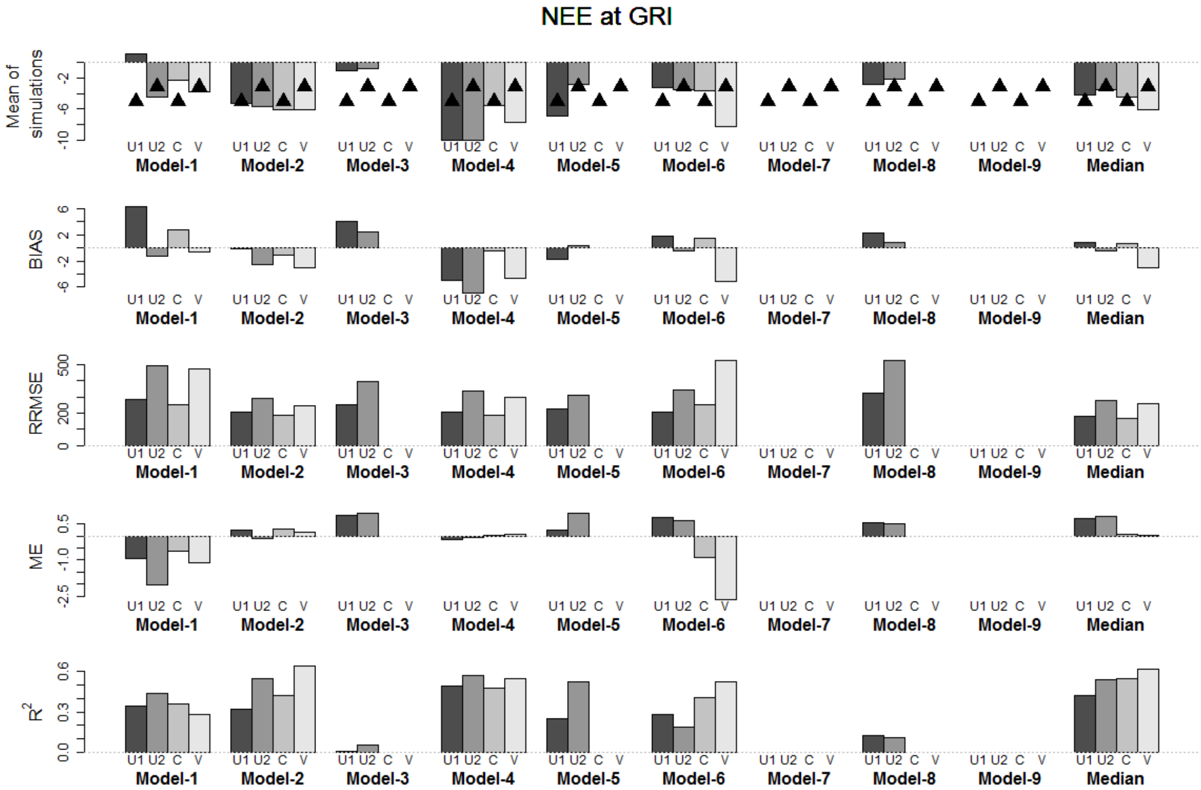

Figure 6. Statistical analyses of the weekly averaged NEE (g C m^{-2} w^{-1}) values at Grillenburg (GRI) with different information levels (U1: uncalibrated, U2: uncalibrated, C: calibration, V: validation) obtained by individual models (models 1 to 9) and multi-model median simulations. Black triangles show the weekly averaged observations. BIAS: mean difference between simulations and observations; RRMSE: relative root mean square error (%); ME: modelling efficiency; R^2: goodness-of-fit of simulations versus observations regression line.

Concluding remarks

This study (blind tests, calibration and validation simulations) focussed on various sites across Europe and peri-Mediterranean regions, extends parallel initiatives on the comparison of grassland models worldwide, such as AgMIP and other international projects (https://colloque.inra.fr/workshop_gra_jpi_facce_eng/2-Model-Intercomparison).

The results shown are illustrative of the methodology adopted for grassland model inter-comparison in MACSUR. The insights gained from this study are relevant for some crop and vegetation models, which in some cases proved comparable to grassland-specific models to simulate biomass data from managed grasslands. The results reported here integrate other published ones (Sándor et al., 2015; Wu et al., 2015). Additional results will be published as they become available together with the sensitivity of models to climate change factors.

Acknowledgements

This report is a contribution to the FACCE MACSUR knowledge hub. We thank all Partner Institutions for providing datasets and running models. We also acknowledge technical support from the European Fluxes Database Cluster (http://www.europe-fluxdata.eu).

References

Amman, C., Flechard, C.R., Leifeld, J., Neftel, A., Fuhrer, J., 2007. The carbon budget of newly established temperate grassland depends on management intensity. Agr. Ecosyst. Environ. 121, 5–20.

Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., Zimmer, D., Sierra, J., Bertuzzi, P., Burger, P., Bussiere, F., Cabidoche, Y.M., Cellier, P., Debaeke, P., Gaudillere, J.P., Henault, C., Maraux, F., Seguin, B., Sinoquet, H., 2003. An overview of the crop model STICS. Eur. J. Agron. 18, 309-332.

Cavallero, A., Talamucci, P., Grignani, C. Reyneri, A., Ziliotto, U., Scotton, M., Bianchi, A.A., Santilocchi, R., Basso, F., Postiglione, L., Carone, F., Corleto, A., Cazzato, E., Cassaniti, S., Cosentino, S., Litrico, P.G., Leonardi, S., Sarno, R., Stringi, L., Gristina, L., Amato, G., Bullitta, P., Caredda, S., Roggero, P.P., Caporali, F., D'Antuono, L.F., Pardini, A., Zagni, C., Piemontese, S., Pazzi, G., Costa, G., Pascal, G., Acutis, M., 1992. Caratterizzazione della dinamica produttiva di pascoli naturali italiani. Rivista di Agronomia 26, n. 3 suppl., 325-343. (in Italian)

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

De Martonne, E., 1942. Nouvelle carte mondiale de l’indice d’aridité. Annales de Géographie 51, 242–250. (in French)

Diodato, N., Ceccarelli, M., 2004. Multivariate indicator Kriging approach using a GIS to classify soil degradation for Mediterranean agricultural lands. Ecol. Indic. 4, 177–187.

Golodets, C., Sternberg, M., Kigel, J., Boeken, B., Henkin, Z., Seligman, N.G., Ungar, D.E., 2013. From desert to Mediterranean rangelands: will increasing drought and inter-annual rainfall variability affect herbaceous annual primary productivity? Climatic Change 119, 785-798.

Graux A.-I., Bellocchi G., Lardy R., Soussana J.-F., 2013. Ensemble modelling of climate change risks and opportunities for managed grasslands in France. Agric. Forest Meterol. 170, 114-131.

Hidy, D., Barcza, Z., Haszpra, L., Churkina, G., Pintér, K., Nagy, Z., 2012. Development of the Biome-BGC model for simulation of managed herbaceous ecosystems. Ecol. Modell. 226, 99–119.

Klumpp, K., Tallec, T., Guix, N., Soussana, J.-F., 2011. Long-term impacts of agricultural practices and climatic variability on carbon storage in a permanent pasture. Global Change Biol. 17, 3534–3545.

Köchy, M., 2008. Effects of simulated daily precipitation patterns on annual plant populations depend on life stage and climatic region. BMC Ecology 8:4, doi:10.1186/1472-6785-8-4.

Ma, S., Lardy, R., Graux, A.-I., Ben Touhami, H., Klumpp, K., Martin, R., Bellocchi, G., 2015. Regional-scale analysis of carbon and water cycles on managed grassland systems. Environ. Modell. Softw., doi: 10.1016/j.envsoft.2015.03.007.

Perego, A., Giussani, A., Sanna, M., Fumagalli, M., Carozzi, M., Alfieri, L., Brenna, S., Acutis, M., 2013. The ARMOSA simulation crop model: overall features, calibration and validation results. Italian Journal of Agrometeorology 3, 23–38.

Prescher, A.-K., Grünwald, T., Bernhofer, C., 2010. Land use regulates carbon budgets in eastern Germany: from NEE to NBP. Agr. Forest Meteorol. 150, 1016–1025.

Sándor, R., Ma, S., Acutis, M., Barcza, Z., Ben Touhami, H., Doro, L., Hidy, D., Köchy, M., Lellei-Kovács, E., Minet, J., Rolinski, S., Ruget, F., Seddaiu, G., Wu, L., Perego, A., Bellocchi, G., 2015. Uncertainty in simulating biomass yield and carbon – water fluxes from grasslands under climate change. Advances in Animal Biosciences 6, 49-51.

Schils, R., Snijders, P., 2004. The combined effect of fertiliser nitrogen and phosphorus on herbage yield and changes in soil nutrients of a grass/clover and grass-only sward. Nutr. Cycl. Agroecosys. 68, 165–179.

Schröpel, R., Diepolder, M., 2003. Auswirkungen der Grünlandextensivierung auf einer Weidelgras-Weißklee-Weide im Allgäuer Alpenvorland. Schuleund Beratung, Heft 11/2003, Seite III-13 bis III-15; Bayerisches Staatsministerium für Landwirtschaft und Forsten, Munich, Germany. (in German)

Silvertown, J., Poulton, P., Johnston, E., Edwards, G., Heard, M., Biss, P.M., 2006. The Park Grass Experiment 1856–2006: its contribution to ecology. J. Ecol. 94, 801–814.

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.

Waha, K., van Bussel, L.G.J., Müller, C., Bondeau, A., 2012. Climate-driven simulation of global crop sowing dates. Global Ecol. Biogeogr. 21, 247-259.

Warnant, P., François, L., Strivay, D., Gérard, J.-C., 1994. CARAIB: a global model of terrestrial biological productivity. Global Biogeochem. Cy. 8, 255-270.

Williams, J.R., Arnold, J.G., Kiniry, J.R., Gassman, P.W., Green, C.H., 2008. History of model development at Temple, Texas. Hydrol. Sci. J. 53, 948-960.

Wohlfahrt, G., Anderson-Dunn, M., Bahn, M., Balzarolo, M., Berninger, F., Campbell, C., Carrara, A., Cescatti, A., Christensen, T., Dore, S., Eugster, W., Friborg, T., Furger, M., Gianelle, D., Gimeno, C., Hargreaves, K., Hari, P., Haslwanter, A., Johansson, T., Marcolla, B., Milford, C., Nagy, Z., Nemitz, E., Rogiers, N., Sanz, M.J., Siegwolf, R.T.W., Susiluoto, S., Sutton, M., Tuba, Z., Ugolini, F., Valentini, R., Zorer, R., Cernusca, A., 2008. Biotic, abiotic, and management controls on the net ecosystem CO_2 exchange of European mountain grassland ecosystems. Ecosystems 11, 1338–1351.

Wu, L., 2007. SPACSYS: Integration of a 3D root architecture component to carbon, nitrogen and water cycling - Model description. Ecol. Modell. 3-4, 343–359.

Wu, L., Whitmore, A.P., Bellocchi, G., 2015. Modelling the impact of environmental changes on grassland systems with SPACSYS. Advances in Animal Biosciences 6, 37-39.