The **CLIMSAVE** Project

Climate Change Integrated Assessment Methodology for Cross-Sectoral Adaptation and Vulnerability in Europe

Report describing the development and validation of the sectoral meta-models for integration into the IA platform

Edited by Ian Holman[1] and Paula Harrison[2]

[1] *Environmental Science and Technology Department, Cranfield University, UK*
[2] *Environmental Change Institute, University of Oxford, UK*

December 2012

Table of contents

1. Introduction

Ian Holman
Environmental Science and Technology Department, Cranfield University, UK

1.1 Background to Deliverable 2.2

Deliverable 2.2 reports on one of the Tasks associated with the development of the CLIMSAVE Integrated Assessment Platform (IAP):

- Task 2.4 - Development and validation of the meta-models within the IA platform.

However, because the development of the meta-models is so intrinsically linked to Task 2.2 (Development of the meta-model specifications), the outcomes from this Task which were reported in D2.1 (Holman & Cojocaru, 2010) are first summarised.

Given the participatory approach to the design and development of the CLIMSAVE Integrated Assessment Platform (van Asselt & Rijkens-Klomp 2002), we anticipate that the IAP and the associated meta-models will undergo modifications throughout the duration of the project in response to progressive stakeholder feedback from the activities of Work Packages (WP) 1 and 3 and from direct stakeholder engagement via the CLIMSAVE website. As such, the activities described in this report represent 'works in progress', rather than being 'set-in-stone'.

1.2 References

Holman, I.P. & Cojocaru, G. (2010). Deliverable 2.1 - A report describing the IA Platform specification, metamodel specifications and the multi-scale approach. CLIMSAVE EC FP7 Project 244031.

van Asselt MBA, Rijkens-Klomp N (2002). A look in the mirror: reflection on participation in integrated assessment from a methodological perspective. Global Environmental Change, 12: 167-184.

2. Summary of the development of the meta-model specifications (Task 2.2)

Ian Holman[1] and George Cojocaru[2]
[1] *Environmental Science and Technology Department, Cranfield University, UK*
[2] *TIAMASG Foundation, Bucharest, Romania*

2.1 Summary of Deliverable 2.1

A meta-modelling approach is being used in CLIMSAVE whereby computationally-efficient or reduced-form models that emulate the performance of more complex models are being developed to deliver the fast run times required by the IA Platform. For efficient development of the CLIMSAVE IAP, each of the meta-models (described in the proceeding sections) are designed to be modular, independent and capable of replacement at any time. A meta-model specification was therefore developed to ensure successful linkage and integration of the meta-models, irrespective of the final algorithms inside each of the meta-models. The specifications have been defined in relation to anticipated stakeholder needs (CLIMSAVE WP1), the vulnerability framework (WP5), the scenario methodology and climate and socio-economic scenario variables (WP3) and the requirements of the adaptive capacity methodology (WP4), plus some redundancy for future development.

The development of the specification went through five distinct stages:

1. Defining the spatial resolution of the data to be transferred between meta-models;
2. Identifying and prioritising meta-model inputs and outputs;
3. Identifying points of contact between the meta-models;
4. Specifying the data dictionaries for each meta-model;
5. Standardising the data dictionaries across all of the meta-models.

For the European scale case study application of the CLIMSAVE IAP, the spatial scale of data transfer between the meta-models represents a compromise between the scale of available harmonised datasets, model runtime and spatial detail of the outputs. The higher the resolution at which the IAP operates, the greater is the number of times that the meta-models have to run and hence the greater the overall runtime of the IAP. It was agreed that the European CLIMSAVE IAP would operate at a resolution of 10' x 10' (10 minute by 10 minute), using the same grid as the Climatic Research Unit's baseline 1961-90 baseline climatology (CRU CL 2.1- Mitchell *et al.*, 2003). This represents over 23,000 land-based grid squares across the CLIMSAVE European case study area. It has similarly been agreed that the Scottish IAP will use a resolution of 5km x 5km.

In order to deliver the fast web-based response time demanded by this application, a process of meta-modelling is being carried out on a set of tried and tested desktop models to abstract the leanest representation consistent with delivering both functionality and speed. Based upon the state-of-the-art sectoral impact models available to the consortium (as outlined in the Description of Work), model inputs and output were identified by the modellers and rated for stakeholder-relevance by the wider CLIMSAVE consortium. For the model inputs, the prioritisation was based on their relevance to adaptation responses, whilst the model outputs were prioritised according to perceived stakeholder relevance (e.g. areas at risk of flooding

and flood damages) and/or policy relevance (e.g. rural land-use allocation for intensive agriculture, extensive agriculture, abandoned land, etc).

Points of contact were also identified between the meta-models (Figure 2.1) – these are the linkages and influences between sectors, and represent data transfers between the models. For example, following the flow arrows from the RUG model in Figure 2.1, the simulated area, location and type of urban development ("artificial surfaces" and "residential/non-residential development" from the urban model – RUG) affects the population exposed to flood risk ("People affected" as estimated by the Flood Model), river basin hydrological response ("Basin flow" from WaterGAP-H), the land available for agriculture and forestry ("landuse allocation" from the land allocation model – SFarmMod) and consequently habitat availability (biodiversity model – SPECIES and LPJ-GUESS).

Within any single simulation of the CLIMSAVE IAP, there will be five components of data reading and transfers:

1. Data transfers from the user to the meta-models, representing the communication of input parameter values from the user (slider bars, timeslice, scenarios, etc) to the models, via the Running Module;
2. Data transfers between the meta-models, where the simulated output from one meta-model is an input to other meta-models;
3. Data transfers from the IAP database to the meta-models containing, for example, the input data for a user-selected scenario;
4. Data transfers between the meta-models and the user Interface, as outputs are selected by the user for display;
5. Data that is read into a meta-model from the meta-model's own internal dataset.

With the exception of (5), all of the above represent transfers of data which need to be clearly defined in a transparent way for the consortium. Data dictionaries have therefore been developed for data associated with (1) – (4), which unambiguously define each variable or parameter and its characteristics. The final step in the process is the standardisation of the data dictionaries across all of the meta-models, so that each end (IAP, database or meta-model) of a data transfer (for example, meta-model to meta-model; or IAP to meta-model) uses the same data dictionary. This then allows the data transfers in terms of model variables and parameters to be defined (Figure 2.2).

The meta-models themselves are implemented as Dynamic-Link Libraries (DLL) developed in various software languages: Microsoft C++, Microsoft C#, Microsoft VB, and Delphi as both managed and unmanaged code. They will be embedded in the main Running module, working as one piece of software. The Running module will feed the DLLs with data, run the DLLs and collect the outputs. The exchange of data will be made available based on structures of data transferred by pointers to minimise the time required for data exchange. In this approach, the meta-model is told where to point data within the internal memory, rather than the data being physically transferred to the model, with consequent time savings given the number of grid cells (>23,000).

2.2 References

Mitchell, T.D., Carter, T.R., Jones, P.D., Hulme, M. & New, M. (2003). A comprehensive set of climate scenarios for Europe and the globe. Tyndall Centre Working Paper 55.

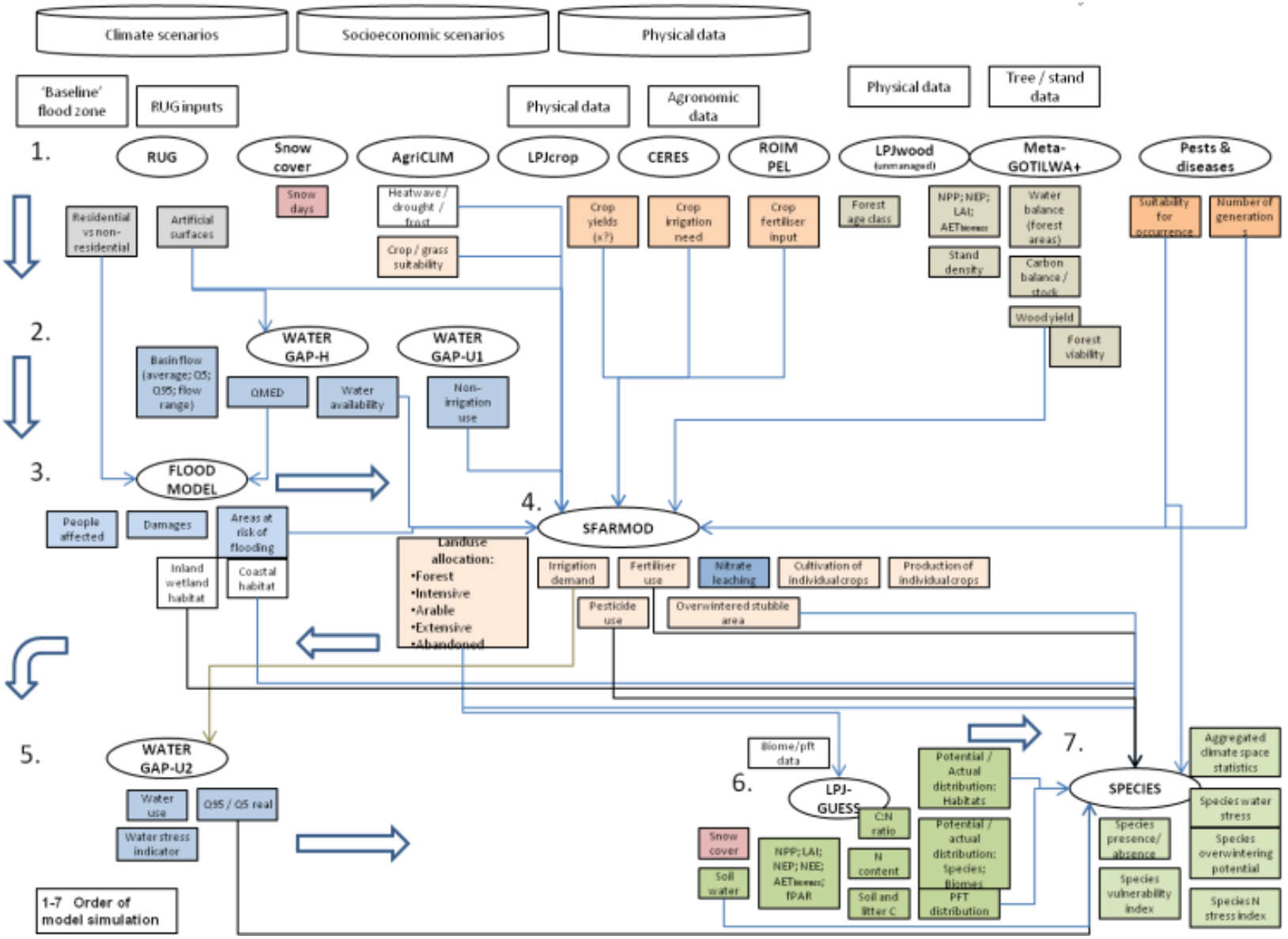
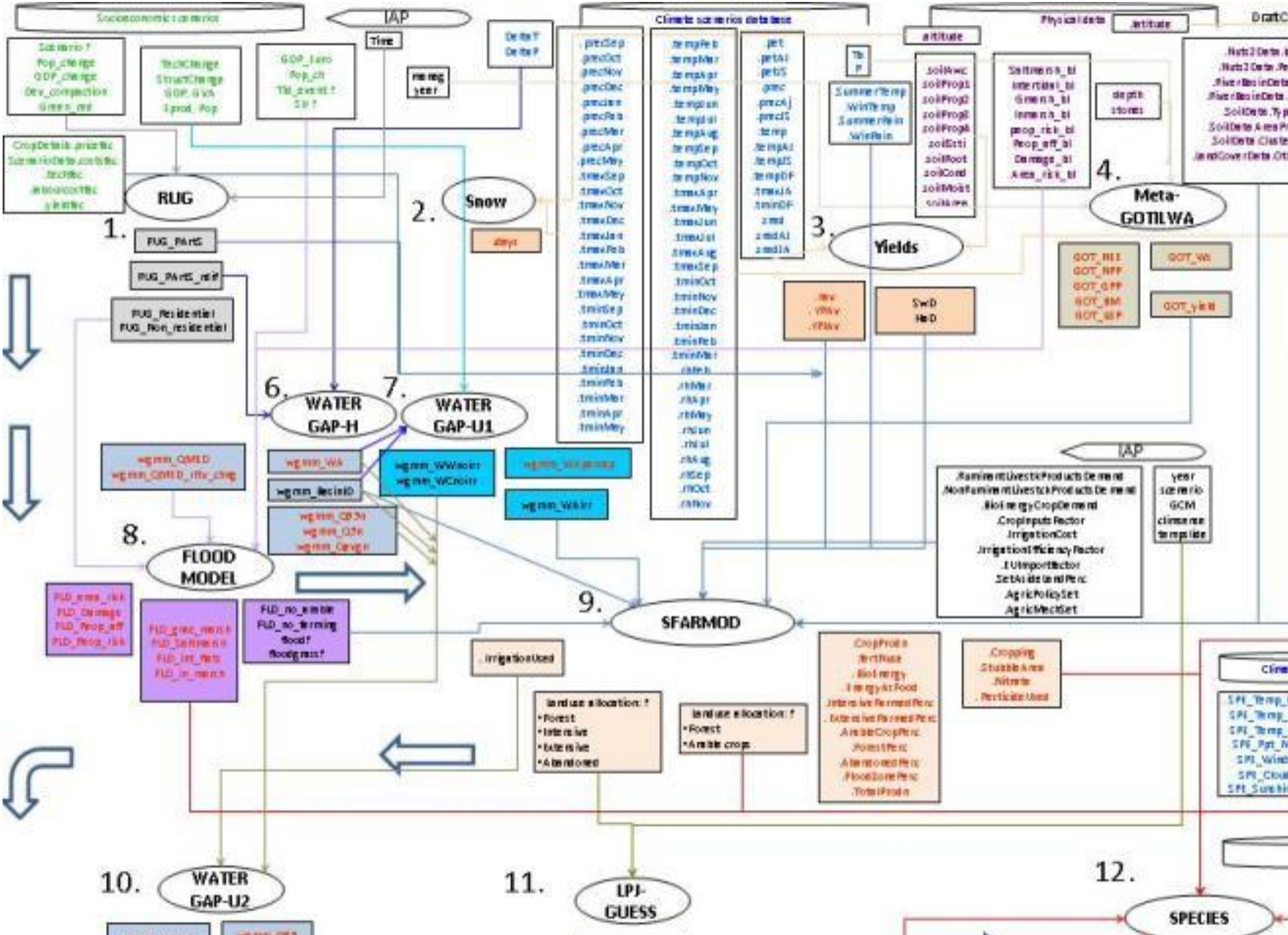


Figure 2.1: Schematic of the data interactions between the meta-models [Ovals - meta-models; open rectangles – data inputs from the databases; shaded rectangles – meta-model outputs; numbering and large open arrows – order of operation of the meta-models].


Socioeconomic scenarios
IAP
Time
Climate scenarios database
Physical data
Attribute
RUG
1.
2.
Snow
3.
Yields
4.
Meta-GOTILWA
WATER GAP-H
6.
WATER GAP-U1
7.
FLOOD MODEL
8.
SFARMOD
9.
WATER GAP-U2
10.
LPJ-GUESS
11.
SPECIES
12.

3. Introduction to Task 2.4 - Development and validation of the meta-models within the IA platform

Ian Holman and George Cojocaru
[1] *Environmental Science and Technology Department, Cranfield University, UK*
[2] *TIAMASG Foundation, Bucharest, Romania*

3.1 Summary of Task 2.4

The CLIMSAVE consortium brings together a number of participants with expertise in developing participatory integrated assessment platforms, such as the Regional Impact Simulator (Holman *et al.*, 2008a;b), CLIMPACTS (Kenny *et al.*, 2000) and SimCLIM (Warrick *et al.*, 2005). Participatory IA platforms are a vehicle for communication, training, forecasting and experimentation (Welp, 2001, Kasemir *et al.*, 2003, Jäger *et al.*, 2008), whose usefulness is enhanced by the integrated assessment approach which enables stakeholders to explore / understand the interactions between different sectors, rather than viewing their own area in isolation. An assessment of stakeholder needs for, and perspectives on, integrated assessment platforms showed that stakeholders desired to be able to perform their own integrated assessment - investigating the impacts and adaptive responses of relevance to themselves, rather than having to rely on the restricted outputs generated from a limited number of simulations chosen arbitrarily by researchers (Holman *et al.*, 2005; 2008a). However, stakeholder involvement is discouraged in most IAs by the complex software and unacceptably long runtimes (Wolfe *et al.*, 2001). Holman et al. (2008a) developed the use of computationally simpler modelling techniques, so called 'meta-models' or 'reduced form models' (Carmichael *et al.*, 2004), within a user-friendly interface and evaluated stakeholder experience (Holman *et al.*, 2008b).

The development of the CLIMSAVE integrated assessment platform, and its constituent meta-models, has learnt from this unique process. The following sections describe the development and validation of each of the meta-models describing key European sectors (agriculture, forests, water, coasts, biodiversity and urban). The meta-models each simulate a range of stakeholder-relevant impact indicators and indicators which translate the outputs from the integrated sectoral models into ecosystem services indicators (Table 3.1). Ecosystem services cover all key European sectors, such as cultivated ecosystems, forest ecosystems, inland water ecosystems, coastal ecosystems, natural ecosystems and urban ecosystems. They closely correspond to the key sectors studied by Working Group II of the IPCC Fourth Assessment Report (IPCC, 2007) and enable climate change impacts to be linked directly to human well-being.

After the following sections which describe each of the meta-models in turn, Section 14 concludes by summarising (Table 14.1) how the stakeholder-relevant indicators simulated by the meta-models link to the ecosystem services in Table 3.1.

Table 3.1: List of ecosystem services according to the Millennium Ecosystem Assessment (MA).

MA category	Ecosystem service
Provisioning services	Food
	Fibre
	Fuel/energy
	Genetic resources
	Biochemical/natural medicines
	Ornamental resources
	Fresh water
Regulatory services	Pollination
	Seed dispersal
	Pest regulation
	Disease regulation
	Climate regulation
	Air quality regulation
	Water regulation
	Erosion regulation
	Natural hazard regulation
	Invasion resistance
	Water purification/waste treatment
Cultural services	Spiritual and religious values
	Education and inspiration
	Recreation and ecotourism
	Cultural heritage
	Aesthetic values
	Sense of place
Supporting services	Primary production
	Photosynthesis
	Provision of habitat
	Soil formation and retention
	Nutrient cycling
	Water cycling

3.2 References

Carmichael, J., Tansey, J. & Robinson, J. (2004). An integrated assessment modelling tool. Global Environmental Change, 14: 171-183.

Holman, I.P., Rounsevell, M.D.A., Shackley, S., Harrison, P.A., Nicholls, R.J., Berry, P.M. & Audsley, E. (2005). A regional, multi-sectoral and integrated assessment of the impacts of climate and socio-economic change in the UK: I Methodology. Climatic Change, 71, 9-41.

Holman, I.P., Rounsevell, M.D.A., Berry, P.M. & Nicholls, R.J. (2008a). Development and application of participatory integrated assessment software to support local/regional impact and adaptation assessment. Climatic Change, 90(1-2), 1-5.

Holman, I.P., Rounsevell, M.D.A., Cojocaru, G., Shackley, S., McLachlan, C., Audsley, E., Berry, P.M., Fontaine, C., Harrison, P.A., Henriques, C., Mokrech, M., Nicholls, R.J.,

Pearn, K.R. & Richards, J.A. (2008b). The concepts and development of a participatory regional integrated assessment tool. Climatic Change, 90(1-2), 5-30.

IPCC (2007). Summary for Policymakers. In: Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Parry, M.L., Canziani, O.F., Palutikof, J.P., van der Linden, P.J., Hansen, C.E. (Eds.)]. Cambridge University Press, Cambridge, UK.

Kasemir, B., Jäger, J., Jaeger, C.C. & Matthew, G. (Eds) (2003). Public Participation in Sustainability Science. Cambridge University Press, Cambridge.

Kenny, G.J., Warrick, R.A., Campbell, B.D., Sims, G.C., Camilleri, M., Jamieson, P.D., Mitchell, N.D., McPherson, H.G. & Salinger, M.J. (2000). Investigating climate change impacts and thresholds: An application of the CLIMPACTS integrated assessment model for New Zealand agriculture. Climatic Change, 46(1-2), 91-113.

Warrick, R.A., Ye, W., Kouwenhoven, P., Hay, J.E. & Cheatham, C. (2005). New developments of the SimCLIM model for simulating adaptation to risks arising from climate variability and change. In: Zerger, A. & Argent, R.M. (Eds.) MODSIM 2005. International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December 2005, pp. 170-176.

Welp, M. (2001). The use of Decision Support Tools in participatory river basin management. Phys. Chem. Earth (B), 26(7-8), 535-539.

Wolfe, A.K., Kerchner, N. & Wilbanks, T. (2001). Public involvement on a regional scale.
Environmental Impact Assessment Review, 21, 431-448.

4. Development and validation of the snow cover meta-model

Miroslav Trnka
Institute of Agrosystems and Bioclimatology, Mendel University, Brno, Czech Republic

4.1 Snow cover model description

The snow cover meta-model is based on the SnowMAUS snow cover simulator (Trnka *et al.*, 2010). The core algorithm used in the snow cover model for agrometeorological use (snowMAUS) was proposed by Running & Coughlan (1988) and was modified by Trnka *et al.* (2010). The snowMAUS model operates on a daily time step, with seven key parameters that govern snow accumulation and melting. Snow melting is usually facilitated by other factors, such as sublimation, sun-driven ablation and often combined with the influence of wind. These factors cannot be directly considered due to the nature of the available input data and were summed into a single empirical factor.

Data was gathered from 1948-2002 from 65 sites across Austria (Figure 4.1), which exhibited considerable variability in elevation (155-3111 m a.s.l.). Of these stations, 65% were located at altitudes below 800 m, where most agricultural activity takes place. Four of these sites within the crop-growing altitude range (Irdning [A], Pabneukirchen [B], Gleisdorf [C] and Hohenau [D]) were randomly selected and the model calibrated for the period 1948-2002 (Figure 4.1). In order to test newly introduced routines and to verify the stability of the selected thresholds, an extensive sensitivity analysis using the Monte-Carlo method was undertaken.

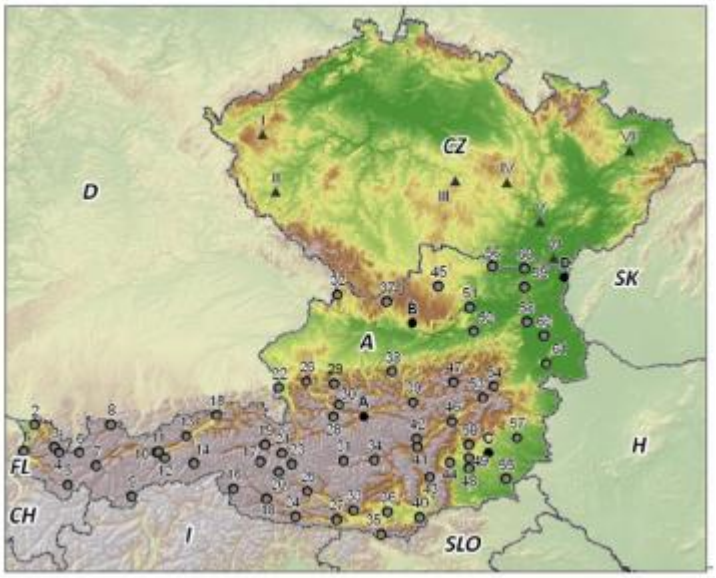


Figure 4.1: Area within which SnowMAUS model was originally calibrated and validated.

The remaining sites served as independent tests of model performance and included several high elevation stations where agriculture production was limited to hay production and/or grazing. The datasets consisted of quality-controlled and homogenised daily surface weather records, including observations of daily maximum and minimum air temperatures at 2 m

above the surface, total daily precipitation, precipitation type, daily values of snow cover height and continuity of the snow cover. Precipitation that was recorded as 'trace' was replaced with 0.0 mm, which had no significant effect on the precipitation totals. The snow cover volume was expressed in terms of water equivalent in mm. Years with incomplete observations of snow cover or precipitation during the winter season were excluded from the analysis. An overview of the station locations is provided in Figure 4.1.

The snowMAUS model effectively captured daily values of snow cover in terms of snow water equivalent (Figure 4.2) across a large altitudinal gradient. The model was able to explain, on average, 73% (ranging from 42 to 89%) of the variability in the number of days with snow during individual seasons and, on average, 81% (ranging from 31 to 97%) of the variability in the seasonal volume of snow between 1948 and 2002. The snowMAUS model captured over 96% and 98% of the between-site variability in the number of days with snow and the volume of precipitation in the form of snow, respectively (Figure 4.2). Despite acceptable overall performance, the model overestimates snow cover at lowland stations and underestimates snow cover at high elevations for some seasons; however on the level of long-term means (as applied in CLIMSAVE) this has marginal importance.

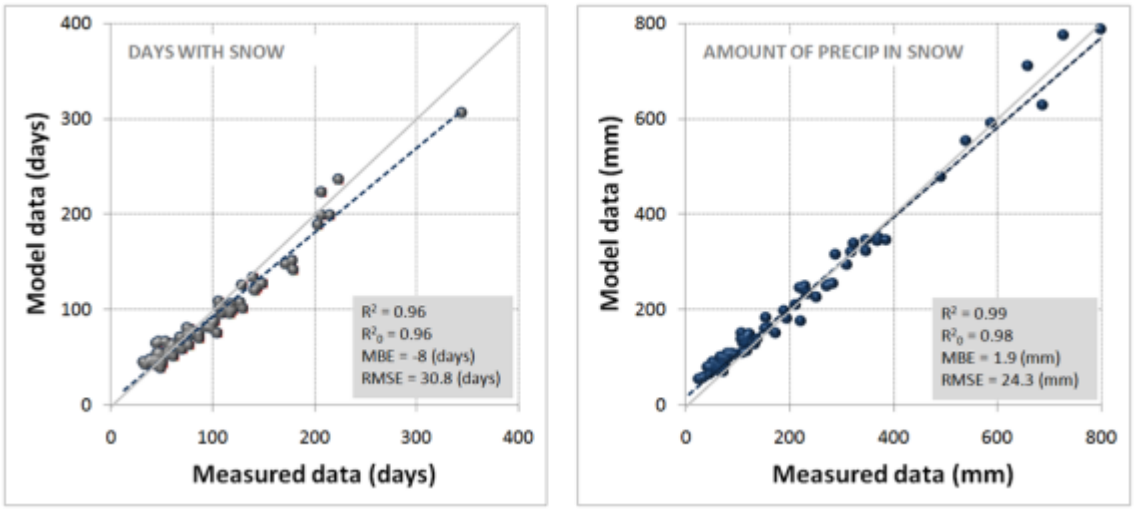


Figure 4.2: Results of the SnowMAUS model validation at 61 sites in terms of long-term climatology (1948-2002) of snow cover.

4.2 Development and validation of the SnowCover meta-model

In order to develop a snow cover meta-model based on SnowMAUS that would be applicable over the wider CLIMSAVE European domain, new datasets were acquired based on the COST734 database (Trnka *et al.*, 2011). In this database, the 83 sites (Figure 4.3) with high quality daily weather data needed for SnowMAUS runs were available both for baseline (1971-2000) climate as well as for climate conditions around 2050 (using three global circulation models runs and A2 emission scenario). The downscaling of these scenarios was based on the pattern-scaling technique combining the MAGICC model with the outputs of three GCMs (HadCM, NCAR and ECHAM). In addition a dataset assuming +5°C warming (and its regionalization through the pattern scaling technique) was developed to calculate snow cover parameters under more extreme warming in order to provide the meta-models with a sufficiently broad range of climate conditions. In all cases (baseline, 2050 A2 driven scenario and +5°C dataset) for each station and each GCM, a stochastic weather generator was used to produce 100-year long daily data series in order to increase the sample size.

Details on the development of the daily scenarios using the pattern scaling technique and the dataset used is available in Trnka *et al.* (2011).

The SnowCover meta-model was based on artificial neural networks (ANNs; Qnet, 2000) that were calibrated and tested using outputs of the more detailed SnowMAUS model. The input variables used for the model development include monthly maximum and minimum air temperatures for months from October till May as well as precipitation means during these months. In addition, an oceanity index (annual temperature range divided by latitude of the grid) is used to account for the moderating effect of the ocean. The model was calibrated on a training set of data that was sampled to cover the whole range of predictors and the predicted variable, i.e. number of days with snow. The sampling of the calibration dataset took into account values outside ± 1 standard deviation from the mean of each parameter. This model was then independently tested on the complementary validation dataset in order to calculate statistics of its performance accuracy. In total, 12 different ANN designs were tested with the most suitable one being selected on the basis of the variability explained (R^2) and the root mean square error (RMSE). For the final design, 20 different initiations for the ANN were tested, but no significant difference in the outputs was found.

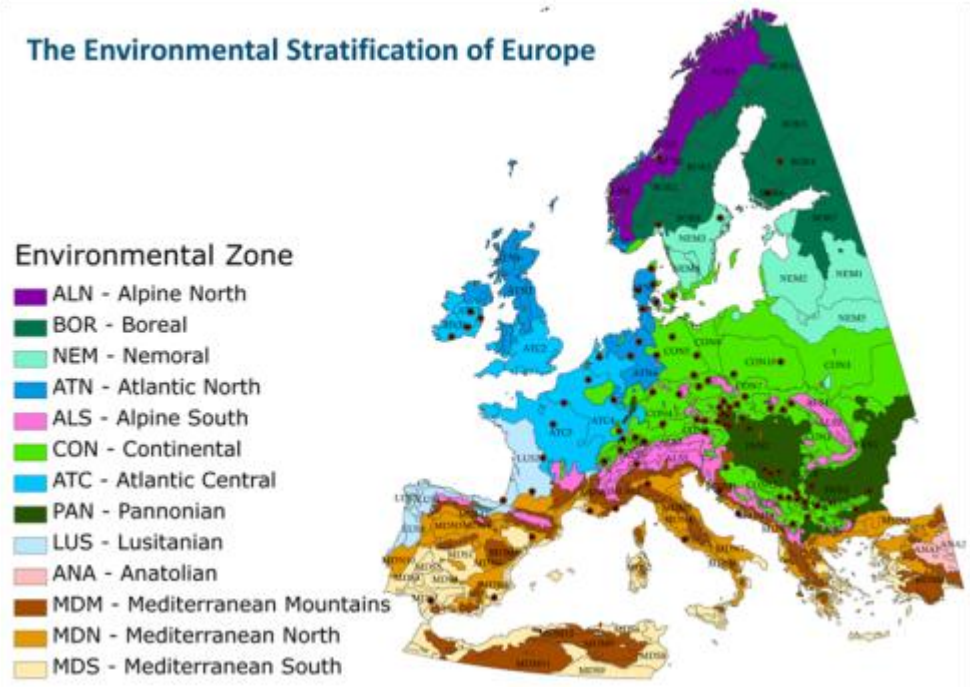


Figure 4.3: Location of the 83 sites (black dots) used for the development of the SnowCover meta-model laid over the environmental stratification of Europe of Metzger *et al.* (2005) and Jongman *et al.* (2006).

Two snow cover meta-models were developed - the first for days with more than 1 cm of fresh snow (i.e. 1 mm of snow water equivalent) and the second for days with more than 10 cm of fresh snow (i.e. 10 mm of snow water equivalent) which would allow leisure activities and provide frost protection for crops. The performance of both snow cover meta-models were evaluated using the explained variability (R^2), mean bias error (MBE) and root mean square error (RMSE) over the validation dataset (Figure 4.4). For days with more than 1 cm of fresh snow (i.e. 1 mm of snow water equivalent), the fit is good, with a MBE of close to 0, a RMSE of 2.1 days and more than 99% of the variability explained. The second meta-model

for days with more than 10 cm of fresh snow (i.e. 10 mm of snow water equivalent) shows similar accuracy (MBE = 0 day; RMSE = 2.6 days and R^2 = 0.99).

Additionally to the set of 83 sites on which the meta-model was developed and validated, data from the ECAD database was used to further test the meta-model performance on a set of 46 sites where all input parameters were available as well as snow cover information. These data included snow cover and weather predictors (i.e. daily maximum and minimum temperature, and precipitation values). Using the ECAD dataset the meta-models were able to explain over 88% of variability of all parameters with a MBE for number of days with snow cover of less than 5 days and RMSE of 16 days.

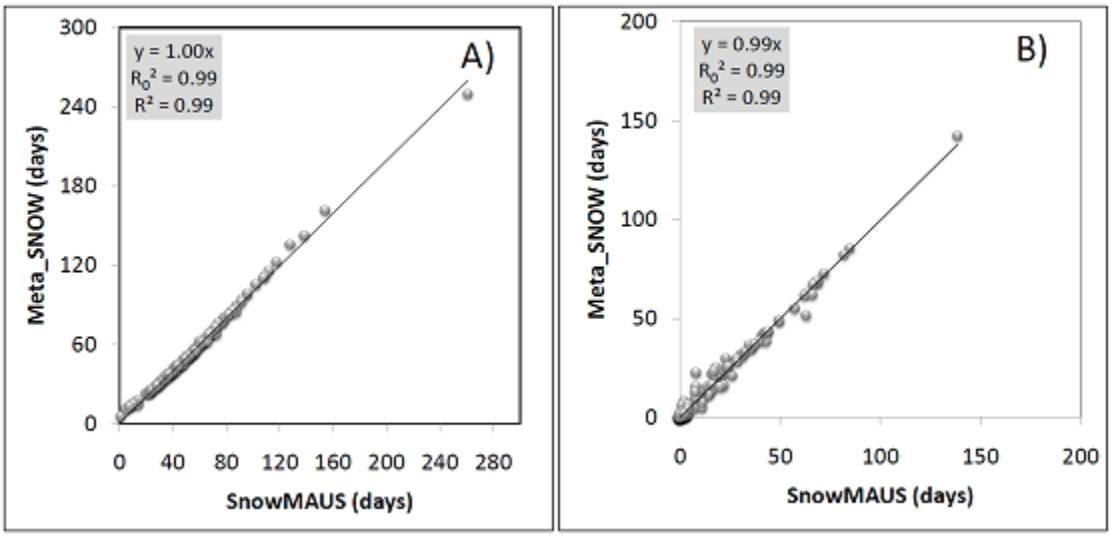


Figure 4.4. Comparison of the validation runs of the snow cover meta-models for snow days with more than (a) 1 cm of fresh snow and (b) 10 cm of fresh snow.

4.3 SnowCover meta-model illustrative results

4.3.1 *Baseline climate*

Once the meta-model was trained and validated, it was then applied across the CLIMSAVE10' European grid to produce a surface of mean snow cover days (Figure 4.5).

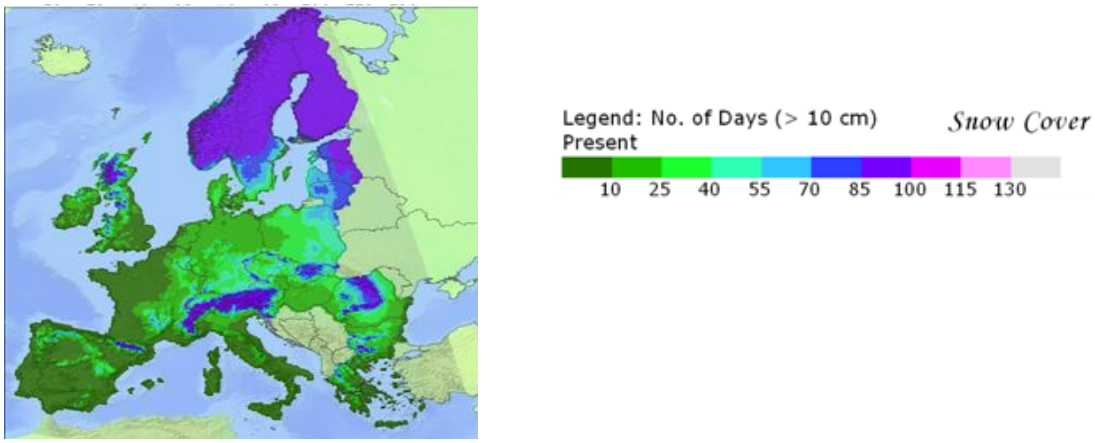


Figure 4.5: Illustrative results for mean number of days with more than 10cm of snow during the period 1961-1990.

4.3.2 *Climate sensitivity*

In order to test the newly developed meta-model routines, an extensive sensitivity analysis was carried out against changes in temperature (across the range from -2 to +6°C) and precipitation (from -40 to +40%). The results indicate that in terms of snow cover days, temperature is the main driving factor. Figure 4.6 illustrates the profound impacts of changes in temperature on the number of days with snow (without any change in precipitation), whilst Figure 4.7 shows the lesser effect of precipitation changes.

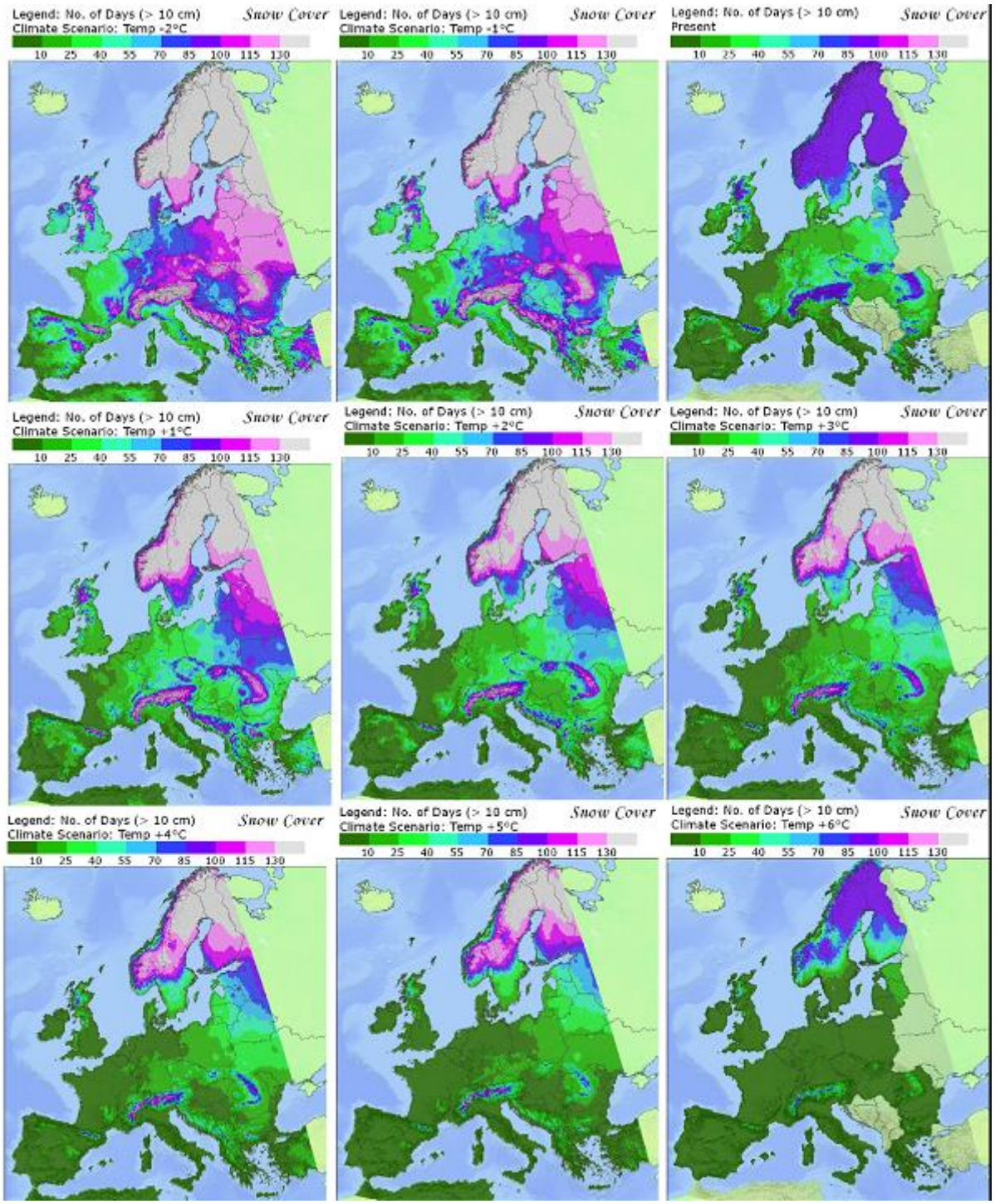


Figure 4.6: Sensitivity analysis of the snow cover meta-model (>10 cm of fresh snow) over the temperature range -2°C to +6°C.

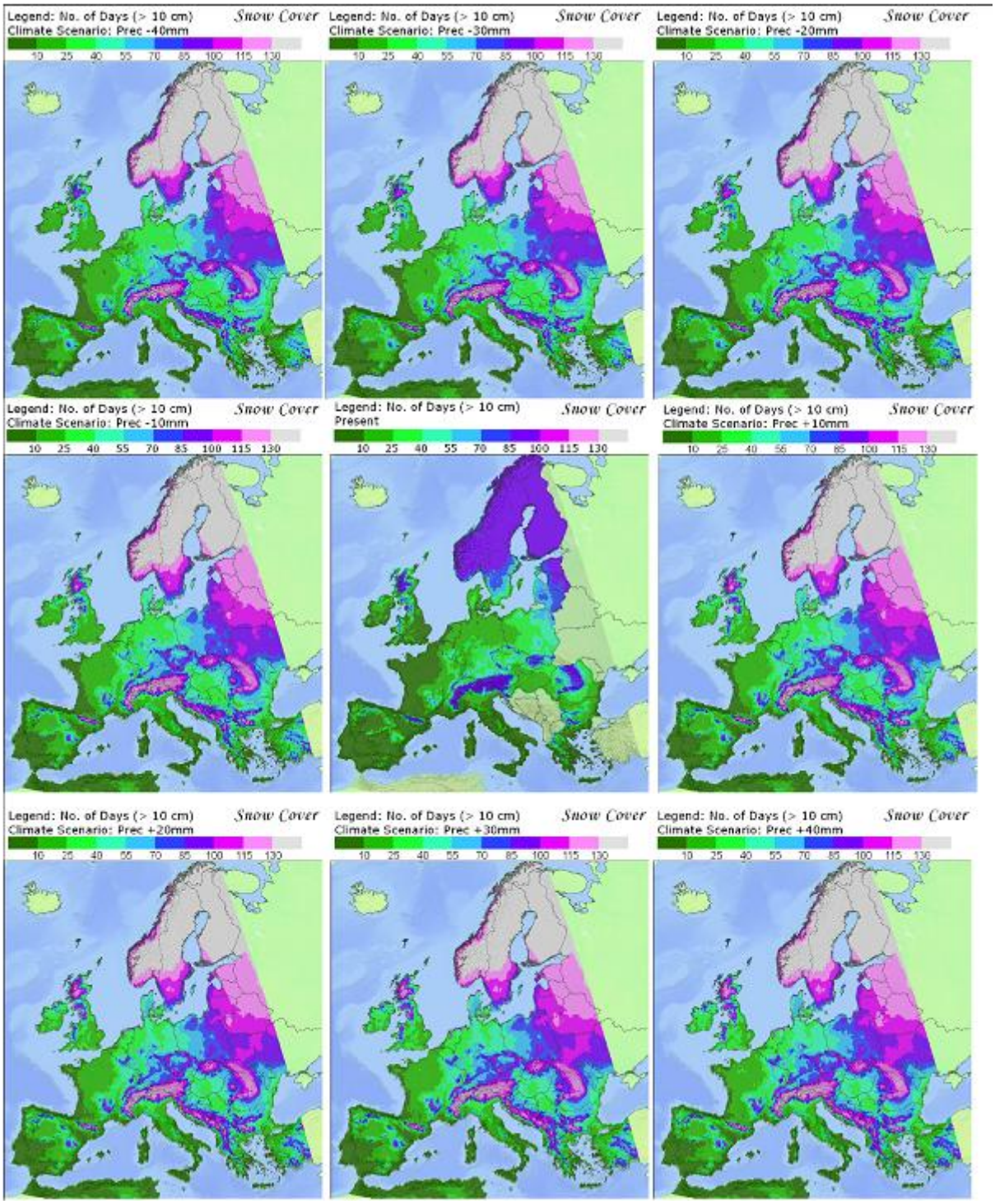


Figure 4.7: Sensitivity analysis of the snow cover meta-model (>10 cm of fresh snow) over the precipitation range -40 mm to + 40 mm per month.

4.4 Climate change scenarios

The meta-model performance was tested for the climate change scenarios available at the IAP and examples are presented in Figure 4.8. As one would expect (and as the SnowMAUS model runs for individual locations confirmed), there is a notable tendency towards a

decrease in the number of days with snow cover that might be considered as suitable for skiing (i.e. at least 10 cm of freshly fallen snow). Figure 4.8 shows there is tendency for a significant decline in the number of snow days that is most pronounced in the Alps and other mountain ranges in southern and central Europe and decreases towards the north and west. According to the GFCM21 scenarios, slight increases are to be expected in northern Sweden. Slight increases are predicted in parts of the UK, eastern France or Belgium, however the increase is negligible, being of the order of a few days, thus making the number of days with significant snow cover close to 10 days or less on average.

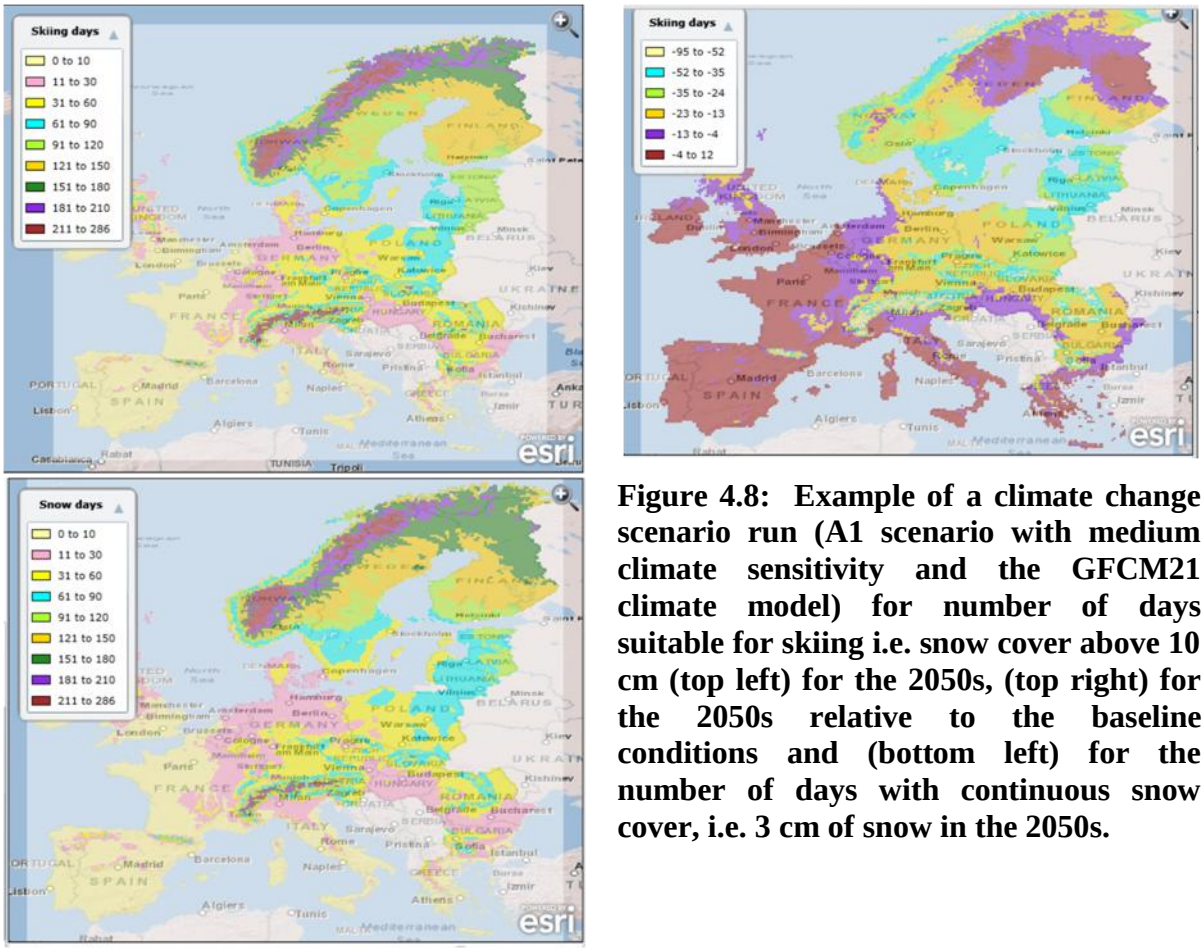


Figure 4.8: Example of a climate change scenario run (A1 scenario with medium climate sensitivity and the GFCM21 climate model) for number of days suitable for skiing i.e. snow cover above 10 cm (top left) for the 2050s, (top right) for the 2050s relative to the baseline conditions and (bottom left) for the number of days with continuous snow cover, i.e. 3 cm of snow in the 2050s.

While the baseline mean of number of snow days above 10 cm across the CLIMSAVE domain is 68 days, this will decrease to 52 days with large regional differences under the GFCM21 scenario. The median period experiences a greater decrease from 45 days under the baseline climate to 25 days in the 2050s. Even more rapid declines of snow cover are expected under some of the other GCM scenarios available in the IAP. As the number of snow days above 10 cm is driven by changes in the temperature and precipitation patterns, there is no effect of the socio-economic scenario (SES). However the particular SES will inevitably influence adaptation options and the coping capacity of individual regions that will be affected by the decline in number of days suitable for skiing or winter oriented tourism.

4.5 Integrating the SnowCover meta-model with the other sectoral meta-models

Currently, the present version of the SnowCover meta-model is considered as stand-alone, providing indicators for ecosystem services related to recreation/tourism (Table 3.1). Outputs may be used to "trim" the results of SFARMOD for particular crops or to define areas that could be used for winter tourism.

4.6 References

Coughlan, C.J. & Running, S.W. **(**1997). Regional ecosystem simulation: A general model for simulating snow accumulation and melt in mountainous terrain. Landscape Ecology, 12: 119–136.

Running S.W. & Coughlan J.C. (1988). A General Model of Forest Ecosystem Processes for Regional Applications: I. Hydrologic Balance, Canopy Gas Exchange and Primary Production Processes. Ecological Modelling, 42, 125-154.

Trnka, M., Kocmánková, E., Balek, J. *et al.* (2010). Simple snow cover model for agrometeorological applications, Agricultural and Forest Meteorology, 150: 1115-1127.

Trnka, M., Olesen, J., Kersebaum, C.K. *et al.* (2011). Agroclimatic conditions in Europe under climate change, Global Change Biology (in print).

5. Development and validation of the RUG urban meta-model

Sophie Rickebusch
Centre for Environmental Change and Sustainability, Edinburgh University, UK

5.1 RUG model description

The Regional Urban Growth (RUG; Rickebusch *et al.*, in prep.) model simulates urban growth as a function of changes in socio-economic variables (population, GDP per capita) and societal values (strictness of planning constraints, household location preferences). The model also takes into account local geography, travel times with the existing infrastructure and city typology (e.g. mono- versus polycentric).

The RUG meta-model in the IA platform consists of a look-up table of maps of the proportion of artificial surfaces per 10' x 10' grid cell. The appropriate map is selected according to the slider values set by the user for percentage change in population and GDP per capita, household preference for proximity to green space versus social amenities, attractiveness of the coast (scenic value versus flood risk) and strictness of the planning regulations to limit sprawl. The RUG meta-model then calculates the relative change in artificial surfaces compared to the baseline map derived from CORINE land-cover 2006 (CLC) and the area of residential and non-residential properties (which are in the same proportion as in the baseline map). The artificial surface maps were produced by running the original RUG model (on a 1 x 1 km grid) with all possible combinations of input values and aggregating the data to the 10' grid.

The original European-wide RUG model (Rickebusch, 2010; Rickebusch *et al.*, in prep.) runs on one NUTS 2 region at a time. It first calculates the expected quantity of artificial surfaces for the region, based on the linear regression model developed by Reginster & Rounsevell (2006), which links the proportion of artificial surfaces to the population and gross domestic product per capita. RUG uses two additional factors, urban type (large city versus smaller city/rural region) and country, in this regression model. RUG then evaluates the potential for settlement in each grid cell within the region, based on the cell's characteristics (e.g. existing artificial surfaces, distance to the coast) and the parameters entered by the user for planning and household preferences (e.g. strictness of planning constraints, attractiveness of the coast). Table 5.1 summarises the internal variables and those set by the user. The new percentage of artificial surfaces returned for each cell depends on its potential for settlement and on the total amount of artificial surfaces expected in the region.

The RUG model currently runs on a "growth-only" assumption, so it cannot simulate shrinkage. If the projected proportion of artificial surfaces is lower than the baseline value, it returns the latter.

Table 5.1: Input variables and parameters for the RUG (meta-) model.

Set by user in IA platform	Internal to model
Change in population	Population
Change in GDP per capita	GDP per capita
Household preferences for green space / social amenities	Current artificial surfaces
Strictness of planning constraints	Distance to coast
Attractiveness of coast	Remoteness from medium & large cities[a]
	Unsuitable areas (e.g. lakes, glaciers)

[a] Rickebusch *et al.* (in review)

5.2 Model calibration and validation

A calibration of all the input parameters was carried out in the previous version of RUG, which covered East Anglia and North-west England. This was done by running simulations using the baseline data. The parameter values were set, by trial and error, so as to minimise the difference between the simulated and observed maps, bearing in mind the significance of each parameter.

When the RUG model was expanded to 25 European countries, further calibration tests were carried out, particularly for variables such as the strictness of planning constraints, which is less likely to be transferable as different countries apply different planning regimes. Figure 5.1 shows an example of the difference between RUG results using baseline data and the observed proportion of artificial surfaces, for different values of the parameter representing strictness of planning constraints. The value of 0.2 used in the previous version of the model still gave the best results, although it led to slightly too high values (up to an average of +4%) in densely-urbanised grid cells. Increasing the parameter value to 0.3 or 0.5 increased the differences in densely-urbanised grid cells. On the other hand, decreasing the parameter value to 0.1 led to higher differences at the other end of the scale.

Figure 5.1 also gives an indication of how the model performs generally, given the parameter finally chosen (0.2, red boxes). The differences between the baseline simulation and the observed data are on average around 2-3%, with most values falling below 7%. There are also a few outliers with differences of over 30%. This is probably inevitable with a general model for Europe, as it cannot capture all the diversity within the simulation area.

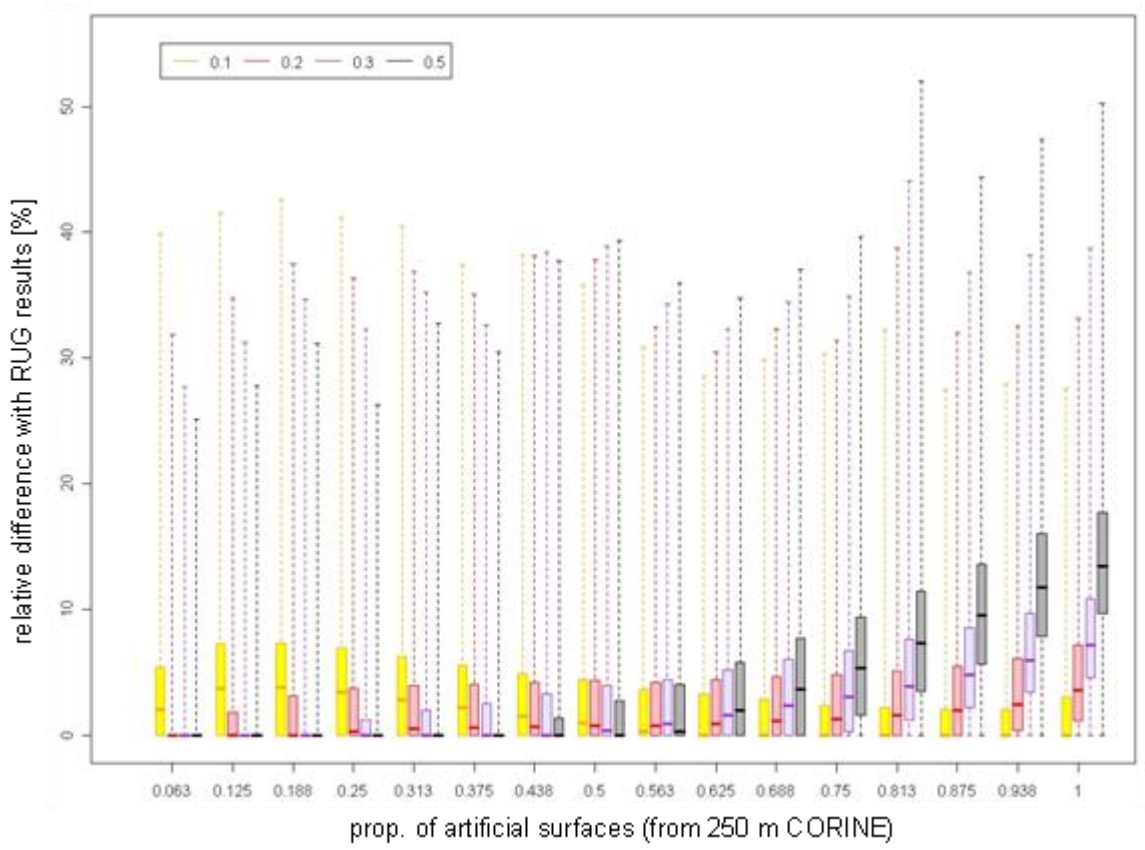


Figure 5.1: Comparison between RUG baseline simulations and observed CORINE data for four values of the "strictness of planning constraints" parameter. The "bars" (boxes) extend from the 1st to the 3rd quartile (with the median shown by the bold line) and the dashed lines are the whiskers which extend to the most extreme data point.

Figure 5.2 shows the proportion of artificial surfaces given by a RUG simulation with "baseline" parameters (no change in population or GDP per capita, household externalities preference = 2, planning constraints & attractiveness of coast = medium). The results are similar to the artificial surfaces found in the CLC map, though RUG tends to over-estimate the artificial surfaces, as shown in the map of the differences between the two (Figure 5.3, left). These differences are absolute values, which accounts for them being generally greater in heavily built-up grid cells. In relative terms, the differences tend to be larger in cells with low densities of artificial surfaces. For example, an absolute difference of 0.8 in a cell which contains 0.6% artificial surfaces according to CLC is equal to +133.3% relative difference. On the other hand, an absolute difference of 6.0 in a grid cell which is 55.0% built-up according to CLC is only +10.9% in relative terms. However, in both cases CLC and RUG show proportions of artificial surfaces of the same order of magnitude.

There are several causes for the differences between the CLC map and the RUG baseline simulation, aside from the fact that no model can ever represent reality exactly, but at best will show similar patterns. RUG is a growth-only model, i.e. it assumes that no artificial surfaces are removed, even if the population decreases for instance, which accounts for its tendency to over-estimate artificial surfaces. Negative differences are small and can be put down to differences in rounding and aggregation from the 1 km to the 10' grid. Using the same model parameters, e.g. for planning constraints, throughout Europe has the advantage of allowing the same model set-up to be applied to the whole study area, but the down side is that the baseline parameter values will be more suitable for some countries or regions than

others. Additionally, the regression function linking artificial surfaces with population, despite including factors for country and large city, only explains 72% of the variation, the rest being down to other factors, e.g. industrial development due to the presence of coal.

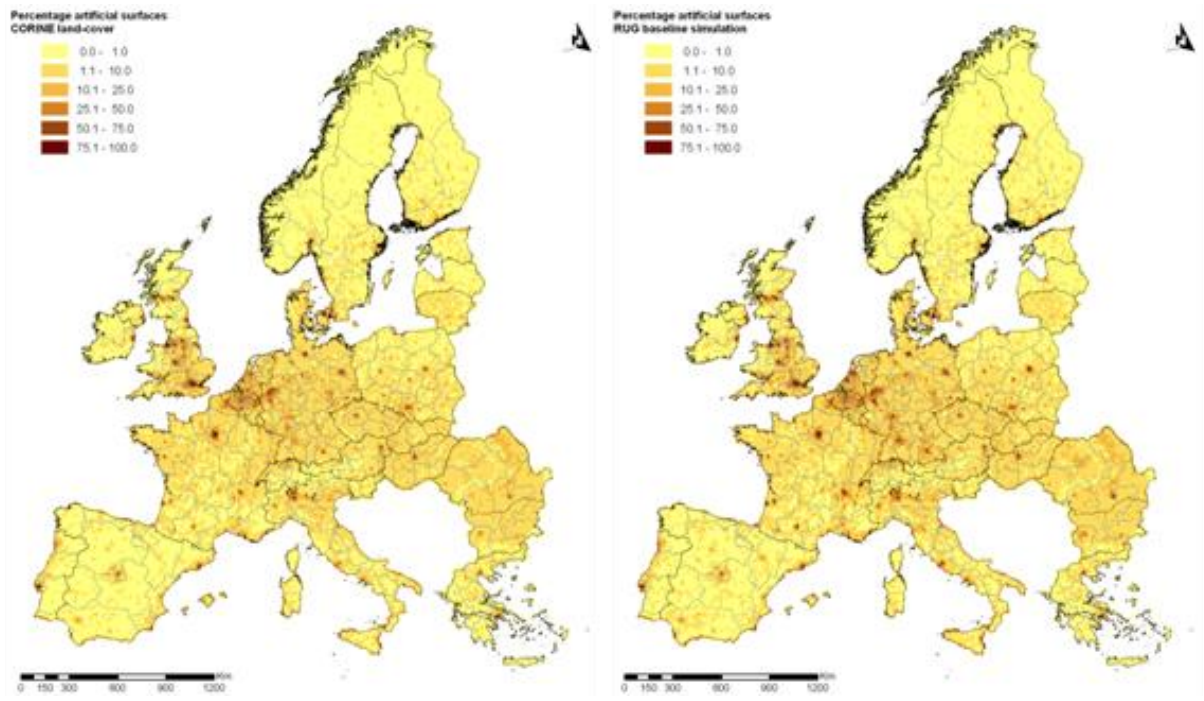


Figure 5.2: Artificial surfaces derived from the CORINE land-cover map (left) and produced by RUG with baseline parameters (right).

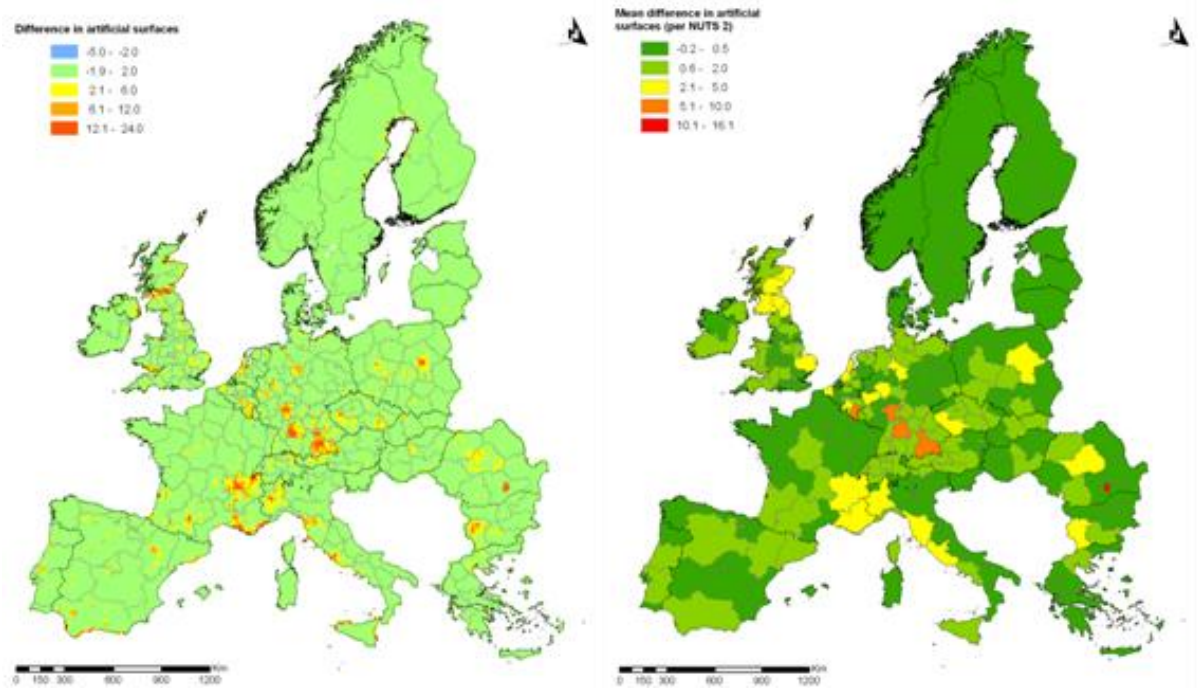


Figure 5.3: Difference in the percentage of artificial surfaces projected by a RUG simulation with baseline parameters and those in the CORINE land-cover map (left). The map on the right shows the same data averaged by NUTS 2 region.

Figure 5.3 (right) shows the mean per NUTS 2 region of the difference in artificial surfaces between the RUG simulation with baseline parameters and CLC. This gives an overview of the regions which are best represented in RUG and those in which the model does not perform as well.

5.3 RUG model outputs and integration with other meta-models

The main variable produced by the RUG model is the proportion of artificial surfaces per 10' x 10' grid cell (Figure 5.4), which has a range of 0 to 1 (0-100 %). It is used as a base to calculate other RUG output variables. It is also an input to the SFARMOD land-use model (Section 10).

From the above, RUG calculates the percentage difference in artificial surfaces relative to the baseline value (derived from CLC) for each cell (Figure 5.4). This is used by the WGMM model (Section 9) to calculate the changes in water flow due to surface sealing.

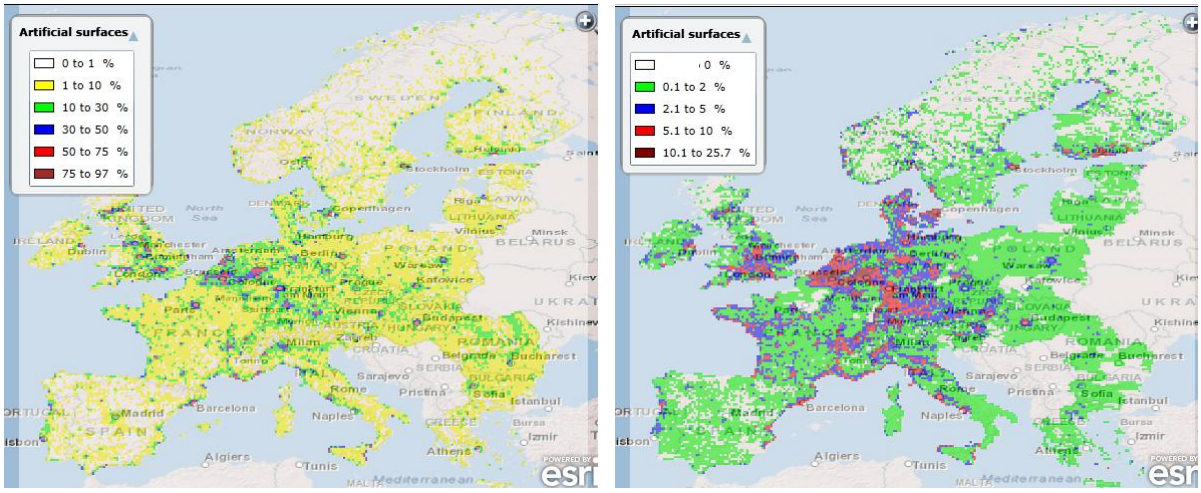


Figure 5.4: Example outputs from the IAP showing the proportion of artificial surfaces (left) and relative change in artificial surfaces (right) for the HadGEM climate scenario (A1 emissions and medium climate sensitivity) combined with the Riders on the Storm socio-economic scenario for the 2050s.

RUG also calculates the surface of residential (CLC category 1.1) and non-residential areas (CLC categories 1.2 - 1.4), in square kilometres, within each grid cell. This is based on the baseline proportions of residential versus non-residential areas in each cell. For example, if a cell has a baseline value of 1 km^2 artificial surfaces of which 75 % (0.75 km^2) are residential areas and RUG predicts the artificial surfaces will double, then there will be 1.5 km^2 of residential areas. These variables are passed to the CFFlood model (Section 7), to assess damage and risk to people.

Finally, RUG calculates the average percentage difference in artificial surfaces relative to baseline value across all cells. This aggregated indicator is displayed on the IA platform, to give the user a quick indication of the general effect of the settings they have chosen.

5.4 References

Reginster I. & Rounsevell M. (2006). Scenarios of future urban land use in Europe. Environment and Planning B-Planning & Design, 33: 619-636.

Rickebusch S. (2010). Maps of land-use change scenario projections for Europe - PLUREL project deliverable 1.4.3, http://www.plurel.net/Reports-76.aspx

Rickebusch S., Fontaine C.M. & Rounsevell M.D.A. (in prep.) Scenario analysis of urban growth in Europe. Computers, Environment and Urban Systems.

Rickebusch S., Helminen V., Ristimäki M., Kontio P., Fontaine C. M. & Rounsevell M. (in review). Commuting patterns and travel time costs: exploring distance to cities for large-extent land-use change scenario modelling. Journal of Transport Geography.

6. Development and validation of the metaGOTILWA+ forest meta-model

Joan Maspons and Santi Sabaté
CREAF, Autonomous University of Barcelona, Spain

6.1 Introduction

MetaGOTILWA+ is used in the IA Platform to simulate the impacts of climate change on forest ecosystems services such as wood production, carbon balance, etc (Table 3.1), and how forest management might play a role to mitigate such impacts on the main forest species that occur over Europe.

MetaGOTILWA+ is based on the GOTILWA+ model. The full GOTILWA+ model requires a lot of computational time to simulate each forest type, in each location (pixel) in Europe under different climates and management regimes. Since the IA platform requires a fast runtime, a new meta-model version has been developed to provide responses in a few seconds. Neural networks have been used to reproduce GOTILWA+ outputs as a function of GOTILWA+ inputs.

6.2 GOTILWA+ model description

The GOTILWA+ model (Growth Of Trees Is Limited by WAter, http://www.creaf.uab.cat/gotilwa+/) simulates carbon and water uptake and fluxes through forests of different tree species and in changing environmental conditions, due to either climate or management regimes. The input data include: climate (maximum and minimum temperature, precipitation, vapour pressure deficit, wind speed and global radiation); stand characteristics (tree structure and diameter at breast height (DBH) class distribution); tree physiology (photosynthetic and stomatal conductance parameters); and site conditions including soil characteristics and hydrological parameters. The processes are described with different sub-models that interact and integrate the results of simulated growth and evolution of the whole tree stand through time (hourly calculations integrated at a daily time step).

The light extinction coefficient is estimated by Campbell's approach (1986), based on an ellipsoidal leaf angle distribution. The photosynthesis equations are based on Farquhar and co-workers approach (Farquhar & Von Caemmerer, 1982). Stomatal conductance uses Leuning's approach that modifies the Ball, Woodrow and Berry model (Leuning, 1995). Leaf temperature is determined based on the leaf energy balance (Gates, 1962; 1980) and transpiration is estimated according to the Penman-Monteith equation (Monteith, 1965, Jarvis & Mcnaughton, 1986). Autotrophic respiration is separated into maintenance and growth respiration. Maintenance respiration is calculated as a proportion of total respiring biomass (structural and non-structural components distinguished), with rates that depend on temperature according to a Q10 approach. Growth respiration is a fraction of available carbohydrates for growth consumed when transformed into new tissues. A constant efficiency of 0.68 is assumed (g of new tissue / g of carbohydrate). Net primary production (NPP) is allocated first to form new leaves and fine roots to compensate for their turnover. The remaining NPP is allocated to the pool of mobile carbon in leaves and woody tissues. The surplus is invested in new tissues (leaves, fine roots and sapwood) according to the pipe model (Shinozaki *et al.*, 1964). Soil is divided into two layers, organic and inorganic

horizons. Soil organic matter (OM) is originated by plant litter: leaves, branches, stems and reproductive organs aboveground and coarse and fine roots belowground. OM is decomposed depending on soil temperature (according to a Q10 approach) and soil moisture (optimal at 60% of the maximum soil water-filled porosity). Soil moisture is calculated based on water inputs and outputs and soil traits. Temperature also affects leaf shedding through a Q10 approach. Root mortality is also dependent on temperature (Q10 approach), soil moisture and the length of the growing period.

6.3 GOTILWA+ validation and application

The GOTILWA+ model has been extensively applied in different European projects such as LTEEF-II, ATEAM, SILVISTRAT and ALARM. To check that the model provides realistic results, it has been tested against empirical data from the Forest National Inventories as well as compared with other process based models (see Kramer et al 2002, Morales et al 2005, Keenan et al 2009a). Within the previous projects, GOTILWA+ has been applied Europe-wide (see Schröter et al 2005; Keenan et al 2009b,c; Keenan et al 2010).

6.4 Development of the metaGOTILWA+ meta-model

Artificial neural networks (ANNs) have been developed to emulate the performance of the GOTILWA+ model but provide results in a few seconds. In order to train the ANN, around 900 cells were selected across Europe to explore the response of GOTILWA+ across all ranges of environmental conditions (Figure 6.1). These cells were selected to ensure the representivity of climatic conditions and to include more extreme conditions by selecting cells with higher and lower values for each input variable (Table 6.1). Simulations were run from 1950 until 2100 using climatic data from the HadCM3 global climate model for the A1B emissions scenario. CLIMSAVE is only simulating impacts until the 2050s. However, including a greater range of projections ensures that extrapolation is avoided because the climatic conditions of the 2050s will be well captured within the GOTILWA+ simulations.

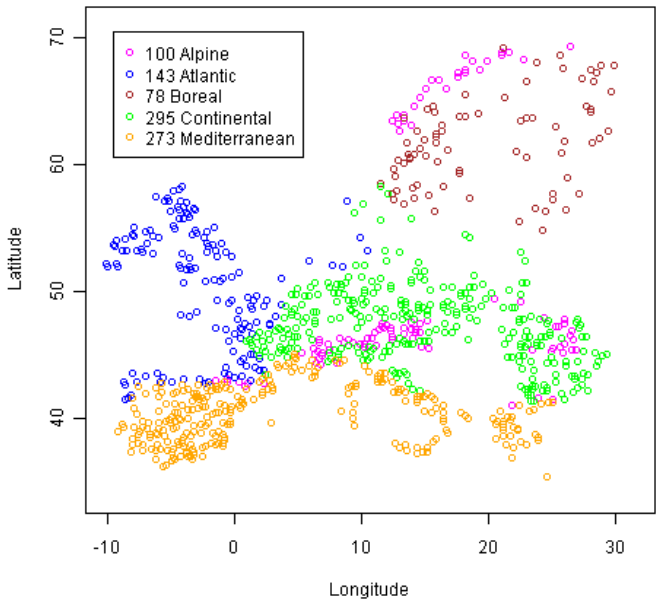


Figure 6.1: Sample cells used to train the Artificial Neural Networks. Colors indicate the region to which the cell belongs.

Table 6.1: Input variables for the metaGOTILWA+ meta-model.

Variable	Definition
Temperature	Monthly mean temperature
Precipitation	Monthly mean precipitation
Effective soil volume	The product of the mean soil depth and the proportion of stones in the soil
CO_2	Atmospheric CO_2 concentration
Forest management	Forest management regime (no management, even aged management or uneven aged management)
Tree species	Dominant tree species in the forest

For each cell simulations were conducted for all characteristic species from the region, all management regimes and with four different levels of effective soil volume to produce the variables listed in Table 6.2.

Table 6.2: Output variables simulated by the metaGOTILWA+ meta-model.

Variable	Definition	Ecosystem Service indicator
Wood yield	Wood yield in managed forests	Wood production
Net Ecosystem Exchange	Carbon balance of the ecosystem	Carbon balance
Net Primary Production	Carbon balance of the primary producers	Forest physiological viability
Gross Primary Production	Total amount of carbon fixed by the trees	Carbon balance
Biomass stock	Sum of soil organic matter, aboveground biomass and below ground biomass	Carbon stock
Water stored in soil	Amount of water stored in soil	Water stress indicator
Length of the growth period	Length of the growth period determined by temperature and water availability	

Fast Artificial Neural Networks library (http://leenissen.dk/fann) has been used to build and run the neural networks. An evolving topology training algorithm (Cascade2) was used which dynamically builds and trains the ANN.

6.5 Meta-GOTILWA+ validation, illustrative application and sensitivity analysis

The predictions of the ANN were tested against data from cells which have not been used for training. Although there is inevitable scatter in the example results for *Pinus sylvestris* (Figure 6.2), there is a strong 1:1 relationship between the outputs of metaGOTILWA+ and GOTILWA+. As an illustrative application of the model, Figure 6.3 shows an example of spatial results across the selected climate zones across Europe in which *Pinus sylvestris* grows (continental, boreal and alpine) for the baseline climate applying even aged management.

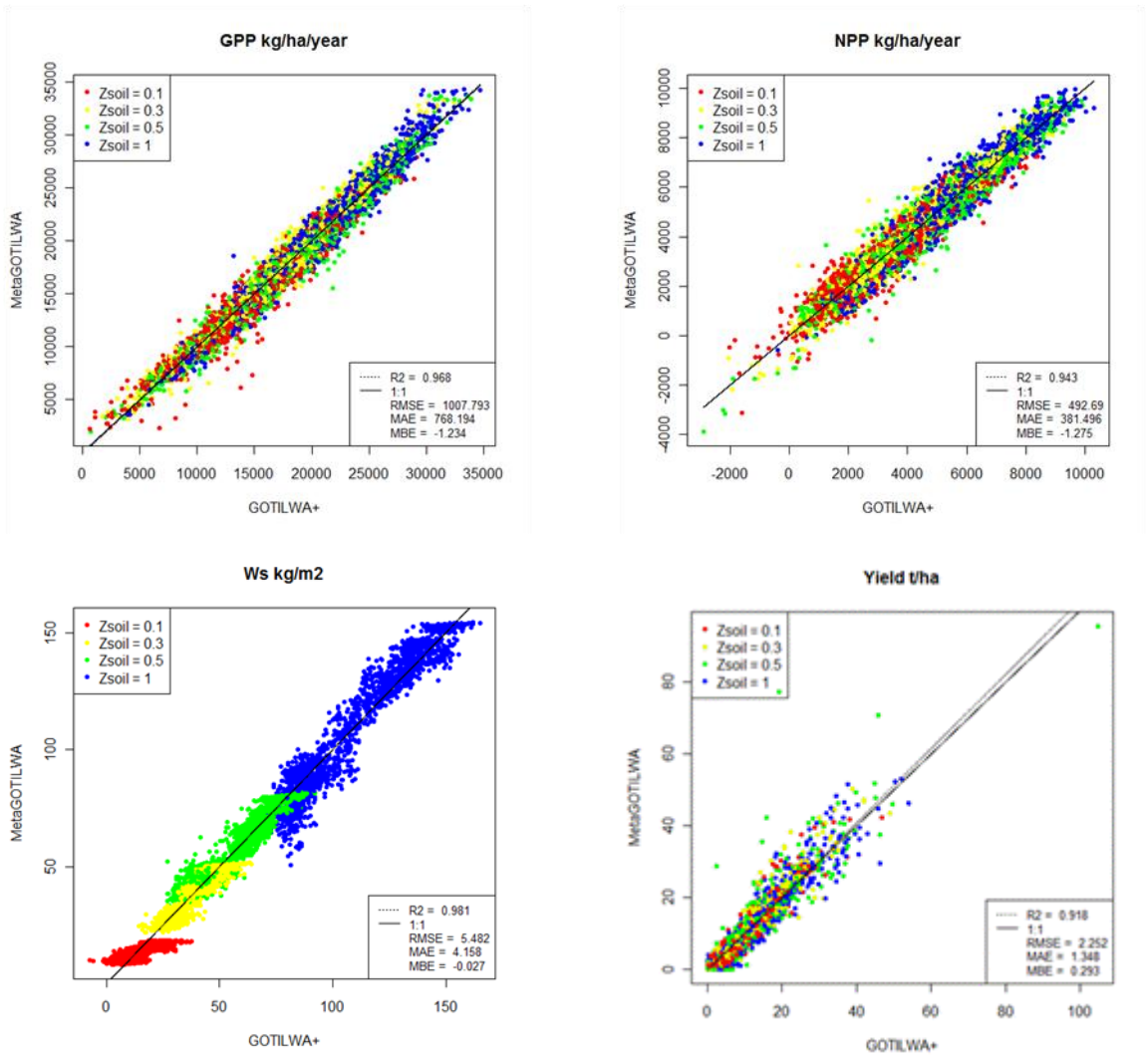
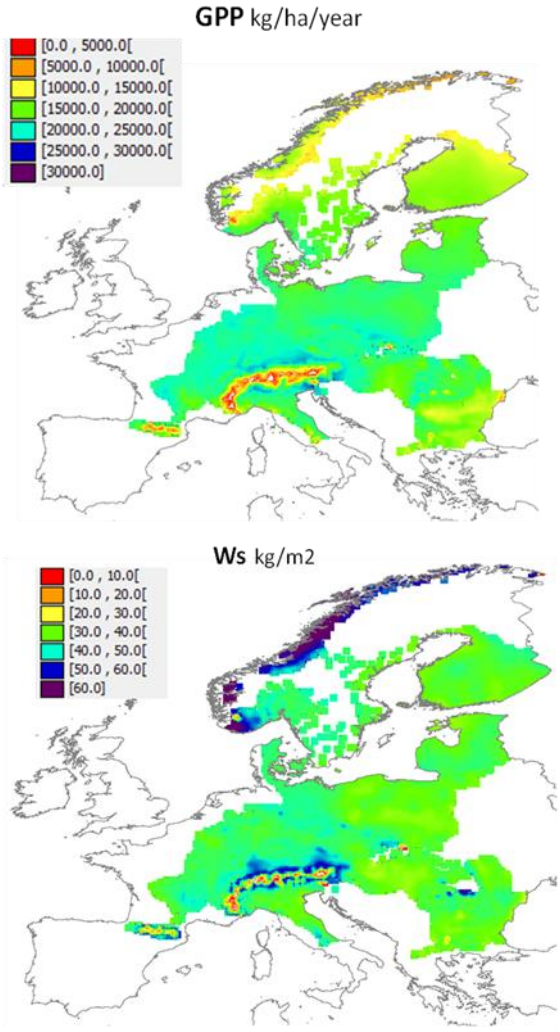


Figure 6.2: Comparison of outputs from GOTILWA+ and metaGOTILWA+ for *Pinus sylvestris* under evenaged management with different effective soil depths [GPP - Gross Primary Production; NPP - Net Primary Production; Ws - Water in soil and Yield].

A sensitivity analysis of metaGOTILWA+ to the main climatic drivers for forest growth, which are precipitation and temperature, has been carried out. From the baseline climate, changes in precipitation (-30%, -15%, 0, +15% and +30%), temperature (-30%, 0, +10%, +20% and +30%), and their interaction were applied. Results for the output variable NPP (kg/ha/year) for the baseline (Figure 6.4) and the more extreme changes (Figure 6.5) are shown. Both precipitation and temperature impact greatly on forest productivity even though precipitation is more determining. Increased temperature together with decreased precipitation has the strongest effect, reducing forest productivity all across Europe. Increased temperature together with increased precipitation has positive effects on productivity in very specific areas but it implies a general slight productivity reduction. Decreased temperature together with increased precipitation increases forest productivity quite uniformly even though the positive effect diminishes in areas which are already very productive. Decreased temperature and precipitation leads to varying affects, increasing productivity in some areas but decreasing it in others. These results show that MetaGOTILWA is sensitive to the tested climate variables.



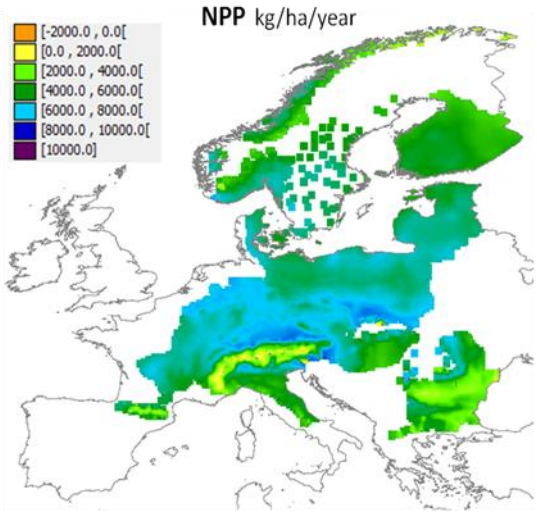


Figure 6.3: Outputs from metaGOTILWA+ for *Pinus sylvestris* for the boreal, continental and alpine regions using the baseline climate, even aged management and an effective soil depth of 0.3m (without stones). [GPP - Gross Primary Production; NPP - Net Primary Production; Ws - Water in soil].

.

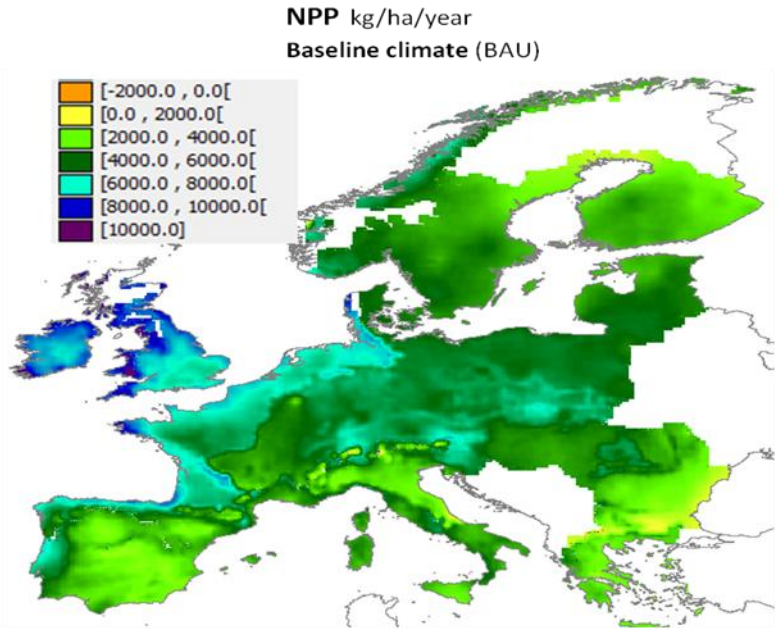


Figure 6.4: Results for metaGOTILWA+ for the baseline climate, self thinning management and an effective soil depth of 0.3m (without stones). [NPP - Net Primary Production].

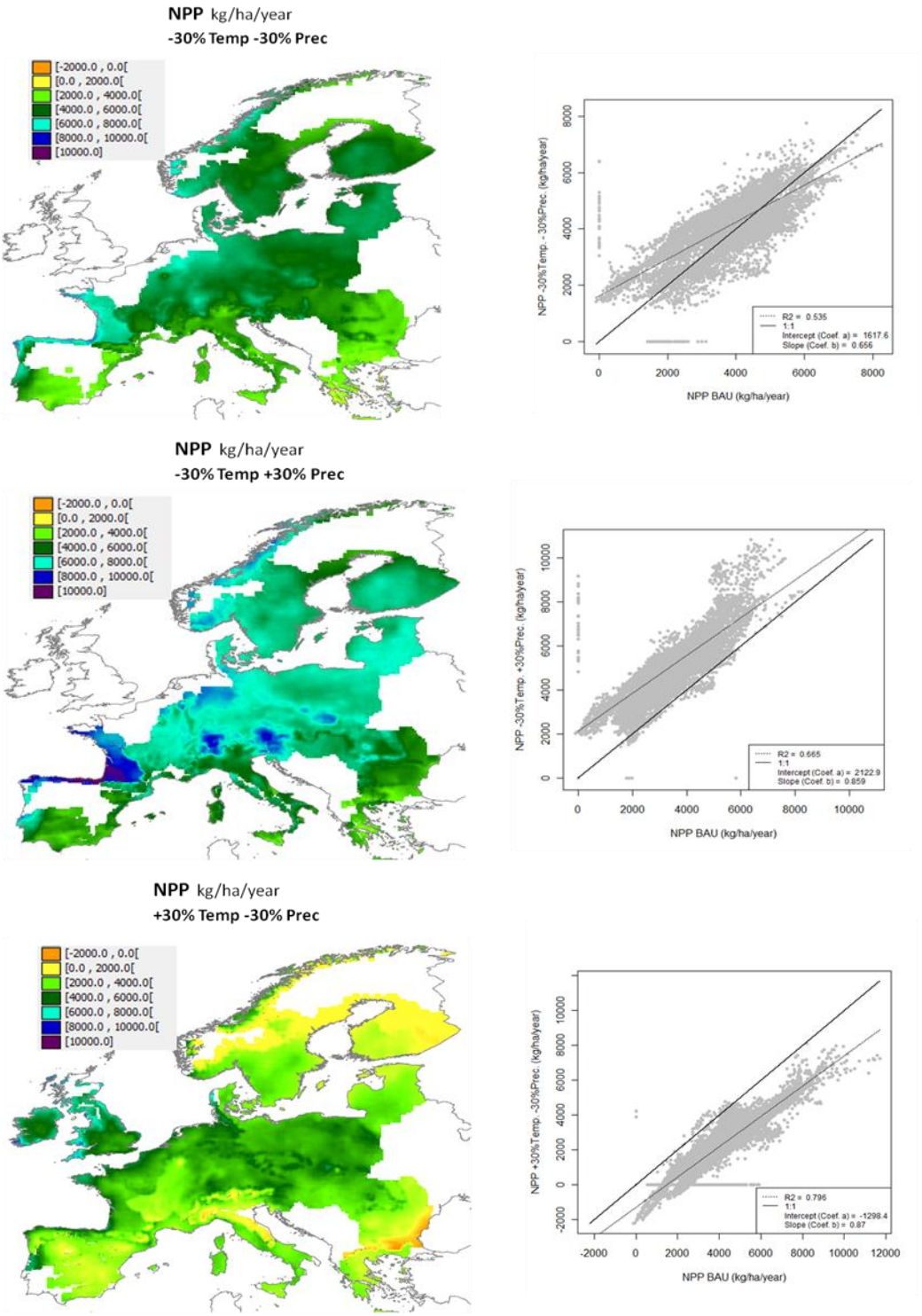


Figure 6.5: Sensitivity analysis for metaGOTILWA+ with self thinning management and an effective soil depth of 0.3m (without stones). [NPP - Net Primary Production]. The most extreme sensitivity tests are shown [-30%Temp-30%Prec; -30%Temp+30%Prec; +30%Temp-30%Prec; +30%Temp+30%Prec].

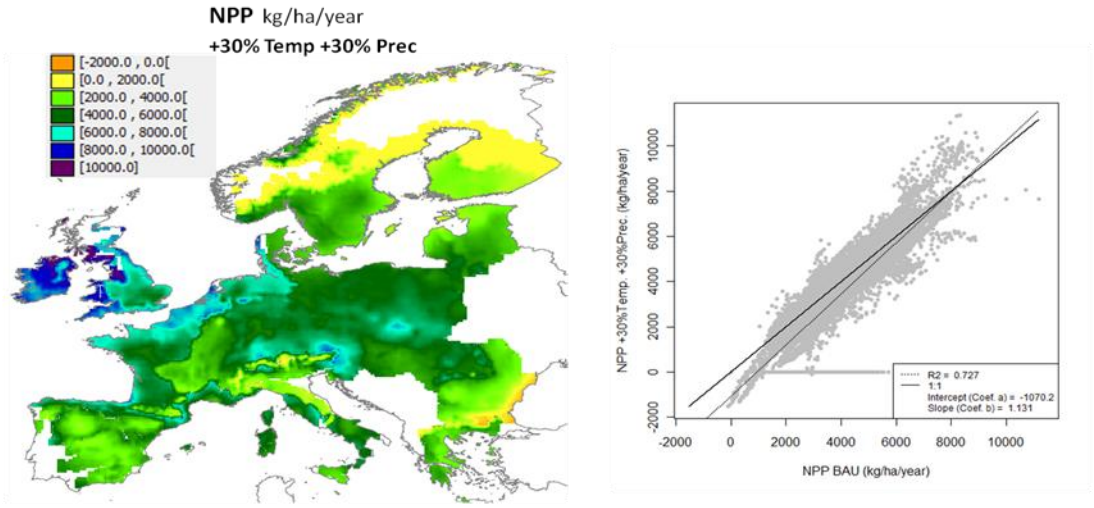


Figure 6.5 cont.: Sensitivity analysis for metaGOTILWA+ with self thinning management and an effective soil depth of 0.3m (without stones). [NPP - Net Primary Production]. The most extreme sensitivity tests are shown [-30%Temp-30%Prec; -30%Temp+30%Prec; +30%Temp-30%Prec; +30%Temp+30%Prec].

6.6 Integration of metaGOTILWA+ with other sectoral meta-models

MetaGOTILWA+ outputs are being used to assess the effects of climate change on European forests and the ecosystem services provided by them. Some outputs such as wood yield are passed according to the climatic conditions, soil depth, management and dominant tree species to the SFARMOD meta-model to include inputs from the forestry sector to optimise land use (see Section 10).

6.7 References

Campbell, G.S. (1986). Extinction coefficients for radiation in plan canopies calculated using an ellipsoidal inclination angle distribution. Agric. For. Meteorol. 36, 317–321.

Farquhar, G.D. & Von Caemmerer, S. (1982). Modeling of photosynthetic response to environment. In: Lange, O.L., Nobel, P.S., Osmond, C.B., Ziegler, H. (Eds.), Encyclopedia of Plant Physiology: Physiological Plant Ecology II, Water Relations and Carbon Assimilation, Vol. 12B. Springer, Berlin, pp. 549– 587.

Gates, D.M. (1962). Leaf temperature and energy exchange. Theoretical and Applied Climatology 12-2, 321-336.

Gates D.M. (1980). Biophysical Ecology. Springer, New York.

Jarvis, P.G. & McNaughton, K.G. (1986). Stomatal control of transpiration: scaling up from leaf to region. Adv. Ecol. Res. 15, 1–49.

Keenan, T., Sabaté, S. & Gracia, C. (2010). Soil water stress and coupled photosynthesis–conductance models: bridging the gap between conflicting reports on the relative roles of stomatal, mesophyll conductance and biochemical limitations to photosynthesis. Agricultural and Forest Meteorology 150, 443-453.

Keenan, T., Garcia, R., Friend, A.D., Zaehle, S., Gracia, C. & Sabate, S. (2009a). Improved understanding of drought controls on seasonal variation in Mediterranean forest canopy CO2 and water fluxes through combined in situ measurements and ecosystem

modelling. Biogeosciences, 6, 1423-1444.
Keenan, T., Niinemets, Ü., Sabate, S., Gracia, C. & Peñuelas, J. (2009b). Process based inventory of isoprenoid emissions: current knowledge, future prospects and uncertainties. Atmospheric Chemistry and Physics, 9, 4053–4076.
Keenan, T., Niinemets, Ü., Sabate, S., Gracia, C. & Penuelas, J. (2009c). Seasonality of monoterpene emission potentials in Quercus ilex and Pinus pinea: implications for regional VOC emissions modeling. Journal of Geophysical Research, 114, D22202,doi: 10.1029/2009JD011904.
Kramer, K., Leinonen, I., Bartelink, H.H., Berbigier, P., Borghetti, M., Bernhofer, C., Cienciala, E., Dolman, AJ, Froer, O. & Gracia, CA and others (2002). Evaluation of six process-based forest growth models using eddy-covariance measurements of CO_2 and H2O fluxes at six forest sites in Europe. Global Change Biology, 3-8, 213-230.
Leuning, R. (1995). A critical appraisal of a combined stomatal–photosynthesis model for C3 plants. Plant Cell Environ. 18, 339-355.
Monteith, J.L. (1965). Evaporation and environment. Symp. Soc. Exp. Biol. 19, 205–234.
Morales, P., Sykes, M.T., Prentice, I.C., Smith, P., Smith, B., Bugmann, H., Zierl, B., Friedlingstein, P., Viovy, N., Sabaté, S. and others (2005). Comparing and evaluating process-based ecosystem model predictions of carbon and water fluxes in major European forest biomes. Global Change Biology, 11-12, 2211-2233.
Schröter D., W. Cramer, R. Leemans, I.C. Prentice, M.B. Araújo, N.W. Arnell, A. Bondeau, H. Bugmann, T.R. Carter, C.A. Gracia, A.C. de la Vega-Leinert, M. Erhard, F. Ewert, M. Glendining, J.I. House, S. Kankaanpää, R.J.T. Klein, S. Lavorel, M. Lindner, M.J. Metzger, J. Meyer, T.D. Mitchell, I. Reginster, M. Rounsevell, S. Sabaté, S. Sitch, B. Smith, J. Smith, P. Smith, M.T. Sykes, K. Thonicke, W. Thuiller, G. Tuck, S. Zaehle, & B. Zierl (2005). Ecosystem Service Supply and Human Vulnerability to Global Change in Europe. Science, 310 (5732): 1333-1337.
Shinozaki, K., Yoda, K., Hozumi, K. & Kira, T. (1964). A quantitative analysis of plant form – the pipe model theory. I. Basic analyses. Japanese Journal of Ecology, 14: 97–105.

7. Development and validation of the fluvial and coastal flood (CFFlood) meta-model

Mustafa Mokrech, Abiy S Kebede and Robert J Nicholls
School of Civil Engineering and the Environment, University of Southampton, UK

7.1 Introduction

The **C**oastal **F**luvial **Flood** (**CFFlood**) meta-model provides estimates of the socio-economic and environmental (i.e., floodplain habitat) impacts of future flooding that are attributed to climate change and sea-level rise in Europe's coastal and fluvial floodplains[1]. It also accounts for future socio-economic changes (e.g., change in population and GDP) by investigating human pressures under a range of socio-economic scenarios as well as at user-defined options for exploratory purposes. The modelling is conducted at multiple scales and aggregated to the 10' spatial grid for the IAP. The baseline datasets are mostly resampled from higher spatial resolution datasets (i.e., 100 m resolution CORINE land use data and 100 m flood maps). The meta-model allows the exploration of a range of plausible adaptation options that are designed to reduce flood risks and/or to minimise losses of key floodplain habitats. The impact assessment methodologies and adaptation options are explained in the following sections.

7.2 CFFlood model description

A conceptual framework of the CFFlood meta-model has been developed to explain the variables and the main steps for implementing the meta-model. The framework consists of three main sub-model components: (1) Coastal flood, (2) Fluvial flood and (3) Habitat change/loss components. These components are coupled and are also integrated to a range of plausible adaptation measures that allow the analysis of plausible responses to climate change and sea-level rise.

7.2.1 *Coastal flood sub-model component*

The framework of the coastal flood component (Figure 7.1) illustrates the main steps implemented for assessing the impacts of coastal flooding. The method uses the estimated Standard of Protection (SoP) parameter for analysing the change in flood risk due to the effect of relative sea-level rise on extreme sea levels. It assumes that SoP decreases and flood frequency increases with a rise of extreme sea level (e.g. Figure 7.2) (Mokrech *et al.*, 2008): baseline extreme sea levels are produced by a combination of astronomical tides and meteorologically-induced storm surges, and future sea levels are increased by sea-level rise.

The flood risk zones are identified by analysing the topography against the regional extreme sea levels, based on present-day extreme sea levels and relative sea-level rise scenarios, as appropriate. Consequently the area at risk of flooding is calculated and an estimate of the people living in the flood risk zones is calculated using population density. A comparison

[1] Note that intra-urban flooding (Evans *et al.*, 2004a; 2004b) which operates at a smaller scale and via different mechanisms (e.g., more intense precipitation and overwhelmed drains) is not considered by the CFFlood meta-model.

between the extreme water levels and the estimated SoP determines the actual extent of flooding within these flood risk zones. Hence, the number of people who experience flooding is determined based on the population within the flooded areas. The flood damages for residential properties (both contents and structure) are also calculated based on urban areas and people at risk of flooding, flood water depths, and Gross Domestic Product (GDP), following the damage curves provided by Linham *et al.* (2010). The changes in urban areas are derived from the RUG model (Section 5).

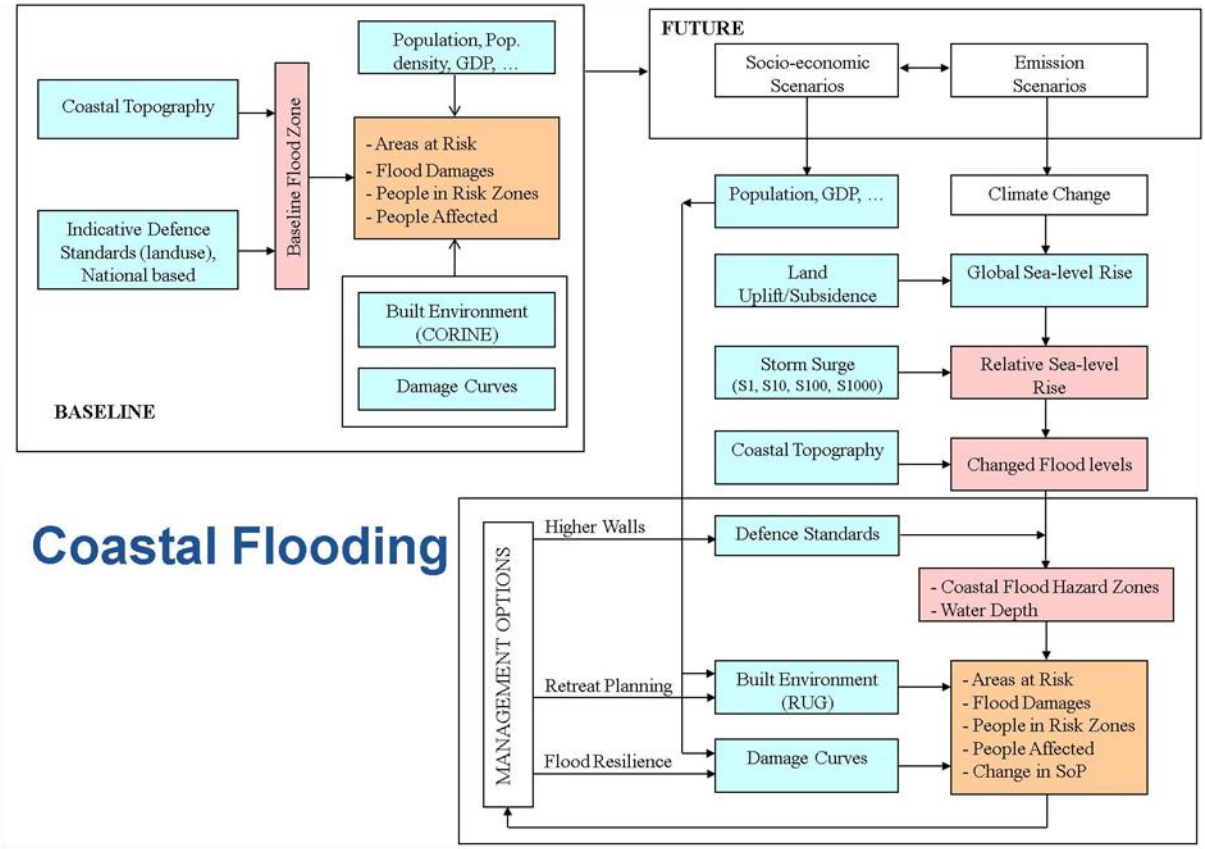


Figure 7.1: Coastal flooding component in the CFFlood meta-model.

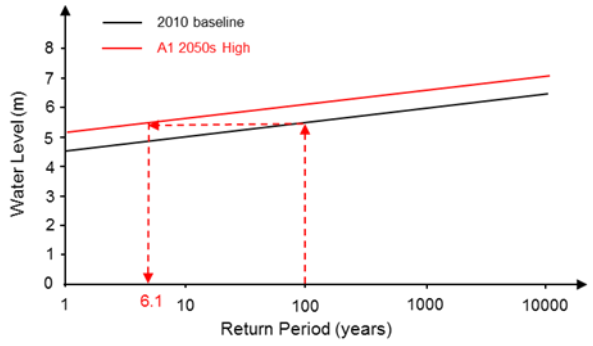


Figure 7.2: Illustrative example of the change in SoP due to sea-level rise. The 1 in 100 year SoP is degraded to a 1 in 6.1 year.

7.2.2 *Fluvial flood sub-model component*

The fluvial flood component follows a similar approach to the coastal flood component (see Figure 7.3). It uses the European fluvial flood maps produced by the JRC Institute using LISFLOOD simulations at 100 m resolution (Feyen *et al.*, 2011). These simulations provide flood maps for fluvial catchments (both extent and water depth) with return periods of 2, 5, 10, 20, 50, 100, 250 and 500 years, assuming no flood defences. These maps have been used as indicative maps of the flood risk zones in the CLIMSAVE project. The fluvial flood model estimates the land area and number of people living in fluvial flood hazard zones, and people affected and economic damages due to fluvial flooding. The flood maps are analysed in conjunction with the CORINE land use data and the results are gridded at the 10' resolution. The estimated Standard of Protection (SoP) parameter is used to analyse the change in flood risk due to changing peak flow (e.g. Figure 4) (Mokrech *et al.*, 2008). The changes in the peak river flow are derived from the WaterGAP model (Section 8).

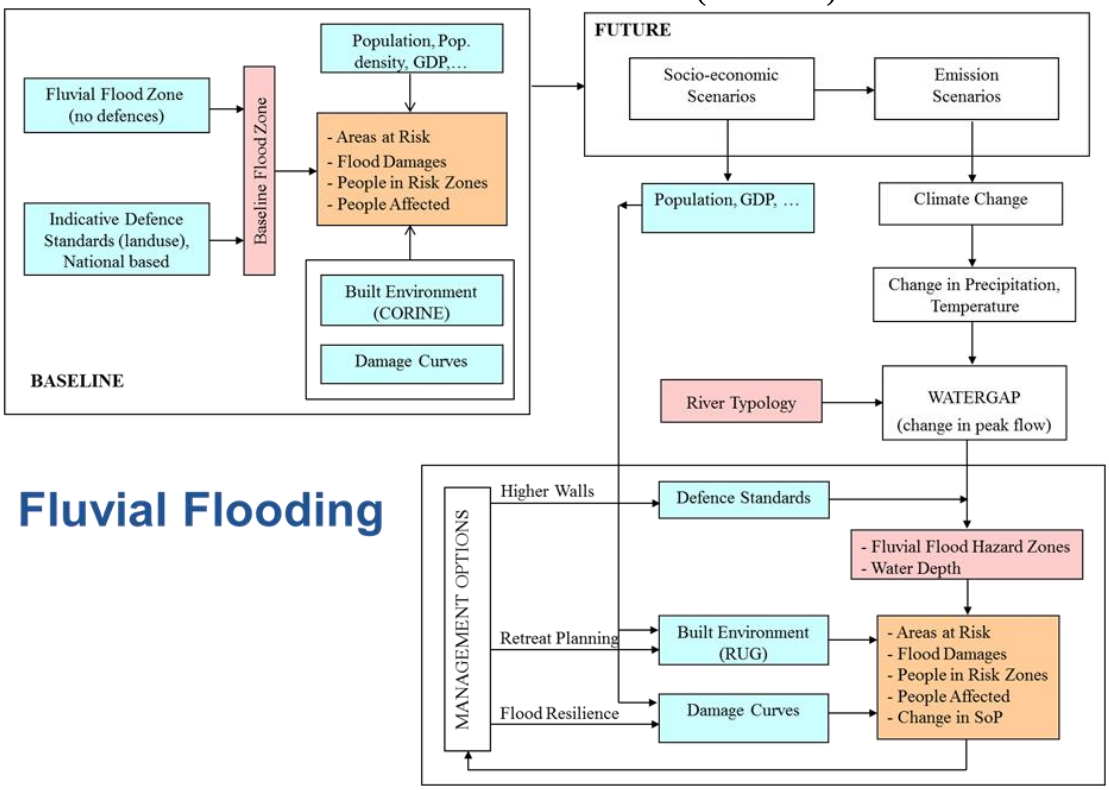


Figure 7.3: Fluvial flooding component in the CFFlood meta-model.

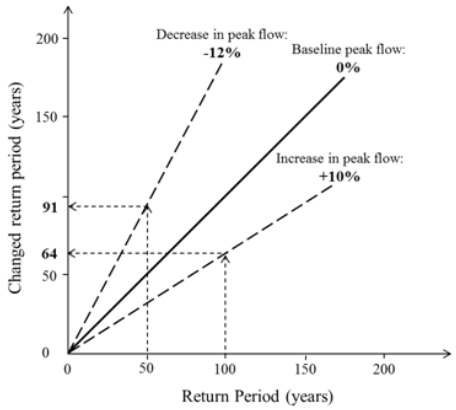


Figure 7.4: Illustrative example of how the fluvial flood model works: the effect of a change in peak flow on the standard of protection (SoP).

7.2.3 *Wetland change/loss sub-model component*

In addition to damages on people and property and flood constraints on agricultural production, the CFFlood model also assesses possible changes in the area of floodplain habitats comprising 'saltmarsh', 'intertidal flats' and what we term here 'coastal grazing marsh' in coastal floodplains, and 'inland marshes' in fluvial floodplains. Changes to these areas are of interest under the Habitats Directive. Saltmarsh and intertidal flats exist seaward of defences and are subjected to tides, while coastal grazing marshes are largely artificial habitats that exist landward of coastal defences in areas that would otherwise be intertidal habitats. The direct impact of sea-level rise on coastal wetlands is assessed following the broad scale model of McFadden *et al.* (2007) (see also Richards *et al.*, 2008). The wetland change/loss component accounts for both habitat loss and habitat change, where the three influencing factors of *accommodation space*, *sediment supply*, and *rate of relative sea-level rise* are considered (see Figure 7.5). Consequently, habitats such as saltmarsh, coastal grazing marsh and intertidal flat can be either lost under high forcing conditions or can experience transition under the low to moderate forcing conditions (as shown in Figure 7.6). The direct effects of sea-level rise and the effects of defence abandonment due to managed realignment are also included. In river valleys, change in inland marshes is a function of change in river flows where existing marshes can increase or decrease as a function of change in floodplains.

The CORINE land cover data is used to establish the baseline of the intertidal habitats: saltmarsh and intertidal flats, and fluvial habitats (inland marshes). However, the designated habitats landward of coastal flood defences are not defined in the CORINE land cover dataset. There is no standard European nomenclature for these areas and they are variously termed: as 'coastal grazing marsh' (in the UK), or 'summer polders' (in the Netherlands/Germany) to give two examples. Therefore, to develop a generic methodology, pasture areas located within the coastal floodplain are assumed to be potential areas for 'coastal grazing marsh' and this term is used for all such habitats in CLIMSAVE. If defences are abandoned or realigned, the new intertidal land experiences a transition to saltmarsh and intertidal flats.

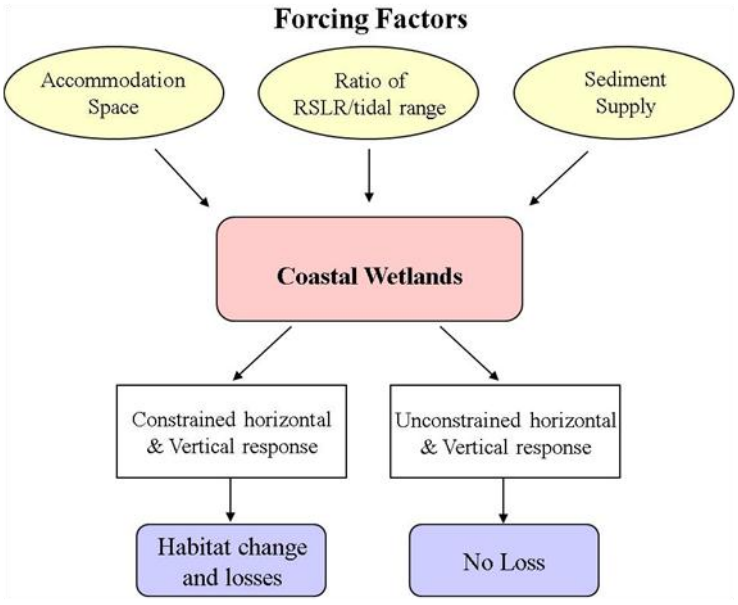


Figure 7.5: The three key forcing factors in the change/loss of coastal floodplain habitats considered in CLIMSAVE.

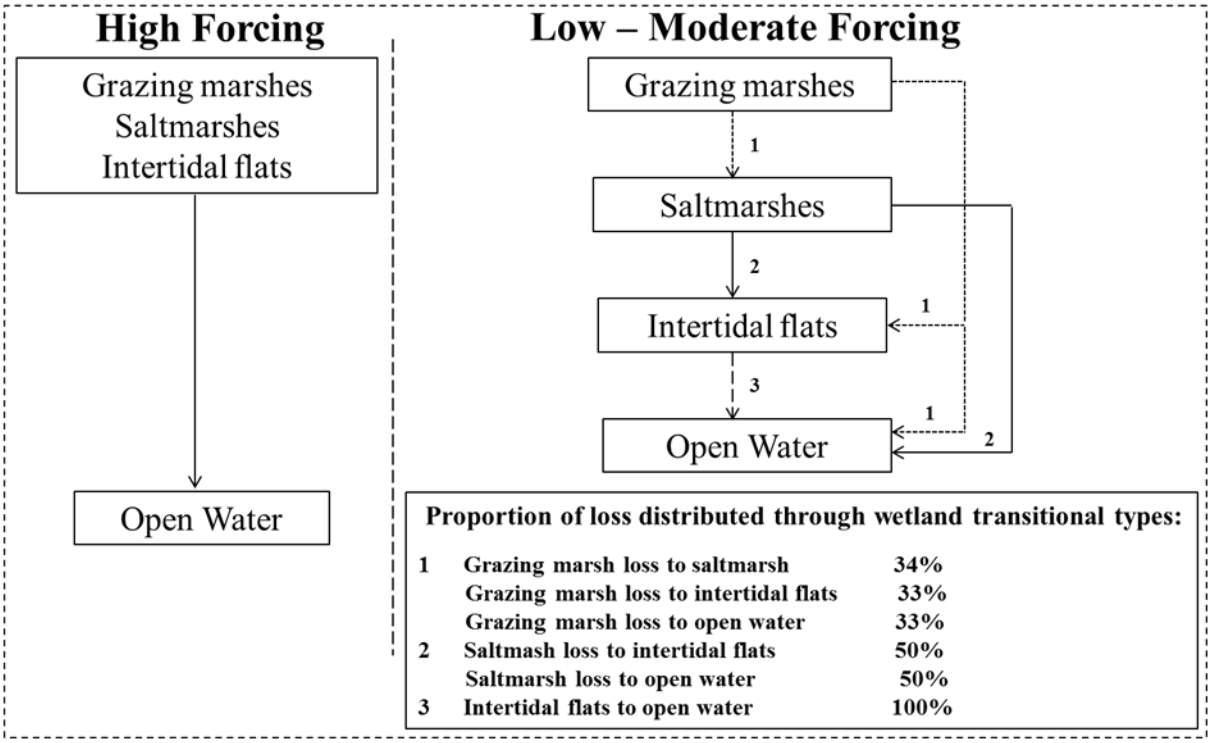


Figure 7.6: Example of modelling wetlands loss/change for coastal areas (adapted from McFadden *et al.*, 2007).

7.2.4 *Data pre-processing and indicators*

Estimating the indicative level of flood protection across Europe

There is no European level dataset on existing flood protection levels for coastal and river areas. Hence, the CORINE land use/cover classes in the impact zones has been used to estimate indicative standards for flood defences (coastal and fluvial) for Europe following the UK DEFRA methodology (MAFF, 1999). The resulting dataset has been calibrated using published data on flood protection in individual regions/nations in Europe – for example, the Netherlands has built an extensive coastal defence system that provides protection up to the 1 in 10,000 year flood event, while the Thames Barrier provides the city of London with protection against a 1 in 1000 year flood event, and we have the national flood defence data for England and Wales. This method provides a consistent approach for establishing a European dataset on flood protection without representing any entitlement or obligation for achieving these protection levels. Table 7.1 shows the indicative standards of protection for five land use bands in fluvial and coastal flood zones considering an indicative range of land use – both the minimum and maximum ranges of fluvial and coastal indicative standard of protection are adopted for the European region. If better local data can be acquired, this data will be included.

Topographical data

The SRTM data at 3 arc second (i.e. almost 90 m) spatial resolution and the Gtopo30 data at 30 arc second (i.e. almost 1 km) spatial resolution have been processed to produce a DTM with full European coverage. The DTM is classified into bands at 0.25 m elevation intervals along the coastline, covering the maximum range of combined sea-level rise, land subsidence

and the extreme storm surge of a 1000 year event. This data set is then gridded at the 10' spatial resolution.

Table 7.1: Indicative standards of protection and land use (from CORINE) (after MAFF, 1999).

Land use band	Description	Land Use (CORINE classes – third level)	Indicative standard of protection	
			Fluvial	Coastal
			Return period (years)	Return period (years)
A	Intensively developed urban areas.	111	50-200	100-300
B	Less intensive urban areas with some high grade agricultural land and/or environmental assets.	112, 121, 122, 123, 124, 131, 141, 142, 211, 212, 213,221, 222, 223	25-100	50-200
C	Large areas of high-grade agricultural land and/or environmental assets with some properties.	132, 133	5-50	10-100
D	Mixed agricultural land with occasional properties at risk of flooding.	241, 242, 243, 244,	1.25-10	2.5-20
E	Low-grade agricultural land (often grass) or seasonally occupied properties at risk.	31, 311, 312, 313, 321, 322, 323, 324, 333	0-2.5	0-5
F		All other classes	0	0

Extreme sea-level data

Four extreme sea-level events (i.e. the 1 in 1, 1 in 10, 1 in 100 and 1 in 1000 return period events) and associated land uplift/subsidence (the local geological component of sea-level change) have been gridded at the 10' resolution. These data are derived from the DIVA database (Vafeidis *et al.*, 2008).

Socio-economic indicators

The socio-economic scenarios are used to develop a series of socio-economic indicators relevant to flooding as follows:

- Change in GDP is used to reflect the change in economic conditions and how these will influence the flood damages of properties' contents.
- Average household size: this indicator allows the number of properties to be estimated as a function of population. The NUTS3 data set provides the average household size for the baseline - this data is gridded at 10' spatial resolution.

- Population density: the population density is used to estimate the number of people in flood risk areas. The NUTS3 data set provides this variable for the baseline - this data is gridded at the 10' spatial resolution.

7.2.5 *Adaptation strategies within the CFFlood meta-model*

The adaptation strategies investigated within the CFFlood meta-model are designed to focus on human safety and/or an environmental emphasis (to sustain or enhance habitats) (see Table 7.2).

Table 7.2: Adaptation measures for the CFFlood meta-model.

Policies	Emphasis on Human Safety	Emphasis on Environment
1. Flood protection upgrade	√	
2. Retreat of flood defences		√
3. Flood resilience measures	√	
4. Mixed response	√	√

Emphasis on Human Safety

These adaptation measures aim to reduce flood risks (for people and properties) through the following three categories:

a) Flood protection upgrade by 50%, 100%, 500% and 1000%: this will be applied directly to the baseline protection levels and applied uniformly for all Europe.

b) Resilience measures: considering that new properties will not be affected by flooding (e.g., by raising them above ground levels) up to a predefined threshold of flood event (i.e., 100 year event), while the old properties will suffer from flood damage.

c) Mixed response: this provides a more realistic adaptation option, where a plausible combination of flood protection improvement (i.e. 100% upgrade) and retreat of defences to maintain habitat is investigated.

Emphasis on Environment

This includes the possibility of either maintaining wetland habitat areas at the baseline level or doubling the area of these habitats. Habitat area losses will be determined using the habitat model by comparing with the baseline stocks, while the rules for determining candidate sites for habitat creation (via retreat) include the following:

Retreat rules:

- Retreat will take place in areas inside the floodplains (coastal and fluvial).
- Non-urbanized areas are considered.

Rules at habitat levels:

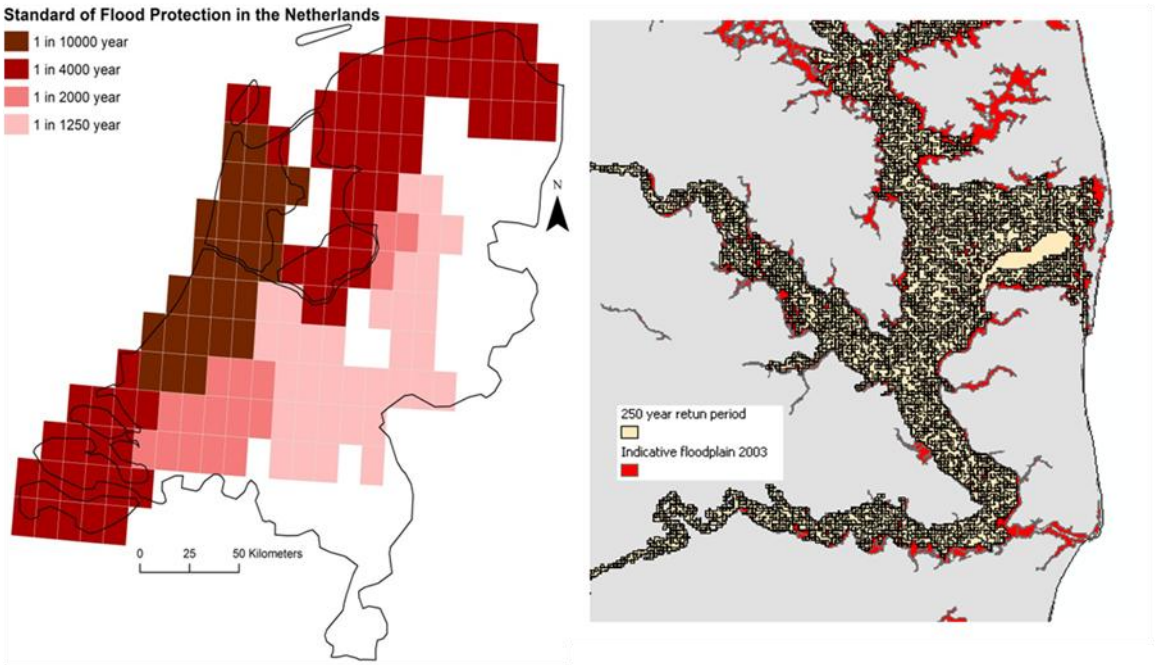
- Saltmarsh and intertidal flat: the coastal grazing marsh areas that will be changed to saltmarsh due to a change in salinity in the habitat model will be considered to be suitable for saltmarsh. Any other areas within the coastal floodplain will be considered candidate for saltmarsh. It is assumed that areas that are not at the correct height are raised or lowered to an optimum height. The available areas can be split

between candidate areas for creating saltmarsh and candidate areas for intertidal flat (e.g. 50% for each).

- Coastal grazing marsh: assumes all pasture can be considered as candidates for coastal grazing marsh.
- Inland marshes: is based on the baseline estimate and considers the change in peak flow. The fluvial model assumes that an increase in peak river flow (Q_{MED}) will lead to larger floodplains and accounts for an increase in the areas of inland marshes of the same proportion.

7.3 CFFlood meta-model calibration and validation

The input parameters into the CFFlood meta-model have been calibrated using published data and studies. Figure 7.7(a) shows an example of the coastal flood protection levels used for the Netherlands (van der Most, 2011) in the calibration of the flood protection dataset produced from the classification of the CORINE land use/cover data. The flood maps are also verified using available flood maps, e.g., Figure 7.7(b) shows a good match against the flood map of the 250 year return period and the 200 year indicative flood map (2003) in the Norfolk region (in England).



(a) Flood protection in the Netherlands.

(b) Screenshot of the 250 year flood map and the 200 year indicative flood map for an area in eastern England (part of Norfolk).

Figure 7.7: Calibrating/validating the input parameters into the CFFlood meta-model.

Figures 7.8 A & B illustrate the effect of the calibrated protection data for the Netherlands, shown in Figure 7.7, on the model outputs of the number of people affected and damage due to flooding. Without protection, the outputs reflect the large area of land at risk of flooding, and hence high impacts on people and properties (as shown in the left top and bottom figures). The model outputs under the minimum and maximum protection realistically represent the existing defence systems (Figure 7.7) that protect the extensive low-lying area behind defences.

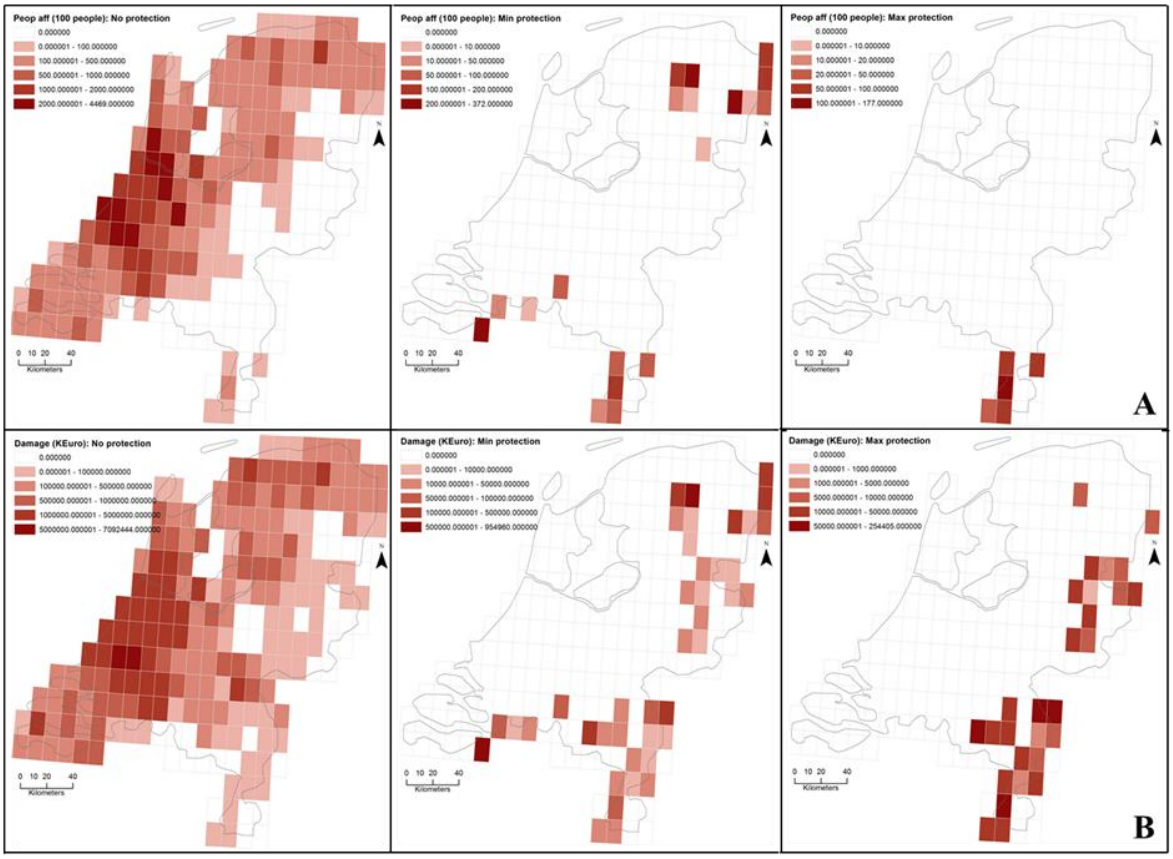


Figure 7.8: Validating output parameters of the CFFlood meta-model: (A) People affected and (B) Damage, due to a 1 in 100 year flood event at the baseline.

7.4 Integration of CFFlood with the other sectoral meta-models

The CFFlood model interacts with a number of the other CLIMSAVE models, either being provided with inputs, or providing outputs to later models in the simulation process:

- *The RUG urban model*: Input data on the areas (in sq. km) of residential (CLC category 1.1) and non-residential areas (CLC categories 1.2 - 1.4) within each grid cell is provided by the RUG meta-model (Section 5) which are used to assess socio-economic impacts.

- *The WaterGAP water model*: Changes in peak river flows (relative change in Q_{MED}) are derived from the WaterGAP model (Section 8) for use in the analysis of fluvial flood risk.

- *The SFARMOD agricultural land use model*: The results from the coastal and fluvial flood analysis are an input to the SFARMOD model (Section 10) – it estimates areas that are not available for arable farming due to a flood frequency of more than once every 10 years. In addition, areas that are flooded more than once a year are considered not suitable for any type of farming (Mokrech *et al.*, 2008).

- *The SPECIES biodiversity model*: Outputs on the areas of floodplain habitats are used as inputs to the SPECIES model (Section 13) to mask the potentially suitable climate space for individual species associated with saltmarsh and coastal grazing marsh habitats.

7.5 Illustrative model sensitivity analysis and scenario results

The following two sections provide illustrative model outputs of selected sensitivity analysis and scenario results of the CFFlood model from the IAP.

7.5.1 *Illustrative sensitivity analysis*

Figures 7.9 through 7.12 present illustrative sensitivity analysis results for the socio-economic and environmental flood impacts of a 1 in 100 year flooding events (both coastal and fluvial) in Europe due to changes in climate and socio-economic drivers including sea-level rise, population, GDP and level of protection. At the baseline, the European total impacts due to a 1 in 100 year flooding are approximately: over 24 million ha of land at risk, over 29 million of people at risk, about 21 million people flooded, and over 133 billion Euros of damages under the minimum baseline flood protection. When no protection is considered, damages due to flooding and the number of people flooded increases by a factor of 1.9 and 1.4, respectively, while a maximum protection decreases the impacts by a factor of 0.5 and 0.7, respectively (see Figures 7.10C and 7.11D).

Considering a "what if" scenario of 2m rise in sea level (by 2100), the European total area at risk of flooding increases by a factor of about 1.1 (Figure 7.9). The number of people at risk (threatened people) also increases by up to about 36 million people – a factor of 1.2 (Figure 7.10). Considering the minimum protection, the impacts on properties (both structural and contents damage) (Figure 7.11A) and people (actually flooded) (Figure 7.10A) grows by factors of about 2.8 and 1.7, respectively.

Similarly, changes in population and GDP affect the potential impacts significantly. For example, a 50% increase in the European population at the baseline increases the total European socio-economic impacts by a factor of about 1.6 (Figures 7.10B and 7.11B). Furthermore, a 180% increase in GDP at the baseline increases economic damages by a factor of 2.4 (Figure 7.11C).

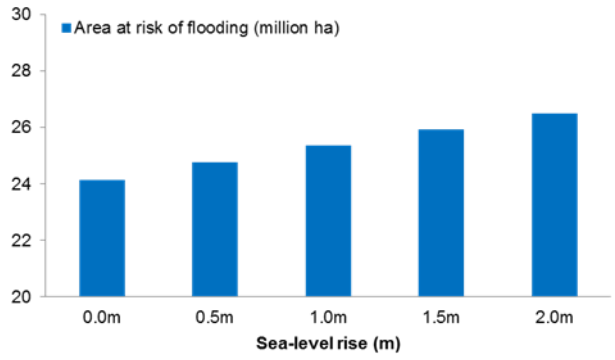


Figure 7.9: The sensitivity of the area at risk of a 1 in 100 year flooding event (million ha) for different amounts of sea-level rise.

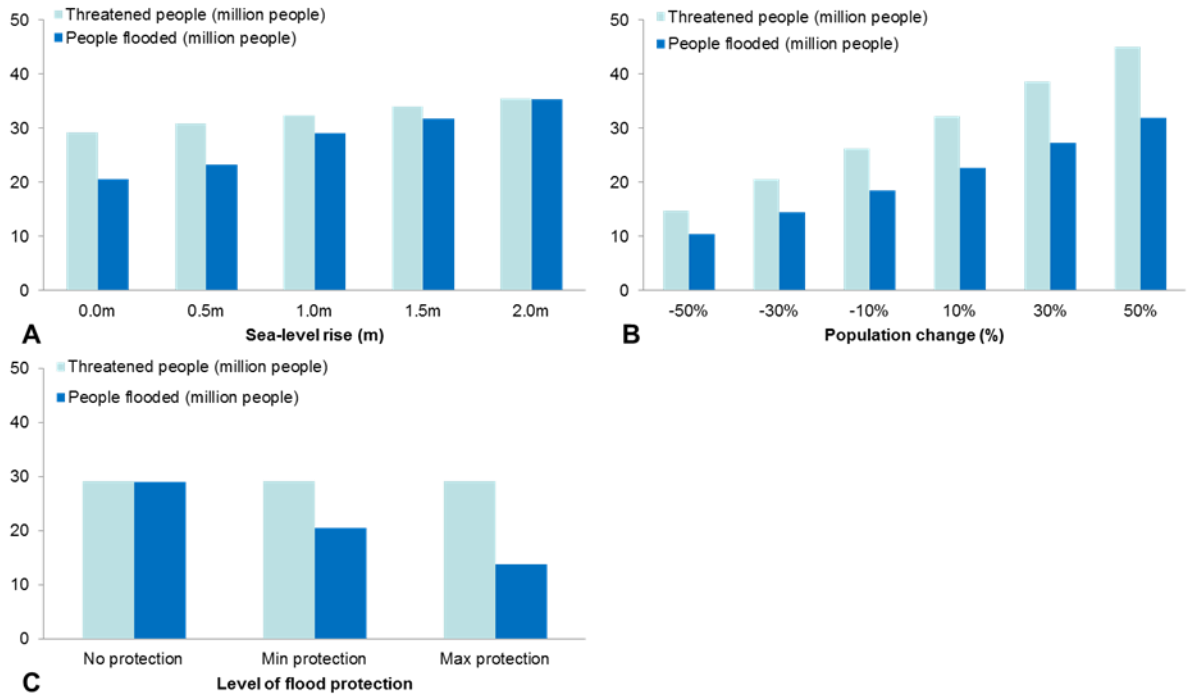


Figure 7.10: The sensitivity of impacts of a 1 in 100 year flooding on people due to: (A) sea-level rise, (B) population change, and (C) change in level of protection. Note: (A) and (B) are run with minimum level of protection.

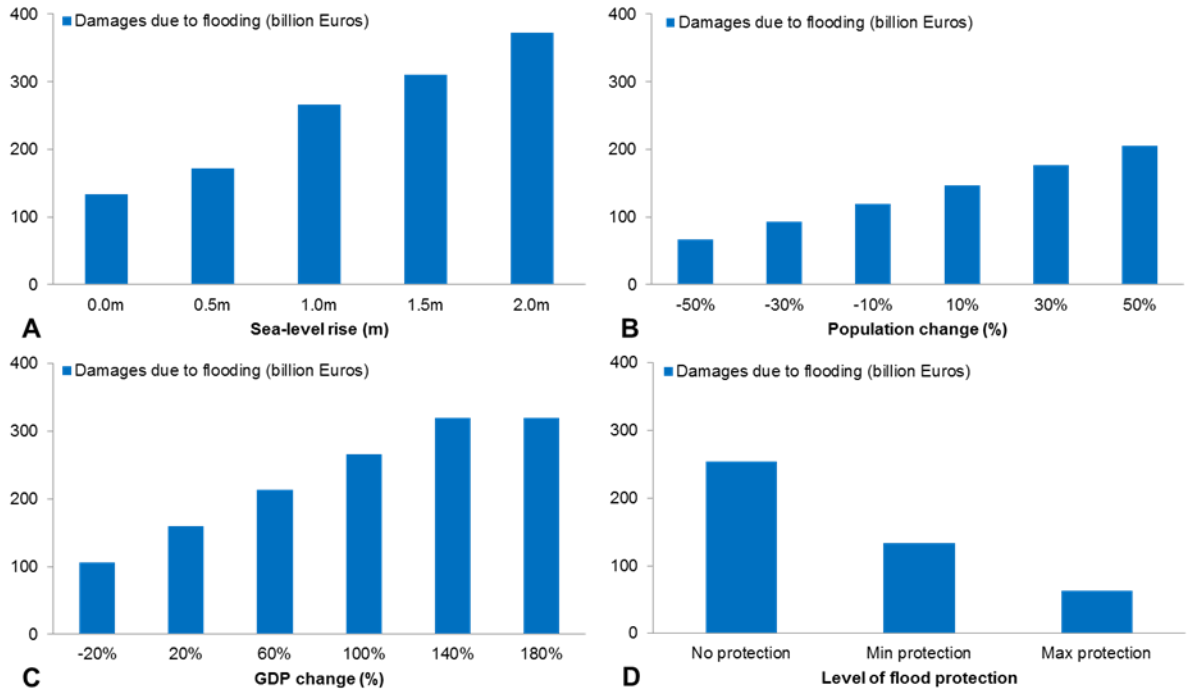


Figure 7.11: The sensitivity of economic damages for properties due to a 1 in 100 year flooding in Europe to: (A) sea-level rise, (B) population change, (C) GDP change, and (D) change in level of protection. Note: (A), (B), and (C) are run with minimum level of protection.

Figure 7.12 illustrates the sensitivity of coastal habitats to sea-level rise (by 2100). Two metres of sea-level rise leads to a significant loss of saltmarshes and intertidal flats from 268 thousands ha and 677 thousands ha (at the baseline) to 40 thousands ha and 179 thousands ha (by 2100), respectively. This is a reduction of 85% and 74%, respectively.

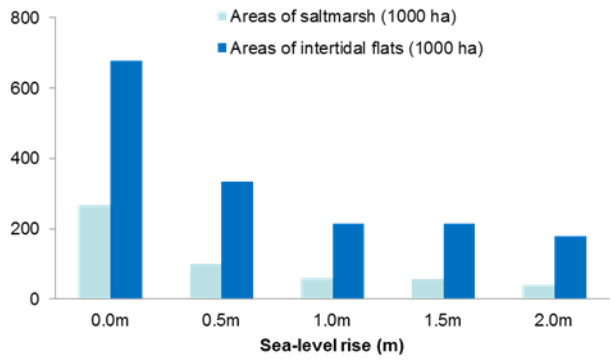


Figure 7.12: Loss of saltmarsh and intertidal flats in Europe due to sea-level rise by 2100, under minimum level of protection.

7.5.2 *Illustrative scenario results*

Figure 7.13 illustrates scenario results of the CFFlood model outputs due to a 1 in 100 year flooding from the IAP based on selected climate and socio-economic scenario default settings for the 2020s and 2050s (see Table 7.3). Under these scenarios, the European total socio-economic impacts range between 19 to 22 million people affected and 142 to 159 billion Euros of economic damage by the 2020s (Figures 7.13 A & C). These impacts increase up to 25.3 million people affected and 278 billion Euros of economic damage by the 2050s (Figures 7.13 B & D). The relatively high number of people affected by flooding (by the 2050s) is experienced under the *Should I Stay or Should I Go* scenario, which is associated with the highest growth in European population and an increase in the number of people living in floodplains. Similarly, higher economic damages are expected under the *We Are the World* and *Riders on the Storm* socio-economic scenarios, which are associated with a high growth in GDP (see Table 7.3).

Table 7.3: Selected climate and socio-economic scenario settings used to illustrate scenario results of the model.

Climate Scenario				
Timeslice	**2020s (Default)**		**2050s (Default)**	
Emissions	A1		A1	
Climate model	CSMK3		CSMK3	
Climate sensitivity	Middle		Middle	
Sea level change	+0.09 m		+0.21 m	
Socio-economic Scenario				
	2020s (Default)		**2050s (Default)**	
Scenario	**Population change (%)**	**GDP change (%)**	**Population change (%)**	**GDP change (%)**
We are the World	-7	26	3	94
Icarus	5	0	-9	0
Should I Stay or Should I Go	5	0	23	-36
Riders on the Strom	5	0	16	54
Level of protection	***Minimum protection***		***Minimum protection***	

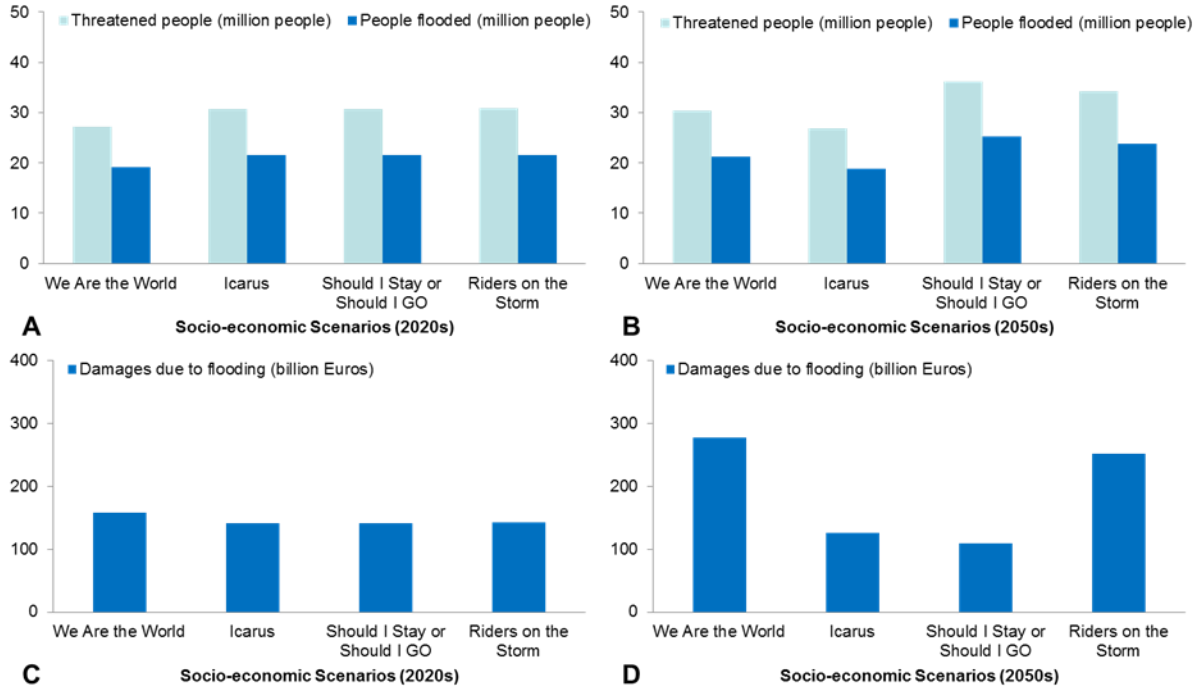


Figure 7.13: The impacts of a 1 in 100 year flooding due to climate and socio-economic change by the 2020s and 2050s under different socio-economic scenarios and CSMK3 climate scenario (see Table 7.3). Note that the default minimum protection is considered here.

7.6 References

Evans, E., Ashley, R., Hall, J., Penning-Rowsell, E., Saul, A., Sayers, P., Thorne, C. & Watkinson, A. (2004a). Foresight. Future Flooding, Scientific Summary: Volume I - Future risks and their drivers. Office of Science and Technology, London.

Evans, E., Ashley, R., Hall, J., Penning-Rowsell, E., Saul, A., Sayers, P., Thorne, C. & Watkinson, A. (2004b). Foresight future flooding. Scientific summary: volume II - managing future risks. Foresight, Office of Science and Technology, London.

Feyen, L., Dankers, R., Bódis, K., Salamon, P. & Barredo, J.I. (2011). Fluvial flood risk in Europe in present and future climates. Submitted to Climatic Change.

MAFF (1999). Flood and coastal defence project appraisal guidance (FCDPAG3), vol. 3. Economic appraisal. Defra publications ref. PB 4650, Defra, London.

Linham, M., Green, C. & Nicholls, R. J. (2010). Costs of adaptation to the effects of climate change in the world's large port cities. AVOID - Avoiding dangerous climate change report AV/WS1/D1/02. London, UK; Department of Energy and Climate Change (DECC) and Department for Environment Food and Rural Affairs (DEFRA).

Linham, M. & Nicholls, R.J. (2010). Technologies for Climate Change Adaptation: Coastal Erosion and Flooding. TNA Guidebook Series, UNEP Risø Centre on Energy, Climate and Sustainable Development, Roskilde, Denmark, 150pp.

McFadden, L., Spencer, T. & Nicholls R.J. (2007). Broad-scale modelling of coastal wetlands: what is required? Hydrobiolgica, 577, 5-15.

Mokrech, M.., Nicholls, R.J., Richards, J.A., Henriques, C., Holman, I.P. & Shackley, S. (2008). Regional impact assessment of flooding under future climate and socio-economic scenarios for East Anglia and North West England. Climatic Change, 90, 31-55.

Richards, J., Mokrech, M., Berry, P.M. & Nicholls R.J. (2008). Regional assessment of climate change impacts on coastal and fluvial ecosystems and the scope for adaptation. Climatic Change, 90, 141-167.

Rupp-Armstrong S. & Nicholls R.J. (2007). Coastal and estuarine retreat: A comparison of the application of managed realignment in England and Germany. Journal of Coastal Research, Volume 23 (6), 1418-1430.

Thorne, C., Evans, E. & Penning-Rowsell, E. (editors) (2007). Future flooding and coastal erosion risks, Thomas Telford, London.

Vafeidis, A.T., Nicholls, R.J., McFadden, L., Tol, R.S.J., Spencer, T., Grashoff, P.S., Boot, G. & Klein, R.J.T. (2008). A new global coastal database for impact and vulnerability analysis to sea-level rise. Journal of Coastal Research, 24 (4), 917-924.

8. Development and validation of the WaterGAP water resources and water use meta-models

Florian Wimmer, Stephanie Eisner and Martina Flörke
Center for Environmental Systems Research, University of Kassel, Germany

8.1 WGMM model description

The WaterGAP meta-model (WGMM) is used in the IAP to assess the impacts of global change on water resources and water use in Europe. WGMM is designed to be a surrogate for the global hydrology and water use model WaterGAP (Water - Global Assessment and Prognosis), which has been developed at the Center for Environmental Systems Research (CESR) to provide an integrated perspective of the impacts of global change on the water sector (Alcamo *et al.*, 2003; Döll *et al.*, 2003). WaterGAP consists of two main components: a global hydrology model and a global water use model.

In order to achieve a very short runtime, the spatial detail of WGMM is reduced from more than 180,000 grid cells of WaterGAP3 for Europe to about 100 spatial units larger than 10,000 km². Those spatial units, hereafter referred to as river basins, are made up either by single large river basins or clusters of smaller, neighbouring river basins with similar hydro-geographic properties. For each river basin, the meta-model simulates the output parameters given in Table 8.1, which are long-term statistics of the corresponding WaterGAP3 results for 30-year time periods. Moreover, WGMM output parameters related to river flow, i.e. Q_{95}, Q_{avg}, Q_5, and Q_{med}, are downscaled to the 10' x 10' grid cells used by other meta-models in the IAP.

Table 8.1: WGMM output parameters.

Model output parameter	Description	Spatial level
Q_{avg} (m³/s)	Long-term average river discharge	Grid cell
Q_{95} (m³/s)	Low flow river discharge (exceeded in 95% of the days)	Grid cell
Q_5 (m³/s)	High flow river discharge (exceeded in 5% of all days)	Grid cell
Q_{med} (m³/s)	Flood flow, median of the annual maximum daily discharge	Grid cell
Ecosystem service indicator (ESI) for flow regulation	Difference of Q_5 and Q_{95} normalized by Q_{avg}	Grid cell
Water availability (mil. m³ / y)	Annual renewable water resources	River basin
Water available for agriculture (mil.m³/y)	Water availability minus water consumption in other sectors	River basin
Water availability per capita (m³/cap/year)	Ratio of water availability and number of people	River basin
Total water use (mil. m³/y)	Total water use (withdrawals and consumption)	River basin
Water stress indicator (-)	Water withdrawals-to-availability ratio	River basin
ESI for drinking water provision	Satisfaction of water demand (withdrawals) in domestic sector	River basin
ESI cooling water	Satisfaction of water demand (withdrawals) in thermal electricity production	River basin

8.1.1 *The WGMM hydrology model*

The aim of the hydrological model WaterGAP is to simulate the characteristic macro-scale behaviour of the terrestrial water cycle in order to estimate water availability. Based on the time series of climatic data, the hydrological model calculates the daily water balance for each grid cell, taking into account physiographic characteristics, such as soil type, vegetation, slope, and aquifer type. Runoff generated on the grid cells is routed to the catchment outlet according to a global drainage direction map (Lehner *et al.*, 2008), taking into account the extent and hydrological effects of lakes, reservoirs, dams and wetlands. The model is calibrated by adjusting one free parameter, which controls the fraction of total runoff from effective precipitation in order to minimise the error in simulated long-term annual discharge.

For the current version, WaterGAP3, the spatial resolution of the model raster has been increased from 30' x 30' to 5' x 5'. Partly enabled by this finer spatial resolution, the process representations of runoff formation and runoff concentration in the hydrological model have been substantially improved, including:

- Revision of the snow dynamics on the sub-grid scale (Verzano & Menzel, 2009);
- Representation of permafrost occurrence directly affecting groundwater recharge (aus der Beek & Teichert, 2008);
- Implementation of a variable flow velocity algorithm (Schulze & Döll, 2004);
- Introduction of a meandering factor to improve the representation of river length (Lehner *et al.*, 2008);
- Estimation of potential evapotranspiration and ground water recharge taking into account Köppen's climatic regions (Weiß, 2009);
- Implementation of dams from the Global Reservoir and Dam Database (GRanD) and the European Lakes and Reservoir Database (ELDRED2) in order to consider anthropogenic flow regulation (Döll & Fiedler, 2009).

These model revisions are a prerequisite for the application of WaterGAP3 to analyse the hydrological extremes in addition to long-term water availability. The model's general ability to simulate flood discharges has been evaluated by Verzano (2009).

The meta-model makes use of a look-up table populated with the results of 273 pre-run WaterGAP3 simulations, aggregated for river basins, driven by monthly CRU climate input (Mitchell & Jones, 2005) with simultaneously modified mean temperature and precipitation. A constant offset was added to all values in the input time series of temperature leading to a shift in mean annual temperature while the dynamics are not changed. The manipulation of precipitation was done in a similar manner except for multiplying the values by a factor instead of adding an offset. The applied temperature variations range from 0.0 to +6.0 K in steps of 0.5 K while precipitation variations range from -0.5 (-50%) to 1.5 (+50%) in steps of 0.05.

During the runtime, the WGMM derives the change in temperature and precipitation for the individual river basins from the 10' climate input provided by the IAP as compared to the baseline. According to these changes, the corresponding river basin-level changes of the flow parameters Q_{avg}, Q_{95}, Q_5, and Q_{med} are taken from the look-up table and are subsequently downscaled to the IAP grid using a pattern scaling technique. The model works with a 10' x 10' raster dataset representing the baseline conditions for all flow parameters. In order to calculate river flow under climate change on a grid cell-level, each raster value is multiplied

by a factor representing the relative change in the given flow parameter. There is exactly one factor for each river basin, while each grid cell entirely belongs to exactly one river basin.

8.1.2 *The WGMM water use model*

WGMM provides simplified estimates of water withdrawals (WW) and water consumption (WC) in the domestic sector, in manufacturing industry, and in thermal electricity production. The modelling approach is based on gridded results (5' x 5') of WaterGAP3 for the base year 2005 (EU FP6 project SCENES), which were aggregated to the river basin level.

In WaterGAP3, the domestic sector includes household use, small businesses and other municipal uses. The basic approach of the domestic water use model is to first compute the domestic water use intensity (m^3/cap-year) and then to multiply this by the population of water users. Changes in water use intensity are expressed by structural changes and technological changes (Alcamo *et al.*, 2003; Flörke & Alcamo, 2004, Flörke et al., in press). The concept of structural change is based on the observation that as average income increases, water consumers tend at first towards a more water-intensive lifestyle. Finally, a maximum level is reached after which the per capita water use is either stable or declines. In this way, human behaviour is covered. The relationship between water use intensity and income (GDP/capita) is derived for each country by a fit to historical data. Water use is then downscaled to river basins according to the spatial distribution of population across Europe.

WaterGAP3 simulates WW in the manufacturing sector on a country scale based on the specific structural water use intensity, i.e. the ratio of water use to the manufacturing gross value added (GVA), which is derived from the base year (Flörke & Alcamo, 2004, Flörke et al., in press). The product of country-specific water use intensity and the scenario values for GVA yields the country wide WW, which are re-scaled to river basins according to sub-national statistics and the spatial distribution of urban population.

In WaterGAP3, the amount of freshwater abstracted for cooling purposes in thermal electricity production is computed for each power plant as the product of the annual thermal electricity production (TEP in MWh) and the water use intensity of the power station (m^3/MWh). The total annual cooling water needs in a river basin are then calculated as the sum of the withdrawals of all power plants located within the region (Vassolo and Döll, 2005; Flörke *et al.*, 2011; Flörke et al., in press).

In the meta-model, WGMM, the water withdrawals in a river basin are calculated as

$$W_{s,r} = \sum_{c=1}^{28} \omega_{s,r} \frac{D_{s,c}}{\partial_{s,c}} s_t s_b F_{s,r,c} \quad (8.1)$$

where ω are the baseline WW in the river basin, D is the main model driver in the scenario and δ is the main model driver in the base year. The subscripts denote the water use sector s, the river basin r, and the country c. The factors s_t and s_b represent water savings due to technological improvements and water savings due to behavioural change, respectively. $F_{s,r,c}$ are weighting factors used to translate the country-level relative change in the main driver to the water withdrawals at the river basin scale. There is one set of weighting factors per water use sector s and river basin r calculated as

$$F_{s,r,c} = \frac{\omega_{s,r,c}}{\omega_{s,r}} \; with \; \sum_{c=1}^{n} F_{s,r,c} = 1 \qquad (8.2)$$

where $\omega_{s,r,c}$ are the baseline WW in sector *s* allocated to the spatial intersection of country *c* and river basin *r* and n is the number of countries covered by the model (28). In order to compute WC in the various sectors, sectoral WW are multiplied by a sector-specific consumption factor, derived separately for each river basin as the WC-to-WW ratio.

For each sector, a different main model driver (*D* and δ) is used. In addition, a number of calculation steps differ among the sectors as described in the following:

Domestic sector

The main model driver for domestic water use is the product of population and structural water use intensity (*S).* The latter depends on income and specifies the domestic WW per person and year. For this purpose, a sigmoidal curve $S = a + b\left(1 - e^{\alpha * I^2}\right)$ was fitted to historical data, where *I* is income and *a, b,* and α are curve fitting parameters.

The minimum threshold for annual domestic water withdrawals, W_{min}, in a river basin is calculated as the number of inhabitants multiplied by 18.25 m³/person (50 litres per capita and day). If the result of Equation 8.1 falls short of this threshold because of small values for s_b and s_t, the final estimate of domestic WW is set to W_{min}.

Manufacturing industry

The main model driver for manufacturing water use is GVA. In order to limit the water savings due to technological improvements in this sector, the lower limit for s_t in Equation 8.1 is set to s_t =0.6. Water savings due to behavioural change are not taken into account in the calculation of manufacturing WW, i.e. a value of s_b=1 is used.

Thermal electricity production

The main model driver for cooling water use in the energy sector is thermal electricity production (TEP in MWh). The lower limit for the technological change factor is defined as s_t =0.8 in this sector. Water savings due to behavioural change are not taken into account in the calculation of cooling WW, i.e. a value of s_b=1 is used.

Technological improvements of water use intensity in the energy sector are mainly achieved by a conversion of the cooling system from once-through cooling to tower cooling. In Europe, average water withdrawals per MWh in once-through cooling systems (~112m³/MWh) are about 45 times higher than in tower cooling systems (~2.5 m³/MWh). At the same time, the consumption factor also strongly depends on the type of the cooling system. The average consumption factor of European power plants with tower cooling (0.53) is about 66 times higher than with once-through cooling systems (0.008). From these figures, the relationship between the average water consumption factor and the percentage share of tower cooling systems can be derived (Figure 8.1a). Further, a given s_t -value can be converted into a fraction of power plants with tower cooling systems, defining s_t =1 for a reference fraction of 57% tower cooling in the year 2005 (Figure 8.1b). In combination, this yields a non-linear relationship between the normalised consumption factor $c_{norm}=c_{scenario}/c_{2005}$ and s_t, which can be approximated by a regression function (Figure 8.1c):

$$c_{norm}(s_t) \approx 0.959\, s_t^{-1.111} \quad (8.3)$$

Finally, the consumption factor in the scenario $c_{scenario}=c_{2005}*c_{norm}(s_t)$ is used to calculate WC in the energy sector.

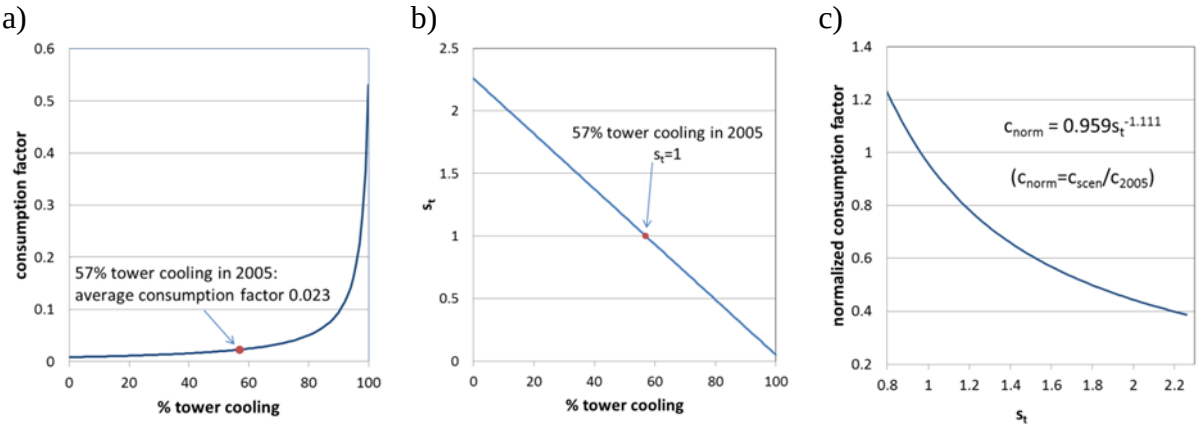


Figure 8.1: (a) Consumption factor as a function of the percentage of tower cooling: 57% of power plants equipped with tower cooling in 2005 corresponds to s_t =1; (b) relationship between the s_t -factor and the percentage of power plants with tower cooling; and (c) non-linear regression function used to calculate the normalised consumption factor as a function of the s_t -factor.

8.2 WGMM model calibration and validation

The modelling approach for hydrological parameters in WGMM is mainly a reproduction of WaterGAP3 results. WaterGAP3 is a state-of-the-art hydrological model for the continental to global scale with a focus on the reliable estimation of long-term water resources and water use. Information on the calibration and validation of WaterGAP3 itself can be found in the literature listed above. In the following paragraphs, it is demonstrated that the additional model uncertainty caused by the major simplifications of the meta-model is still acceptable for the purpose of the IAP.

The daily WaterGAP3 simulations of river discharge that are used to derive Q_{med}, Q_{95} and Q_5 (see Table 8.1) are based on monthly precipitation input, i.e. only total monthly precipitation and the number of rain days per month are known. The model disaggregates this kind of precipitation input to daily values using a statistical approach that leads to a considerable reduction in the day-to-day variability in the resulting 'pseudo-daily' precipitation time series as compared to observations. However, a comparison of simulated versus observed discharges for European gauging stations where daily time series for the period 1971-2000 are available shows fairly good agreement for Q_{med}, Q_{95} and Q_5 (Figures 8.2, 8.3 and 8.4).

Another simplification of the meta-model approach is related to the technique to transfer river basin changes of river discharge to the 10' grid of the IAP. This downscaling is done by multiplyling gridded baseline values by the relative changes at the river basin outlet. Implicitly, this method assumes a uniform relative change in discharge for all segments of a river network although runoff generation and river routing is actually a non-linear process. Hence, there is in general a difference between WGMM results and corresponding

aggregated WaterGAP3 output on the grid cell level as soon as climate input differs from the baseline. Note, that the baseline grids are derived by spatial aggregation of WaterGAP3 output (5') to the IAP grid (10') using the same aggregation routine. The maps in Figure 8.5 show the relative deviation of Q_{avg} simulated by WGMM from aggregated WaterGAP3 output for Q_{avg}. The maps indicate that: (i) in major parts of Europe the deviation is between ±5%, (ii) WGMM tends to overestimate Q_{avg}, and (iii) the overestimation of Q_{avg} increases with increasing precipitation.

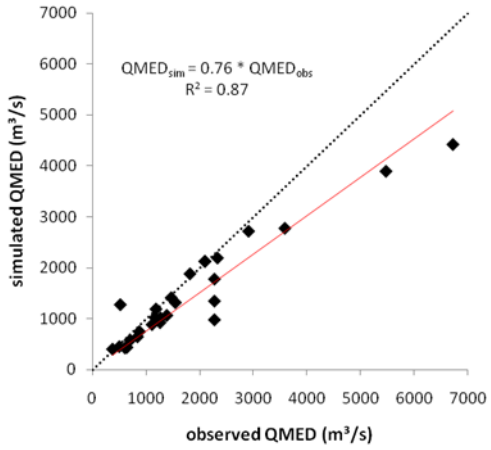


Figure 8.2: Simulated vs. observed flood parameter Q_{med} for 25 gauging stations across Europe, dashed line = 1:1 line, red (solid) line = linear fit.

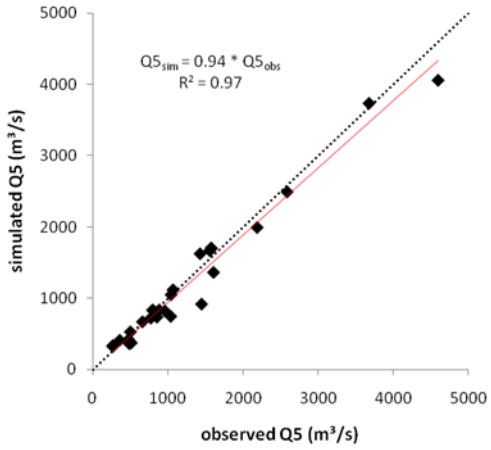


Figure 8.3: Simulated vs. observed high flow parameter Q_5 for 25 gauging stations across Europe, dashed line = 1:1 line, red (solid) line = linear fit.

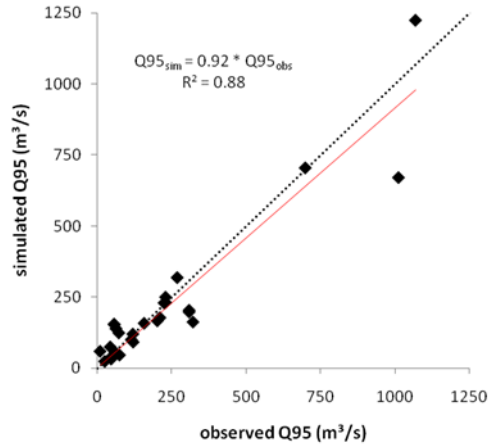


Figure 8.4: Simulated vs. observed low flow parameter Q_{95} for 25 gauging stations across Europe, dashed line = 1:1 line, red (solid) line = linear fit.

For the baseline, the estimates of WW and WC by the meta-model are simply aggregates of the original model output of WaterGAP3, which was validated against the best data sources available (Flörke *et al.*, in press). In a scenario calculation, the estimates for the baseline are scaled proportionally to the relative change in the main drivers of water use. In order to test the performance of the meta-model to approximate the original WaterGAP3 results for a scenario run, the model outcomes of both models were compared for the "Economy First" scenario, which was defined in the EU FP6-Project SCENES (Figure 8.6). The comparison shows that the results of the meta-model very closely match the WaterGAP3 results for manufacturing WW (R^2=0.998). For WW in the domestic sector (R^2=0.975) and thermal

electricity production (R^2=0.975), the model accuracy is somewhat lower but the model still provides a reasonable estimate of future water demand based on given scenario assumptions.

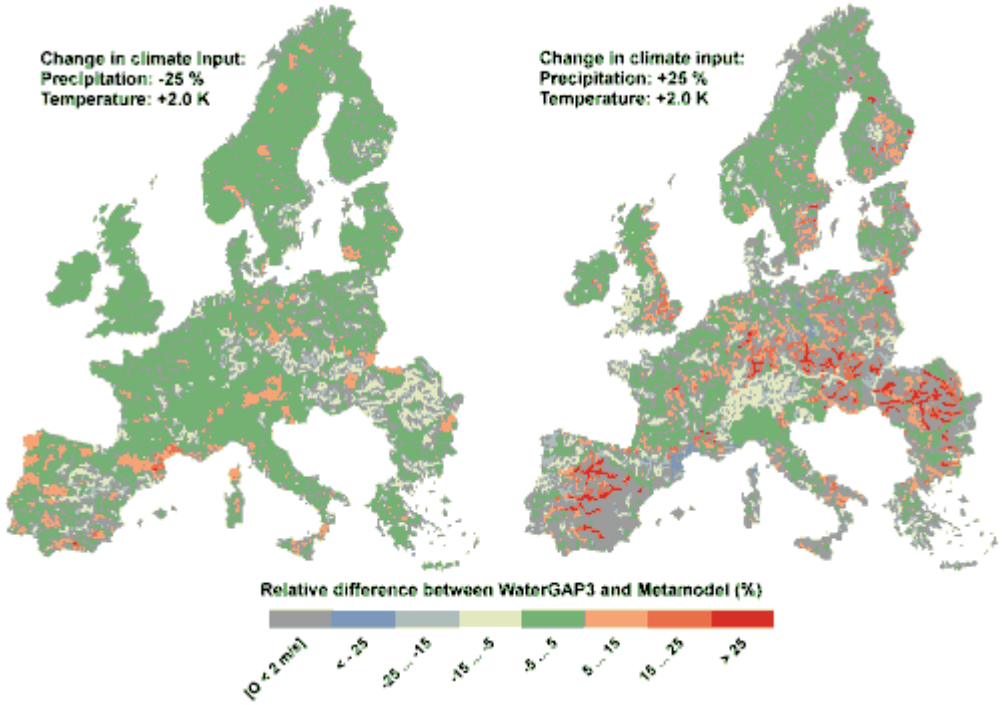


Figure 8.5: Relative deviation of Q_{avg} simulated by WGMM from Q_{avg} simulated by WaterGAP3 (Q_{avg}(WGMM)/Q_{avg}(WaterGAP3)-1) assuming uniform changes in temperature and precipitation across Europe. Left: temperature +2°C / precipitation -25%; right: temperature +2°C / precipitation +25%. (Grid cells with Q_{avg} < 2 m³/s not greyed out).

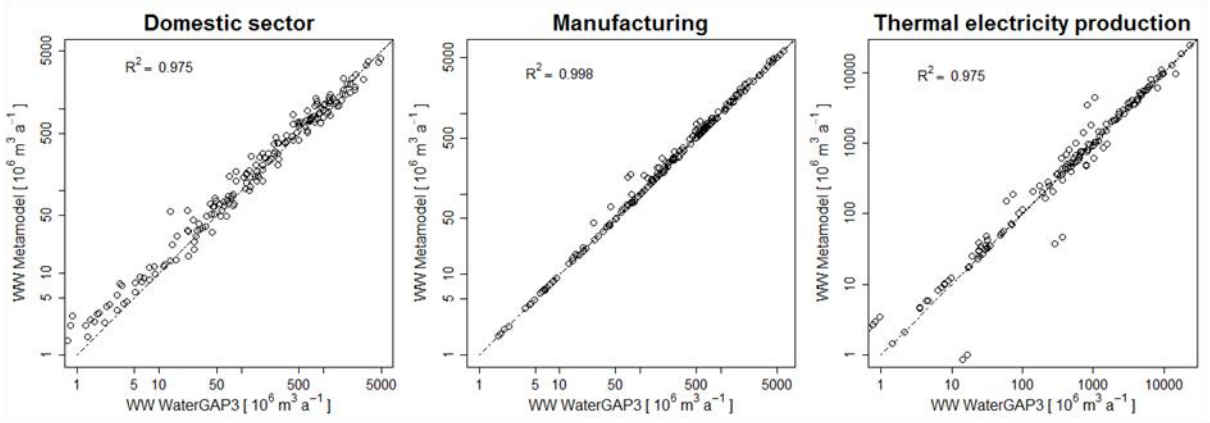


Figure 8.6 Scatter plots comparing meta-model results and WaterGAP3 results of WW in the domestic sector, manufacturing, and thermal electricity production on river basin level for the scenario "Economy first" (FP6-Project SCENES) in 2025 and 2050.

8.3 Integrating WGMM with the other sectoral meta-models

Water use in the agricultural sector is not covered in the WGMM since it is calculated by the agricultural land use meta-model SFARMOD (Section 10). Nevertheless, SFARMOD takes

into account an estimate of WGMM regarding the available water for agricultural use, i.e. mainly irrigation, as the maximum allowed water withdrawals for irrigation.

In order to estimate the amount of water available for agriculture on the river basin scale, WGMM balances the water availability and a "first guess" of total water consumption. The latter is the sum of the projected non-agricultural water consumption plus the agricultural water consumption in the base year. If this demand can be satisfied, water availability for agriculture in the river basin is calculated as water availability reduced by non-agricultural water consumption. In the case of a potential water shortage, a 'water sharing rule' is applied uniformly across all affected river basins to distribute the available water resources to different sectors. The share of water resources falling upon agriculture is passed to SFARMOD. The default rule is to split water resources proportional to the base year conditions. However, the user of the IAP can choose between several rules (Table 8.2). Finally, SFARMOD returns the amount of water actually used in agriculture, which is taken into account by WGMM to correct the "first guess" water use estimates if necessary.

WGMM is also linked to the meta-models SPECIES (biodiversity; Section 13) and CFFlood (flood damages; Section 7). In these cases, WGMM provides input for SPECIES (Q_{avg}, Q_{95}, Q_5) and CFFlood (Q_{med}) but no feedback to WGMM is taken into account. For further information on how WGMM output is used by these meta-models see Sections 7 and 13.

Table 8.2: Water sharing rules currently implemented in WGMM. Abbreviations: WC is water consumption (in 2005), r is the sector share, dom=domestic, man=manufacturing, ele=thermal electricity production, Q95 is the river discharge exceeded in 95% of the time.

Name	Priority sector	Maximum share priority sector	Maximum share other sectors	Water for nature
Baseline	-		Proportional to WC in 2005	-
Prioritising environmental needs	-	-	Relative share as in 2005	$2 * Q_{95}$
Prioritising food production	Agriculture (irrigation)	up to 0.8 if needed	$r_{dom}=WC_{dom}/WC_{sum}$ $r_{man=}WC_{man}/WC_{sum,}$ $r_{ele}=WC_{ele}/WC_{sum}$ with $WC_{sum}= WC_{dom}+WC_{man}+WC_{ele}$	-
Prioritising most important sector	Sector with highest WC in 2005	Up to 0.9 if needed	For each non-priority sector i up to $WC_i/(WC_{total}-WC_{priority\ sector})$	-

8.4 Illustrative examples of scenario simulations

In this section, results for the CLIMSAVE scenario "We are the world" in combination with the climate projection by the CSMK3 model (A1 emission scenario, middle climate sensitivity) are used to illustrate the model output parameters of WGMM. The simulation was carried out with a standalone version of the model. Hence, agricultural water use is not considered because of the missing link to the SFARMOD meta-model (WW and WC in agriculture=0).

In WGMM, simulated water availability only depends on climate conditions. The climate scenario used for this example projects a temperature change between +1°C in Western Europe and more than +6°C in the very north of Scandinavia (Figure 8.7). At the same time, the precipitation change ranges from -30% in southern Europe to +20% in northern Europe (Figure 8.7). These changes in climate correspond to a change in simulated water availability between -76% and +21% (Figure 8.7).

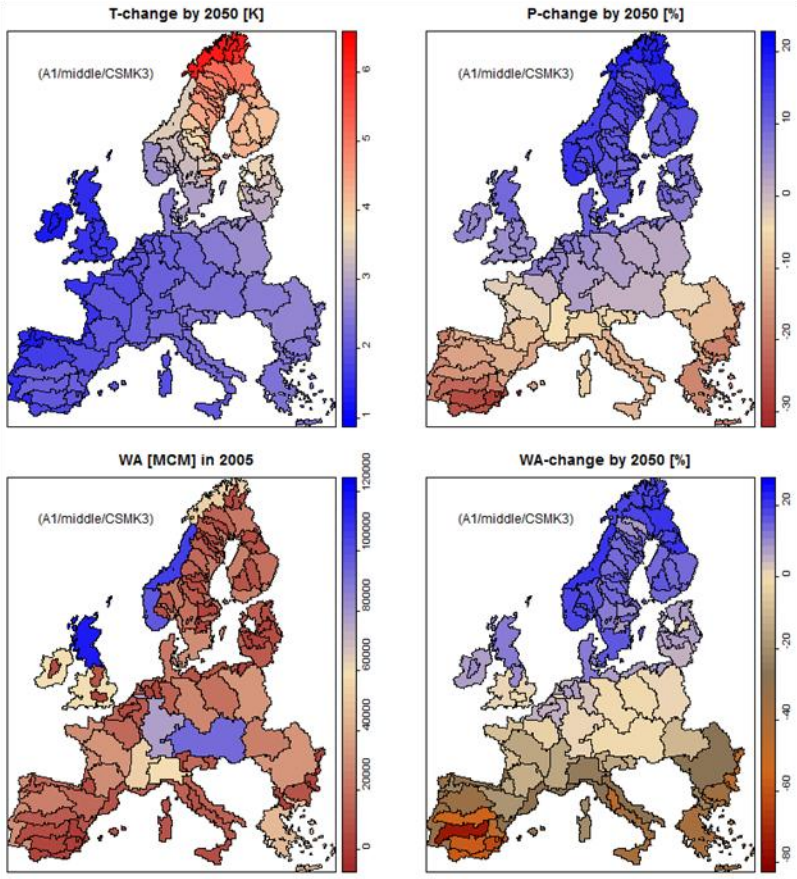


Figure 8.7: Climate input to WGMM, given as T(emperature)-change (top-left) and P(recipitation)-change (top right); water availability (WA) in the year 2005 (bottom left) and change in WA by 2050 (bottom right) for the climate scenario CSMK3/A1/middle sensitivity.

The relevant socio-economic inputs from the scenario "We are the world" (percentage change until 2050) for modelling of water use was:

- 3% population growth → factor 1.03
- 88% increase in income (GDP/cap)
- 94% increase in GVA (gross value added) → factor 1.94
- 29% water savings due to technological improvements → factor 0.71
- 45% water savings due to behavioural change → 0.55
- 78% reduction of TEP (thermal electricity production) → factor 0.22

Domestic WW are reduced considerably between 2005 and 2050 (Figure 8.8, top row) because of the water savings due to technological improvements and behavioural change. In total, the savings (61%) over-compensate by far for the additional WW due to population growth and increasing income. An 88% increase in income results in changes in the country-specific structural water use intensity across Europe by a factor of 1.0-1.33. Hence, the

overall change in domestic WW in the various river basins was between -60% (0.71*0.55*1.03*1.0≈0.40) and -47% (0.71*0.55*1.03*1.33≈0.53).

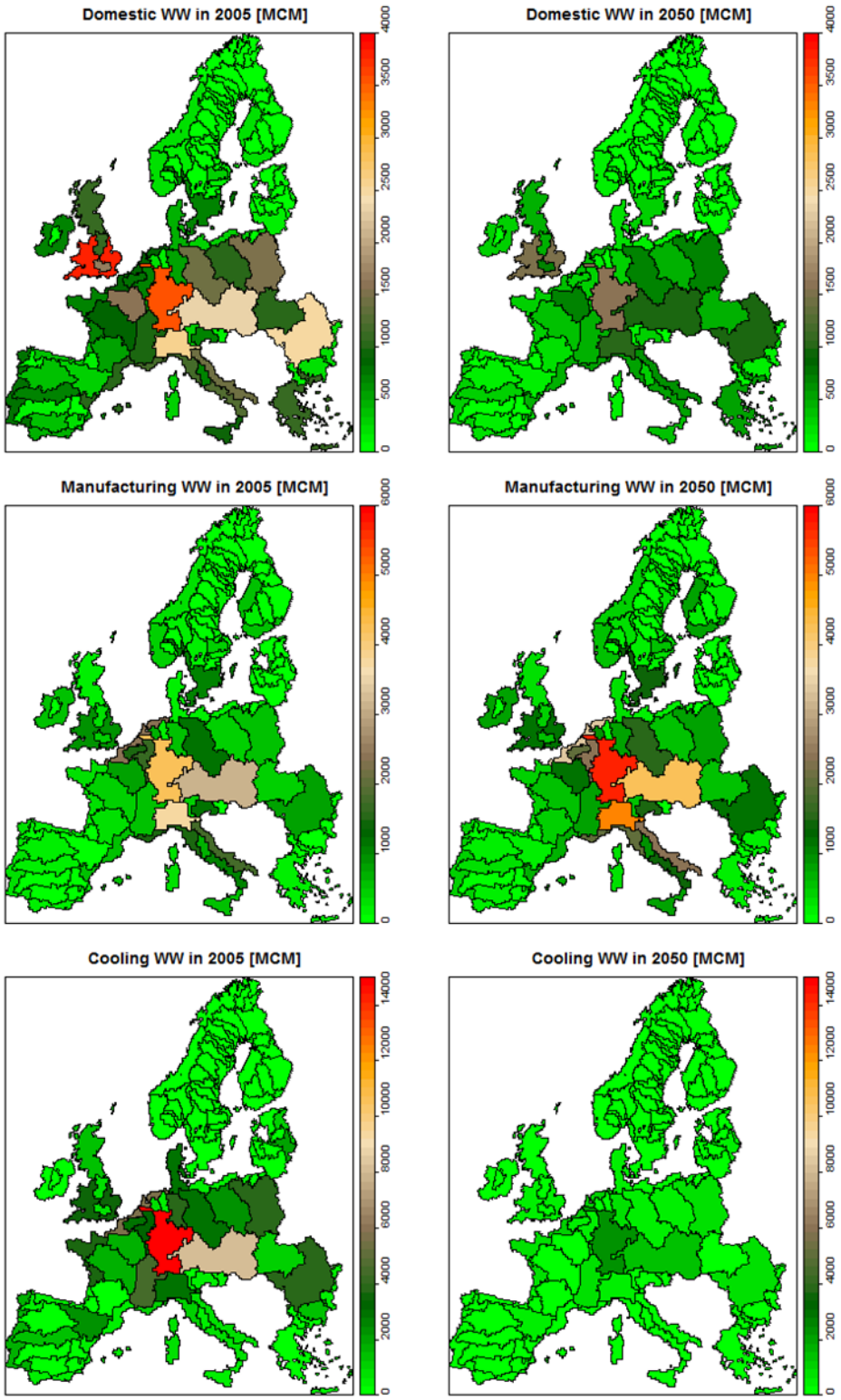


Figure 8.8: Water withdrawals (WW) in the domestic sector (top row), manufacturing industry (middle row), and for cooling in electricity production (bottom row) in the year 2005 (left column) and 2050 in "We are the world" (right column).

A uniform increase in GVA by 94% across all European countries in combination with 29% water saving due to technological improvements results in a 38% increase in manufacturing WW in all river basins (1.94*0.71≈1.38) (Figure 8.8, middle row). Water withdrawals in thermal electricity production show a uniform decrease by 84% because thermal electricity production decreases by 78% in combination with 29% water savings (0.22*0.71≈0.16).

Summing up the different sectors, the total change in WW in the domestic sector, manufacturing industry, and thermal electricity production ranges between -87% and +12%. The actual value in a river basin depends on the percentage share of WW falling upon the individual sectors.

The water exploitation index (WEI) integrates climate change impact on WA and change in water use due to socio-economic change. This water stress indicator is defined as the ratio of WW to WA. WEI values below 0.2, between 0.2 and 0.4, and above 0.4 indicate low, medium, and severe water stress, respectively. Figure 8.9 shows the WEI in the baseline and in the year 2050. A considerable decline in WW in combination with a slight increase in WA reduces water stress levels from "severe" to "medium" in a number of river basins, e.g., in northern France and the Benelux countries. In regions with declining WA, e.g. south-eastern Spain, the water stress level increases despite the reduction in WW. Note that water stress levels are underestimated in this example simulation because agricultural water use is not considered.

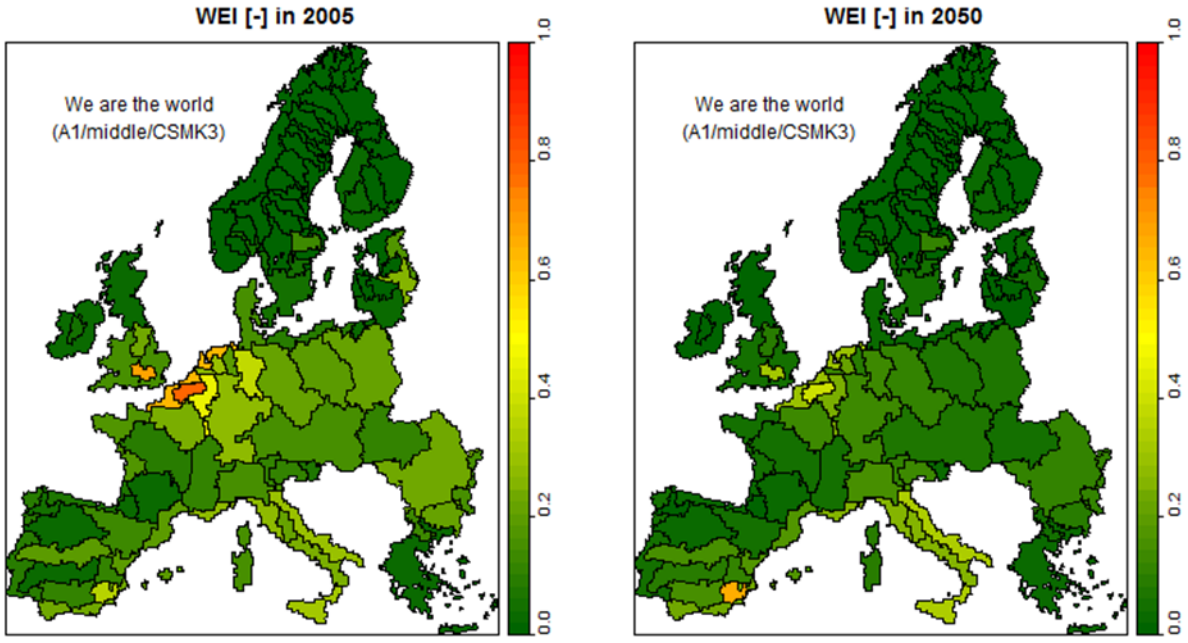


Figure 8.9: Water exploitation index (WEI) in the baseline (2005) (left map) and the year 2050 (right map) under the "We are the world" socio-economic scenario and the CSMK3 climate scenario with A1 emissions and middle climate sensitvity. Water stress levels are underestimated because agricultural water use is not considered in the stand-alone WGMM.

8.5 References

Alcamo, J., Döll, P., Henrichs, T., Kaspar, F., Lehner, B., Rösch, T. & Siebert, S. (2003). Development and testing of the WaterGAP 2 global model of water use and availability, Hydrolog. Sci. J. 48, 317-337.

Aus der Beek, T. & Teichert, E. (2008). Global permafrost distribution in the past, present and future, Proc. 9th International Conference on Permafrost, Fairbanks, USA, 29.6.-3.7.

Döll, P., Kaspar, F. & Lehner, B. (2003). A global hydrological model for deriving water availability indicators: model tuning and validation, J. Hydrol., 270, 105-134.

Döll, P. & Fiedler, K. (2009). Global-scale analysis of river flow alterations due to water withdrawals and reservoirs, Hydrol. Earth Syst. Sci., 13, 2413-2432.

Flörke, M. & Alcamo, J. (2004). European Outlook on Water Use, Technical Report prepared for the European Environment Agency. Kongens Nytorv. 6. DK-1050. Copenhagen, DK URL: // http: scenarios.ewindows.eu.org/reports/fol949029.

Flörke, M., Bärlund, I. & Teichert, E. (2011). Future changes of freshwater needs in European power plants. Management of Environmental Quality: An International Journal 22(1): 89-104.

Lehner, B., Verdin, K. & Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data, Eos, Transactions, AGU, 89(10), 93-94.

Mitchell, T.D. & Jones, P.D. (2005). An improved method of constructing a database of monthly climate observations and associated high-resolution grids. Int. J. Climatology, 25, 693-712.

Schulze, K., Hunger, M. & Döll, P. (2005). Simulating river flow velocity on global scale, Adv. Geosci., 5, 133–136.

Vassolo S. & Döll P. (2005). Global-scale gridded estimates of thermoelectric power and manufacturing water use. Water Resources Research, 41, W04010, DOI:10.1029/2004WR003360.

Verzano, K. & Menzel, L. (2009). Hydrology in Mountain Regions: Observations, Processes and Dynamics, IAHS-Publication 326, Snow conditions in mountains and climate change - a global view, pp 147-154.

Verzano, K. (2009). Climate change impacts on flood related hydrological processes: Further development and application of a global scale hydrological model, Ph.D. thesis, International Max Planck Research School on Earth System Modelling, University of Kassel, Germany.

Weiß, M. (2009). Modelling of global change impacts on hydrology with focus on Europe and Africa, Ph.D. thesis, University of Kassel, Germany.

9. Development and validation of the crop yield meta-models

Miroslav Trnka
Institute of Agrosystems and Bioclimatology, Mendel University, Brno, Czech Republic

9.1 Development of the crop meta-models

The development of the crop meta-models focused on estimating five variables needed by the whole farm model SFARMOD (Section 10), i.e.

- Mean water- and nutrient-limited yield (Yield_Av);
- Mean water-limited yield (Yield_POT);
- Mean water- and nutrient-unlimited yield (YieldPOTI); and
- Sowing date (Sowing) and harvest date (Harvesting).

The development of the crop meta-models for the IAP was affected by the following considerations:

1. A relatively demanding time scale for the crop meta-models to be available and integrated into the IAP;
2. Pan-European coverage for all major crops was required;
3. Several yield levels (e.g. potential as well as water- and nitrogen-limited yields) were required; and
4. The meta-models should include the CO_2 fertilisation effect.

As a result of the above requirements, the CLIMSAVE team opted to use the full agricultural model ROIMPEL that has been validated in previous studies (e.g. Mayr, 1996; Rounsevell *et al.*, 2003, Audsley *et al.*, 2006; Alexandrov, 2006) and used in similar though smaller scale studies (e.g. Audsley *et al.*, 2008; Henseler *et al.*, 2009). In addition ROIMPEL was applied as the principal crop model in earlier FP5 projects e.g. ACCELERATES and ACELCEEC and its outputs used in a number of others (e.g. CECILIA, AGRIDEMA, ADAGIO).

The major advantage of using ROIMPEL is the considerable amount of results available from past EU projects. The data available for the development of meta-models included outputs of the full ROIMPEL model for EU-15 and most of the central and eastern European countries for the baseline climate and 2050 Low, Medium and High climate scenarios. Runs for the period centered around 2080 were also available for more than 50% of the territory. The available outputs of ROIMPEL are actual, potential and irrigated crop yields and crop sowing and maturity dates. The strengths of ROIMPEL are its modularity, the fact that it was developed specifically for GIS-based regional and sub-regional land-use evaluation projects (unlike most detailed crop models) and that initial detailed screening of soil/climate conditions for land suitability for a given crop is performed. The daily dynamics of development stages and of water-, temperature-, and nitrogen stresses are the main crop processes simulated in ROIMPEL which determine the land suitable for a given crop. The accumulation of biomass is based on radiation use efficiency and net photosynthetically active radiation, which is sensitive to CO_2 concentration. The radiation-potential daily biomass increase is corrected according to the temperature, water and nitrogen stresses. Additional penalties on crop yields are included through alarm criteria (for example, for

unfavorable weather parameters during the most sensitive development stages) based on crop specific physiology.

Meta-models have been developed for the following crops:

- Winter wheat and spring wheat;
- Winter barley and spring barley;
- Winter oil seed rape;
- Potatoes;
- Grain maize;
- Sunflower;
- Soybean;
- Cotton;
- Grass; and
- Olives.

Sets of soil and climate predictors for the meta-models were selected based on available databases to emulate the full crop model results. The soil data were characterised by:

- The available water content in the rooting depth (1 parameter);
- The proportion of this water available between five suction levels between Wilting Point and Field Capacity (4 parameters);
- Surface soil texture index, estimated using the formula Int[(4c+2z+d-78)/22.2] where c, z, and d are the percentage clay, silt and sand respectively and Int[x] is the integer part of x. The index increases as the soil becomes heavier – more clayey than sandy (1 parameter); and
- Rooting depth, surface horizon hydraulic conductivity and wilting point soil moisture water content (3 parameters).

The climate data used by the full crop model consists of daily air temperature (maximum and minimum), precipitation, potential evapotranspiration and solar radiation. These daily data are generated from monthly means and the climate data were therefore characterised by:

- Mean annual potential evapotranspiration (PET);
- Mean sum of PET from April to June;
- Mean sum of PET from July to September;
- Mean annual sum of precipitation;
- Proportion of precipitation from April to June;
- Proportion of precipitation from July to September;
- Mean annual temperature;
- Mean temperature from April to June;
- Mean temperature from July to September;
- Mean temperature from December to February;
- Mean maximum temperature from June to August;
- Mean minimum temperature from December to February;
- Mean annual sum of global radiation;
- Proportion of global radiation from April to June;
- Proportion of global radiation from July to September; and

- Ambient concentration of carbon dioxide in the centre of the particular time-slice.

The preparation of the crop meta-models was a two-step procedure. The first versions of the crop meta-models were based on step-wise regression models. These produced outputs in the expected range, allowing the identification of the best set of predictors, but lacked precision and reliability. The second version which were integrated into the IAP are based on artificial neural networks (ANN) combined with temperature thresholds to prevent crops growing in unsuitable territories.

9.2 Calibration and validation of the crop meta-models

The crop meta-models were calibrated on a training set of data from the results of the original ROIMPEL runs mostly carried out under the ACCELCEEC and ACCELERATES projects. Calibration datasets were always sampled to adequately cover the whole range of both predictors and the predicted variables, e.g. sowing date or actual yield. The sampling of the calibration dataset took into account values outside ± 1 standard deviation from the mean of each parameter (both input and output). From the interval between 1 and 2 standard deviations, two-thirds of the data were used for model calibration and of those data points above/below 3 standard deviations 90% were used for model calibration. After calibration, each meta-model was independently tested on a complementary validation set in order to assess performance accuracy.

As the training and validation datasets include over 150,000 data points, a custom-made software application for the development and training of the ANNs for the 60 meta-models (12 crops x 5 output variables) was developed. The procedure for the meta-model development is summarised in Figure 9.1. This application aids the effective selection of the most suitable ANN design (e.g. input parameter selection, number of layers and hidden layers) and, based on 100 iterations of the best design, selects the top five ANNs based on the R^2, RMSE and MBE to prepare an ensemble of ANNs. As the run-time of the meta-models increases considerably with the number of ANNs in the ensemble, five was selected as an acceptable balance between model performance and runtime. The outputs from each of these five ANNs are then combined together in order to generate a final composite projection. There is a large body of statistical theory and practical work showing the superiority of ensembles over the use of any single model (Naftaly *et al*., 1997; Sharkey, 1999; Granitto *et al*., 2005). When needed, the ANNs are combined with temperature thresholds that are designed to "prevent" a given crop growing at sites which are not considered suitable (but in which the limiting factors are not covered by the input parameters, e.g. in the case of winter wheat, the mean annual temperature must be over 4.3°C and mean temperature from April to June above 8.25°C. Using these criteria, the number of locations at which the meta-models wrongly predicted possible cropping decreased by 60-75%. In order to increase the stability and robustness of the predictions of individual yield levels (i.e. mean water- and nutrient-limited yield (Yield_Av), mean water-limited yield (Yield_POT) and mean water- and nutrient-unlimited yield (YieldPOTI)) and in particular to preserve rules that govern the relationship between the yield categories (i.e. Yield_POTI>Yield_POT>Yield_Av) in each grid cell, additional meta-models estimating the differences between YieldPOTI-YieldPOT and YieldPOT-Yield_Av were developed as well as the individual Yield_POTI, Yield_POT and Yield_Av meta-models. The final estimates of the individual yield levels are based on the integrated results of all 5 meta-models in order to preserve the relationship between individual yield levels within each grid cell.

The results of the 60 meta-models (for mean water- and nutrient-limited yield, mean water-limited yield, mean water- and nutrient-unlimited yield, sowing date and harvesting date for each of 12 crops) are summarised in the Table 9.1. The meta-models show excellent performance in predicting sowing and harvest dates, with usually more than 90% of the variability explained. The meta-models were less successful in reproducing crop yields (nutrient- and water-limited, water-limited and unlimited) but in all cases the results are considered acceptable. Overall the RMSE for the yield estimates is in most cases below 0.5 t/ha and the MBE that is close to 0 indicating that there is low/no systematic bias.

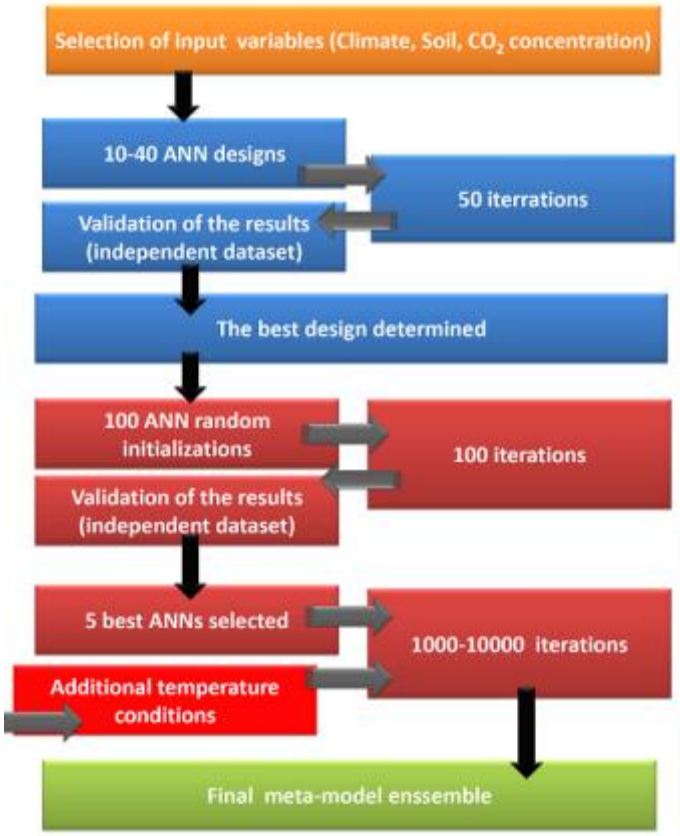


Figure 9.1: Overview of the ANN development for the crop meta-models.

9.3 Crop meta-model illustrative results

Figures 9.2 to 9.6 show complete results of the meta-models for winter wheat in comparison to the outputs of ROIMPEL. Figure 9.7 shows results from each of the best five ANNs and the ANN ensemble mean in comparison to the outputs of ROIMPEL. Given the complexity and variability of conditions across Europe, it was not possible to achieve the level of accuracy reported by Audsley *et al.* (2008) for the much smaller area of eastern England. However, the validation statistics shown in Table 9.1 are acceptable and it is likely that the uncertainty arising from using ANNs instead of ROIMPEL will be smaller than that reported, as the final IAP will use a clustering approach such that aggregation will likely lead to higher accuracy of the meta-models. The complete overview of validation maps for all yield levels and crops is available at http://www.climsave.eu/internal/Data/Yield_Metamodel/Validation_maps_baseline/

Table 9.1: Meta-model validation performance statistics for the 1980-1990 period of the ensemble mean of the five best performing artificial neural networks (ANN) for mean water- and nutrient-limited yield (Yield_Av), mean water-limited yield (Yield_POT) and mean water- and nutrient-unlimited yield (YieldPOTI), sowing date (Sowing) and harvest date (Harvesting).

Metric	Meta-model output	Winter wheat	Spring wheat	Winter barley	Spring barley	Winter oil seed rape	Potatoes	Grain maize	Sunflower	Cotton	Soybean	Grass	Olives
R^2	Yield_Av	0.81	0.74	0.82	0.75	0.86	0.93	0.86	0.85	0.86	0.86	0.81	0.99
	YieldPOT	0.78	0.72	0.75	0.76	0.84	0.9	0.82	0.81	0.91	0.88	0.8	0.99
	YieldPOTI	0.88	0.83	0.87	0.86	0.95	0.96	0.94	0.97	0.91	0.98	0.98	0.99
	Sowing	0.99	0.98	0.99	0.98	1	0.99	0.95	0.99	1	0.96	0.75	0.99
	Harvesting	0.99	0.99	0.98	0.99	0.99	0.99	0.72	0.97	0.82	0.9	0.83	1
RMSE	Yield_Av	0.55	0.53	0.5	0.48	0.45	1.7	0.55	0.12	0.22	0.43	0.43	0.05
	YieldPOT	1.02	1.02	1.01	1.06	0.82	3.01	1.06	0.37	0.42	0.5	1.59	0.05
	YieldPOTI	0.88	0.93	0.86	0.94	0.74	3.53	0.85	0.19	0.89	0.32	1.25	0.05
	Sowing	1.94	3.7	2.11	3.68	0.86	2.82	2.38	1.32	2.7	1.69	8.11	2.55
	Harvesting	2.07	1.86	2.13	2.13	3.56	3.53	8.22	4.96	11.61	3.18	8.21	1.54
MBE	Yield_Av	0	0	0	0	0	0.02	0.01	0	0	-0.01	0.01	0
	YieldPOT	-0.01	0	-0.01	-0.01	-0.01	0.03	0.01	0	0	-0.01	0	0
	YieldPOTI	-0.01	0.01	0	0	0	0.04	0.03	0	0.02	0	-0.02	0
	Sowing	0.01	-0.09	0	0	0	0.12	-0.01	0.02	0.01	0.02	-0.04	0.01
	Harvesting	-0.02	0.12	-0.01	0	-0.12	0.02	0.24	-0.07	0.03	-0.02	0	-0.01

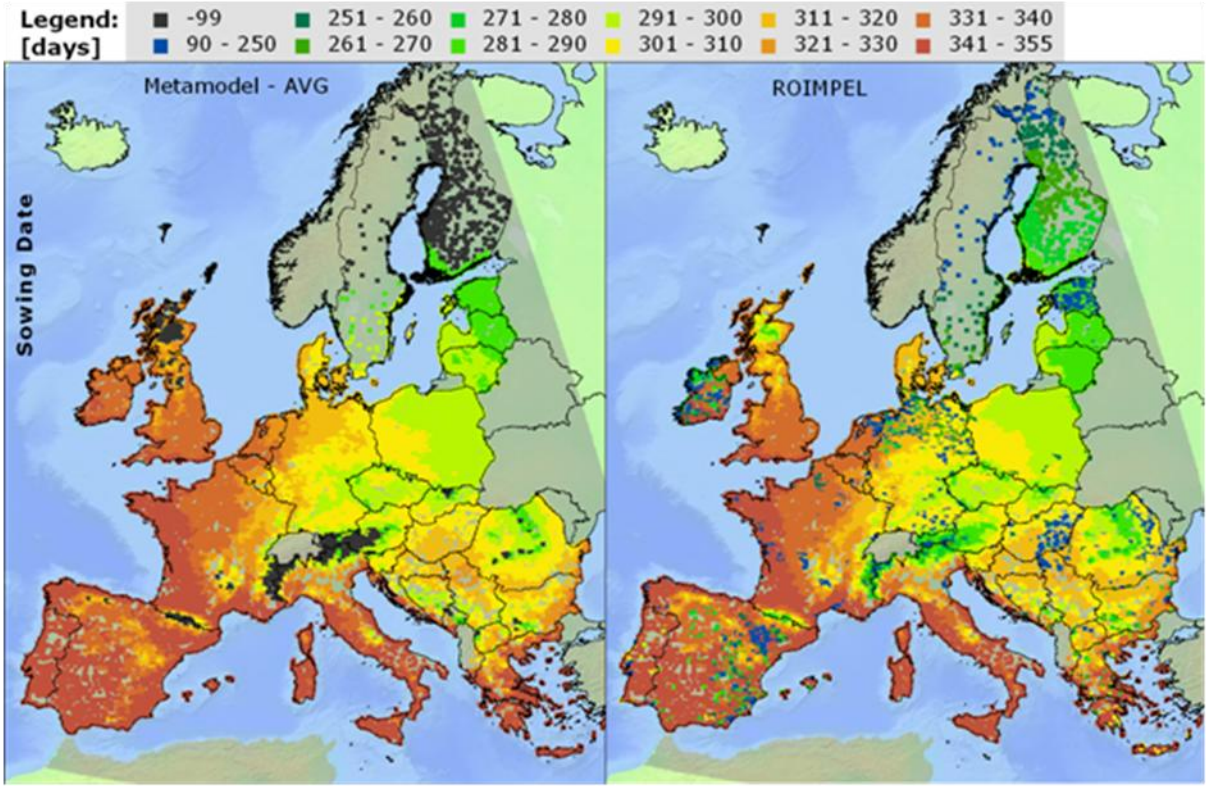


Figure 9.2: Comparison of sowing date (Julian Day – 1st Jan = Day 1) for winter wheat as predicted by (left) the mean of the meta-model ANN ensemble and (right) ROIMPEL.

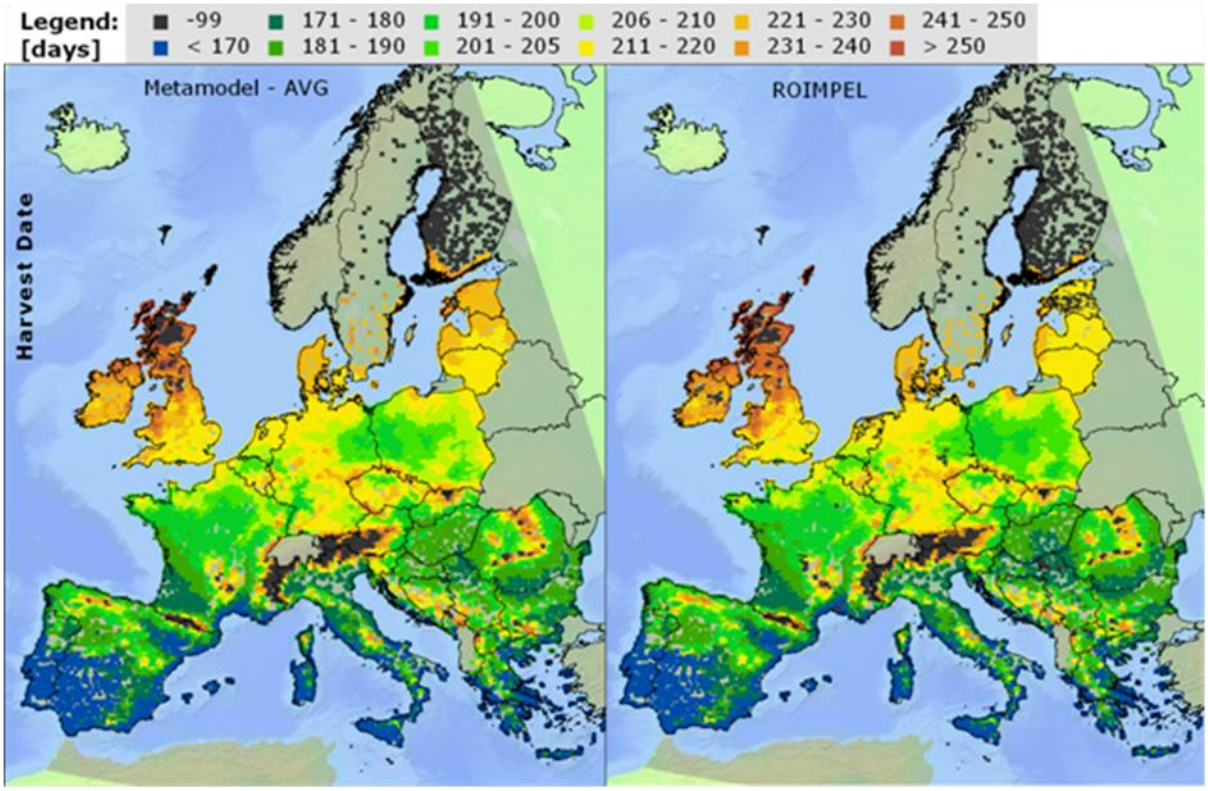


Figure 9.3: Comparison of harvest date for winter wheat as predicted by (left) the mean of the meta-model ANN ensemble and (right) ROIMPEL.

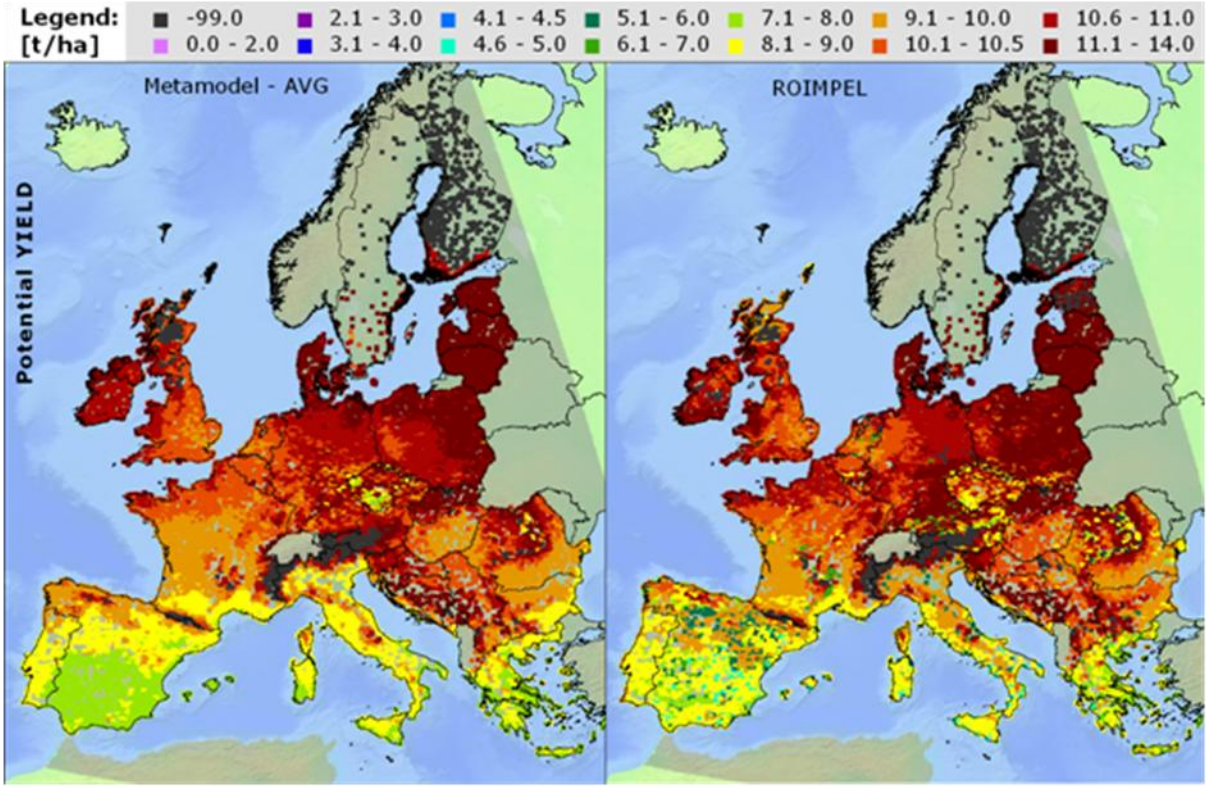


Figure 9.4: Comparison of potential (water- and nutrient-unlimited) yield for winter wheat as predicted by (left) the mean of the meta-model ANN ensemble and (right) ROIMPEL.

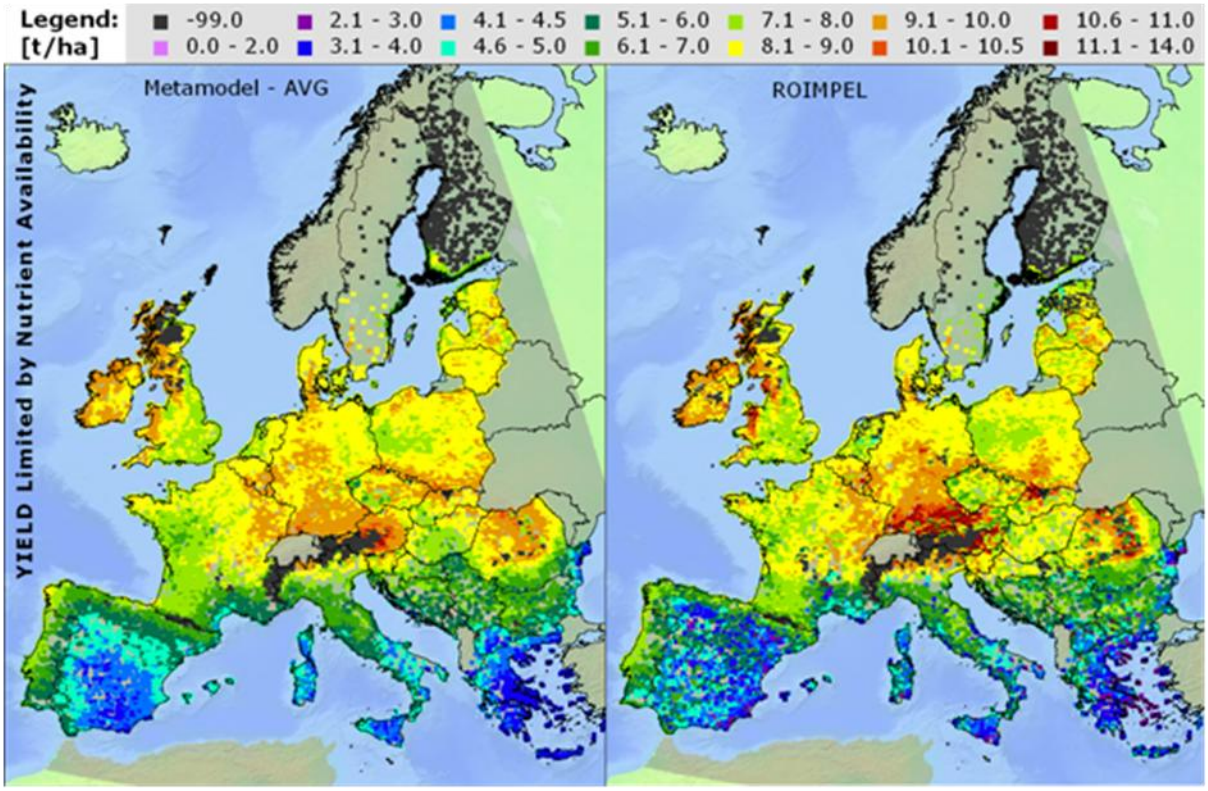


Figure 9.5: Comparison of yields limited by nutrient availability for winter wheat as predicted by (left) the mean of the meta-model ANN ensemble and (right) ROIMPEL.

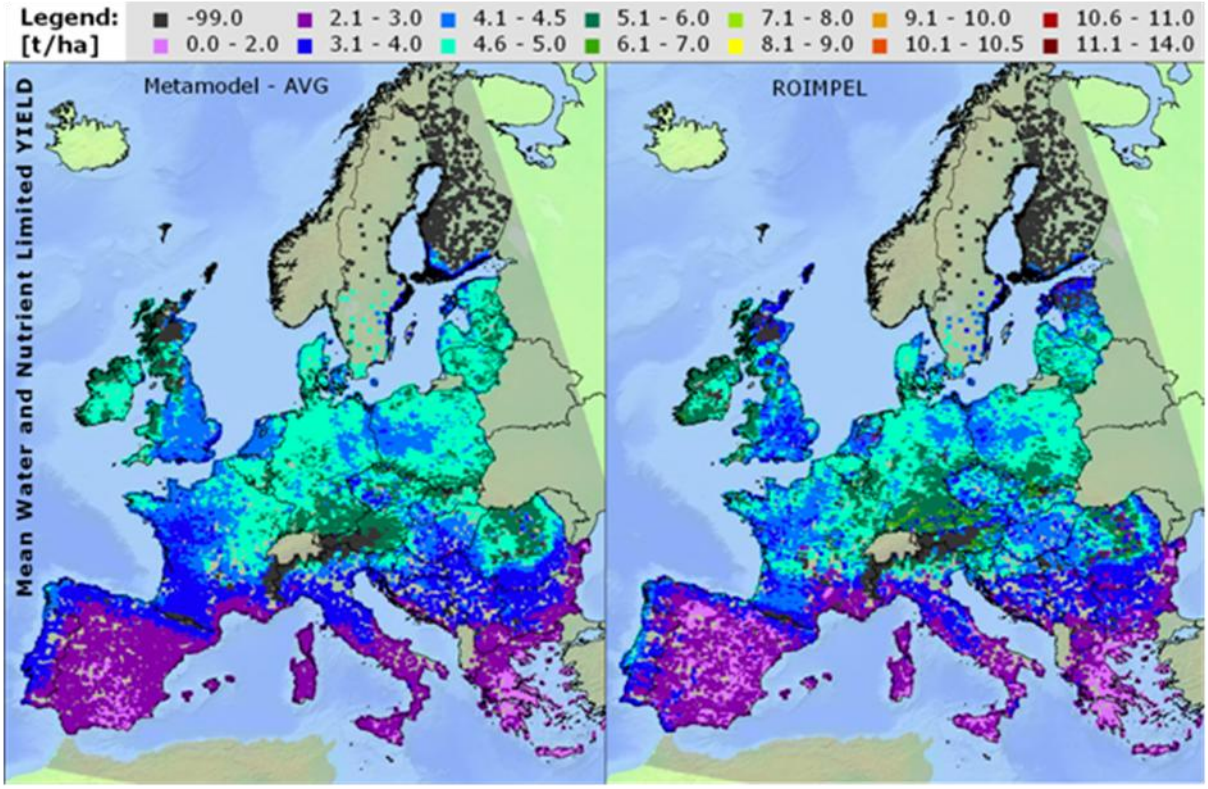


Figure 9.6: Comparison of water- and nutrient-limited yield for winter wheat as predicted by (left) the mean of the meta-model ANN ensemble and (right) ROIMPEL.

9.4 Integrating the crop meta-models with the other sectoral meta-models

In the IAP design the crop meta-model outputs are not used directly but only in association with the agricultural land use or farm model (SFARMOD; Section 10). Only after evaluation of the farm model gross margins which, given the crop yields, can be calculated from the crop prices, subsidies and variable costs, is it possible to estimate crop production in a particular area. Interaction between individual sectors and the crop meta-models is therefore provided by the SFARMOD meta-model and discussed in Section 10.

The statistical analysis, that is summarised in Table 9.1, was in reality much more detailed and included careful assessment of numerous indicators of meta-model performance, both for the training as well as validation subsets of data. It included testing of the meta-models across the multiple climate change scenarios and comparisons with the original model. Given the large number of meta-models, only an overview is available within this Deliverable with complete data for all crops, individual ANN runs and variables available at http://www.climsave.eu/internal/Data/Yield_Metamodel/Statistics/.

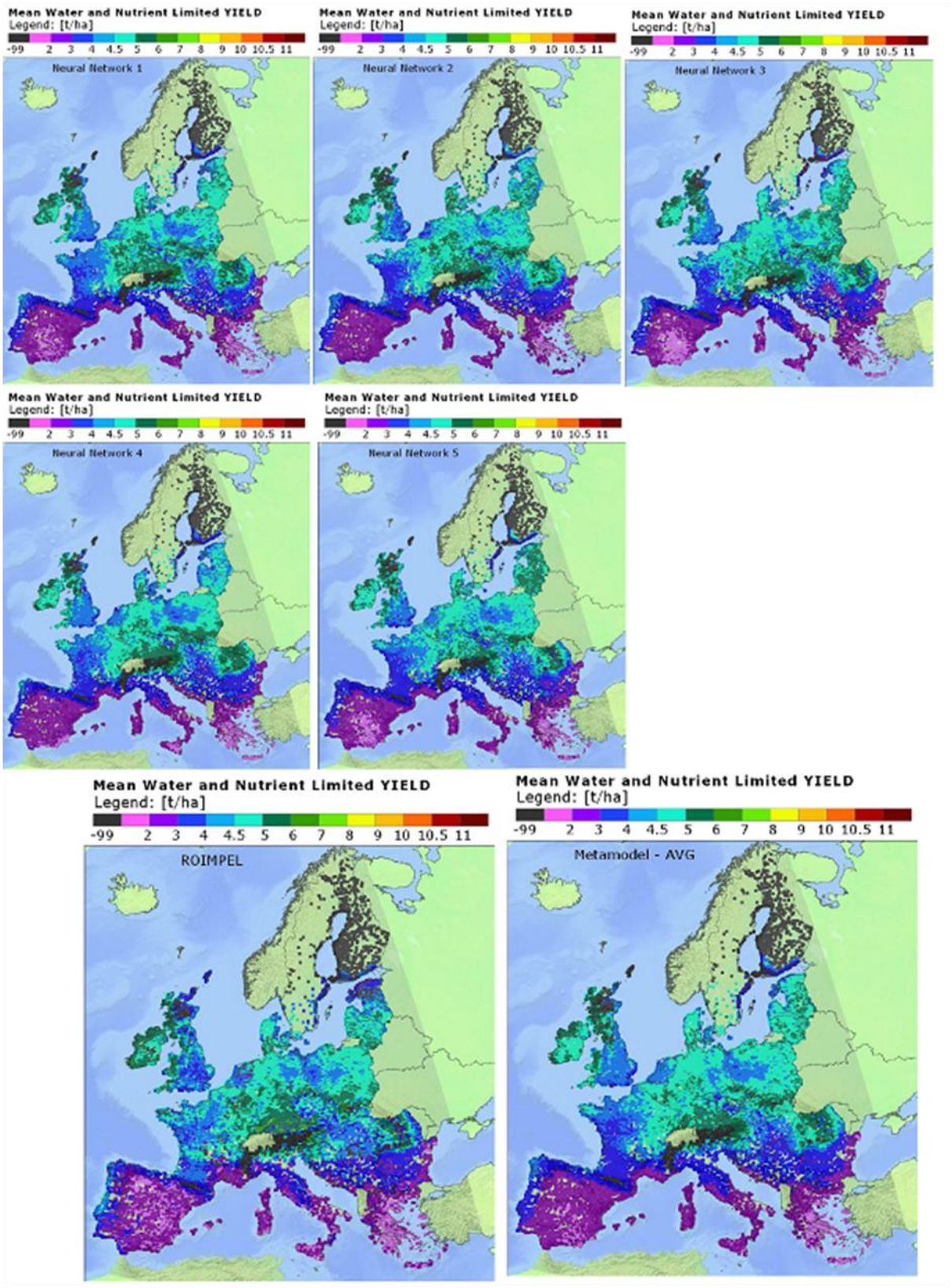


Figure 9.7: Comparison of winter wheat yields limited by the nutrient- and water-availability as predicted by ROIMPEL and by the five best ANNs and their mean used in the final meta-model.

In order to assess in detail the sensitivity of all three yield levels to changes in all input parameters, a thorough sensitivity analysis was carried out on the response of the original model (ROIMPEL) and the meta-models which is available at http://www.climsave.eu/internal/Data/Yield_Metamodel/Sensitivity_Charts/. Figure 9.8 shows examples of the sensitivity response of winter wheat to changes in carbon dioxide concentrations (Figure 9.8 left) and to mean temperature during the period from April to June (Figure 9.8. right).

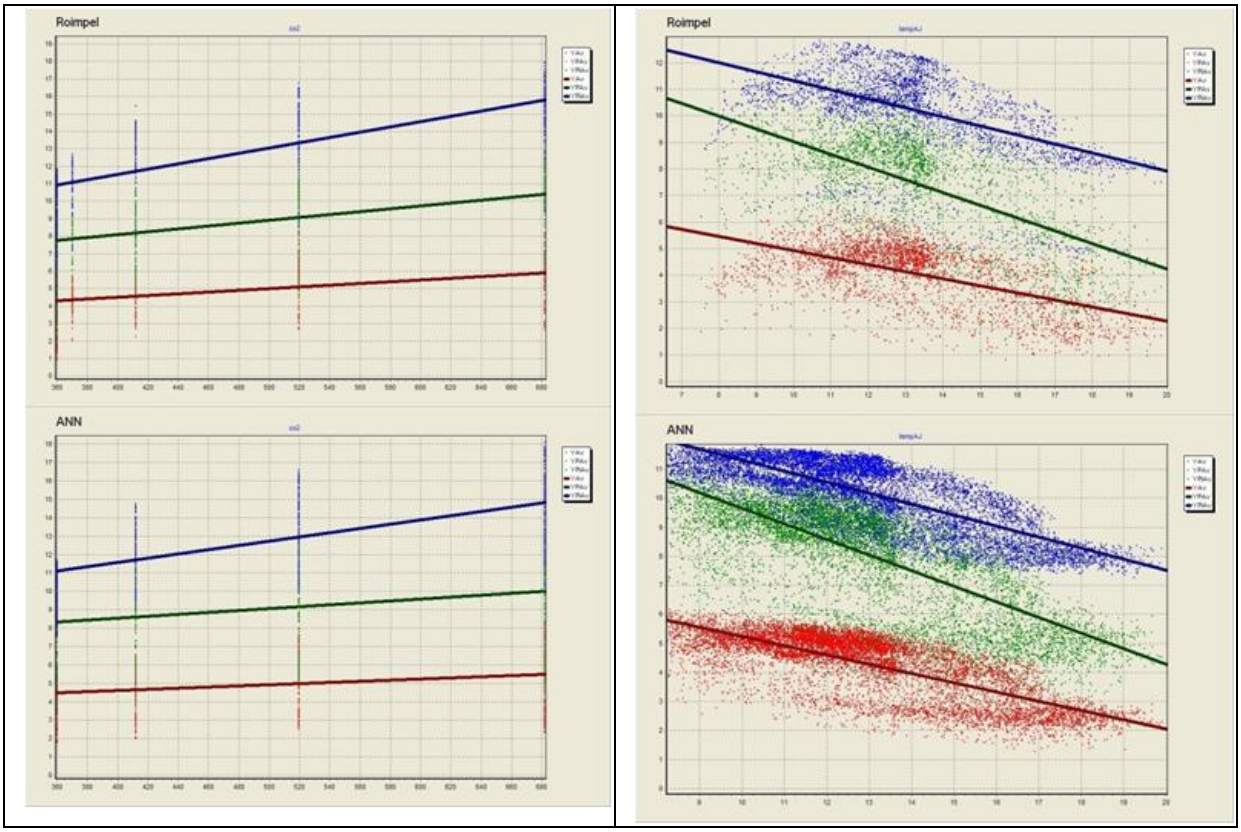


Figure 9.8: Response of the original winter wheat yield model (upper figures) and the meta-model (bottom figures) to changing gradient of CO_2 and mean temperature from April to June for all three yield levels (Yield_Av – Red, Yield_POT – Green and Yield_POTI – Blue).

The sensitivity analysis revealed that the responses of the meta-models to variation in input parameters (used as input variables to the meta-models) are very similar to the original model for all crops and yield levels. This might be viewed as a considerable success given the complexity of the meta-models and sheer number of input parameters.

9.5 References

Alexandrov, V. (2006). Climate variability and change and related impacts on ecosystems in southeast and central Europe as well as in southeast USA, Bulgarian Academy of Sciences, Doctoral Thesis, p. 199.

Audsley, E., Pearn, K.R., Simota, C., Cojocaru, G., Koutsidou, E., Rounsevell, M.D.A., Trnka, M. & Alexandrov, V. (2006). What can scenario modelling tell us about future

European scale land use, and what not? Environmental Science & Policy , 9, 148-162.
Audsley, E., Pearn, K.R., Harrison, P.A. & Berry, P.M. (2008). The impact of future socio-economic and climate changes on agricultural land use and the wider environment in East Anglia and North West England using a metamodel system. Climatic Change 90: 57-88.
Granitto, P.M., Verdes, P.F. & Ceccatto, H.A. (2005). Neural network ensembles, evaluation of aggregation algorithms. Artificial Intelligence, 163, 139-162.
Henseler M., Wirsig A., Herrman S., Krimly T. & Dabbert S. (2009). Modeling the impact of global change on regional agricultural land use through an activity-based non-linear programming approach. Agricultural Systems 100 (2009) 31-42.
Mayr, T.R., Rounsevell, M.D.A., Loveland, P.J. & Simota, C. (1996). Agro-climatic change and European suitability: regional modeling at monthly time steps. International Agrophysics 10 (3), 155-170.
Naftaly, U., Intrator, I. & Horn, D. (1997). Optimal ensemble averaging of neural networks. Network: Computation in Neural Systems, 8, 283-296.
Rounsevell, M.D.A., Annetts, J.E., Audsley, E., Mayr, T. & Reginster, I. (2003). Modelling the spatial distribution of agricultural land use at the regional scale. Agric Ecosyst Environ, 95(2–3): 465–479.
Sharkey, A.J.C. (Ed.). (1999). Combining Artificial Neural Nets, Springer, London.

10. Development and validation of the SFARMOD rural land use allocation meta-model

Daniel Sandars and Eric Audsley
Environmental Science and Technology Department, Cranfield University, UK

10.1 Introduction

SFARMOD is the generic name given within CLIMSAVE to the routine for selecting rural (non-urban) land use. The concept is that the profitability of the competing uses for land is estimated using a general linear programming model and it is assumed that in the long-term the use that is most profitable will be the one selected. This procedure was used in Audsley *et al.* (2006). There are basically three land uses: agriculture, forestry or unused. If profit is above a threshold it will be used for intensive agriculture, which can be either arable or grassland (dairy) cropping. Above a second threshold land is extensive grassland, which is considered as grazing sheep or beef. Below this land is described as abandoned, which could also be forest but could equally be simply unusable for agriculture such as bare rock.

The full model used to develop the SFARMOD meta-model within CLIMSAVE is the SFARMOD optimising linear programme (hereafter referred to as the SFARMOD-LP) of whole farm planning, based on profit maximisation subject to the constraints of soil, precipitation and sound agronomic practice. This calculates the profitability of arable and intensive grass cropping on the land. Further details of the SFARMOD-LP can be found in Audsley (1981), Holman *et al.* (2005) and Annetts & Audsley (2002).

10.2 Description of the SFARMOD-LP – Silsoe Whole Farm Model

SFARMOD-LP (also known as the Silsoe Whole Farm Model) is a mechanistic farm-based optimising linear programming model of long-term strategic agricultural land use. Crops are defined by their gross margin, the amount and timing of the labour and machinery they require, restrictions on crop rotations, and their sowing and harvest dates. Gross margins are determined from the yield, which is a function of soil and climate, plus in some cases the amount of irrigation. Soil workability which determines labour and machinery availability is a function of soil and climate. In addition farmers have uncertain future knowledge of actual prices and yields, and this is simulated in the full model by ten combinations of yields and prices from which the average cropping represents the expected land use. The decision variables are crop areas, crop rotations, amount of labour and machinery, and operational timing within its feasible period which determine the farm profit given this soil and climate.

The inputs to the full SFARMOD-LP model are:

- Soil type, as an index reflecting the trafficability and available water capacity ranging from 2.5 on heavy land to 0.5 on sand, by soil polygon covering Europe.
- >30 year mean annual precipitation and evapotranspiration.
- Gross margins determined by:
    - Prices and support regime rules.
    - Yields
    - Input costs.

- Harvest and sowing date - also if it is feasible to grow the crop in terms of it being able to reach maturity. The model decides whether it is economic to grow the crop.

The outputs produced are:

- Net profit, at the farm level.
- Cropping as percentage of area.
- Environmental burdens: nitrate leaching, pesticide use, nitrogen use.
- Measures for biodiversity indicators required by the SPECIES model in Section 13, such as over-winter stubble and use of pesticides.

Yields under soil and climate conditions are given by the yield model described in section 9. This provides:

- Yield under normal fertiliser input and no irrigation.
- Yield with no nutrient limitation and no irrigation
- Yield with no nutrient or water limitation.
- Sowing and harvest dates.

Tests of the differences between these yields and yields reported by Eurostat resulted in SFARMOD-LP using a calibration procedure to adjust the nutrient-unlimited yields to allow for reduced inputs in CEEC countries, to allow for the effect of disease pressure on yields in high rainfall situations and to update the yields to modern levels. These are applied at a NUTS2 scale.

The SFARMOD-LP model simulates the impact of, and adaptation to, climate and socio-economic change by modifying input parameters. Because changes to yields and suitability can lead to large increases in the area of some crops, it is necessary to modify prices in response to the need to meet demand for a commodity. For example, if climate change meant that northern Europe could profitably grow sunflowers, there would be oversupply unless the price was lower. Thus, the model reduces prices until demand is matched.

10.3 Development of the SFARMOD meta-model

The objective of the SFARMOD meta-model is to simulate the behaviour of the full SFARMOD-LP model described above, as applied to all soil-climate combinations. The procedure must allocate land to categories of land use and calculate the total expected production of each type of crop outputs for each scenario. It must be able to respond to scenario and adaptation options. The following options have been defined:

- Increase in crop yields due to technology.
- Reduction in labour and machinery due to technology.
- Reduction in irrigation needed due to technology.
- Increase in arable land set-aside for the environment.
- Increase in arable land used to produce bio-energy.
- Reduction in use of nitrogen for diffuse pollution.
- Reduction in the consumption of ruminant meat.
- Reduction in the consumption of non-ruminant meat.
- Change in imports of agricultural products.

- Increase in cost of inputs due to oil price.
- Increase in cost of labour due to wealth (GDP).
- Increase in demand due to population growth.
- Reduction in water available for irrigation.

10.3.1 *Data available*

CLIMSAVE operates on 10' x 10' grids and all data input and output are provided on this common framework for all models. There are the 23,871 grid squares defined.

1. The soil data file is derived from an inter-section of the European soil map with the CLIMSAVE 10' grid. There are 143,955 soil type-grid combinations, with up to 47 different soil types (officially known as Soil Typological Units) within each grid square, and a total of 5,107 different soil types. This needed to be simplified to facilitate efficient application of the meta-model:
   a. The soil attribute database for each soil type was reduced to those parameters required by the meta-models for crop yield, forestry and SFARMOD such as Available Water Capacity (AWC) at four suctions from Saturation to Permanent Wilting Point, stoniness, and soil texture. On this basis many soil types are identical and the total is reduced to 582 distinct soils.
   b. A proportion of each grid can be identified as urban or not possible for agro-forestry using the CORINE database (e.g. the land use category Bare Rock). These categories were used as far as possible to eliminate the zero soil or very shallow soil types. Even so there was a small proportion of land with no soil data which was not identified as such by CORINE.
   c. A clustering procedure was applied to the soil data (Figure 10.1) to produce 182 similar soil types, with the procedure aiming to not cluster soil types of over 5,000,000 ha unless they are very similar. This used the Akaike Information Criteria (AIC) optimum for loss of information. However, note that it is actually possible to cluster more or less tightly depending on run time.

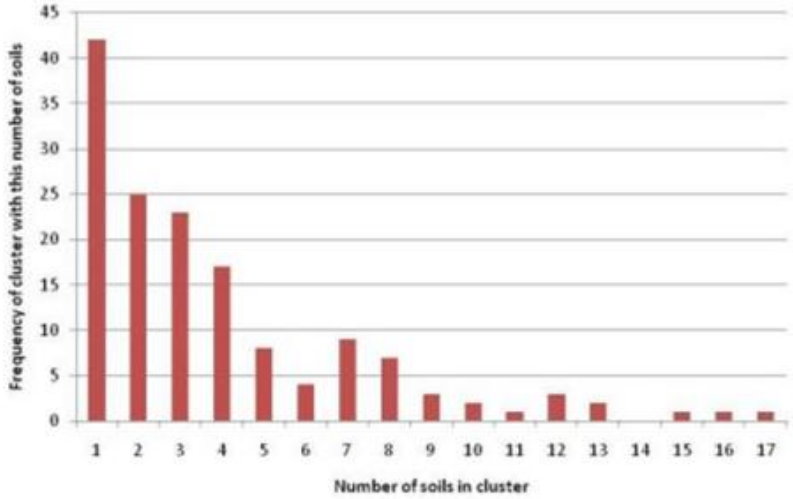


Figure 10.1: Distribution of the number of soils per soil cluster.

2. A similar clustering procedure was also applied to the baseline climate data for the 23,871 grid squares, assumed uniform over a grid, which produced 170 clusters (Figure 10.2). The clustering was based on grouping grid cells with similar average summer and

average winter temperature, potential evapotranspiration and precipitation, and days (from 1st January) until average temperature > 0°C and 6°C. As with soils it is possible to use more or less clusters than the AIC optimum.

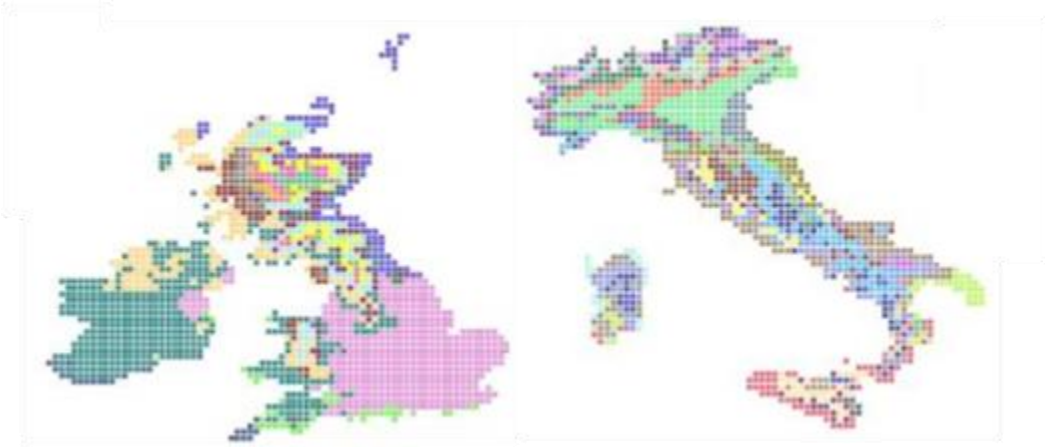


Figure 10.2: Example of UK and Italy meteorological clusters.

It is assumed that the climate change data has the same clustering. Analysis of the clusters showed that there was very little difference if the change was applied to all grids within a cluster. Cluster climate files were thus produced for each grid climate file.

3. Each grid is allocated a water basin by the WGMM meta-model.

4. Each grid is allocated a climate type (e.g. Maritime) by the forestry meta-model.

5. Combining climate and soil clusters and allowing for water basins, forest climate regions and the fact that calibration factors between EEC and CEEC countries can be very different, there are 16,058 distinct climate-soil clusters, a factor of 10 reduction, since not all soils occur in all climate regions. Due to the diversity of soil types within a single grid square, there are multiple climate-soil clusters within a grid square (but all having the same climate). The crop yield meta-models (Section 9) and forestry meta-model (Section 6) produce data on the same soil-climate clustering (not per grid).

10.3.2 *Construction of the SFARMOD meta-model*

The approach taken to develop the meta-model is to use the full SFARMOD-LP to systematically populate the input parameter space and then to create a meta-model that relates the input parameters to the SFARMOD-LP outputs. In order to fully cover the parameter input space, SFARMOD-LP was run with 20,000 randomly selected sets of input data:

- Gross margins for each crop.
- Net precipitation used in the SFARMOD-LP workability formula.
- Summer temperature which modifies the harvest and sowing dates for each crop.

These results were then used to create the meta-model. A number of approaches were taken for the meta-modelling, but the most reliably successful proved to be a regression rather than a neural network approach. The regression is broken into steps to allow the effect of scenario variables to be included. The steps estimate first the percentage of the area of each crop, then

the costs of dairy cows (concentrates) then the fixed costs of labour and machinery. The profitability of this soil and climate is the difference.

The procedure derives a regression model for each crop using various combinations of input parameters. The input parameters to the regressions were:

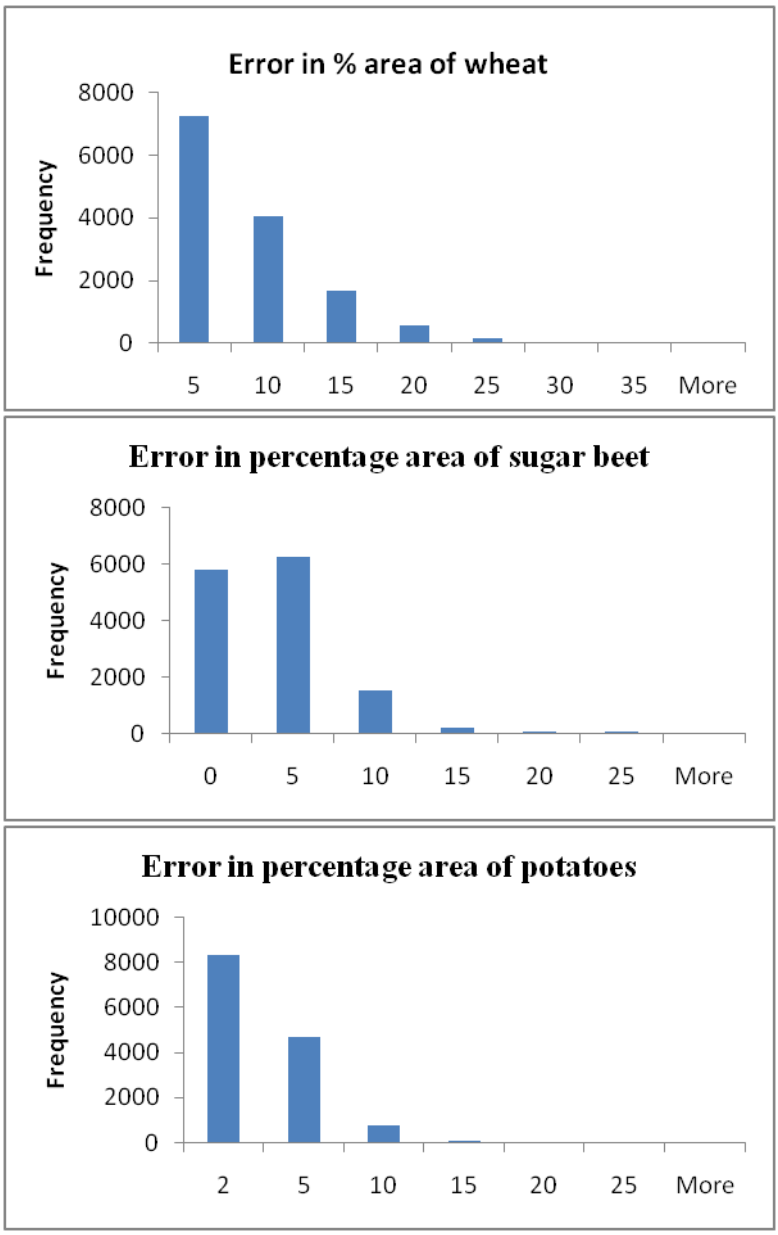
1. The gross margin of the target crop (in €1000), or DM yield for forage crops.
2. The effective precipitation measure used by the LP to calculate workability.
3. The distance of the sowing date from the start of the year (fortnights).
4. The distance of the harvest date from May 1st.
5. The soil type on a scale of 1 (sand) to 9 (heavy clay).
6. If the latitude is greater than 4 degrees = 1 else 0.
7. A measure of summer temperature on a scale of 0 to 1.
8. The ratio of the gross margin of other crops to the target crop omitting spring versions of wheat, barley and rape which are very strongly correlated to their winter version.
9. The ratio of the target crop to winter barley and of winter wheat to winter barley.
10. Other crop specific parameters, e.g. sugarbeet to oilseed rape gross margin ratio.

With squares and inverses of the above variables, there are 23 or 24 input values for each crop regression. Examples of the errors from three fits are shown for percentage crop areas for wheat, sugar beet and potatoes in Figure 10.3. Where points are a very bad fit, these were examined and found to be cases where extreme gross margins existed. The total percentage of crops is then scaled to be 100%.

Given the crop areas, number of dairy cows and their gross margins, the net income of the farm is known. The capital and labour costs are higher where the land is heavier and precipitation is high due to fewer workable hours, which is exacerbated where the crop is winter sown or harvested later in the year. The fixed cost regressions are shown in Table 3. The profit is then the net income minus the dairy cow cost and fixed costs.

Where the profit is above a threshold (set at €350/ha) this soil/climate is deemed to be used for intensive agriculture (either arable or dairy farming). Otherwise the land can be used extensively (for livestock grazing) and the profit is then re-calculated without the arable crops. This profit is then compared with the profit from managed forests (Section 6). If the resulting profit is greater than a second threshold (set at €120/ha) then this soil/climate is used for forest or extensive grazing. Otherwise the land is not used and if the Net Primary Productivity of unmanaged forests (section 6) is positive the land is deemed to be forest, otherwise the land is abandoned and un-forested. Other output parameters are determined from the cropping, in particular irrigation requirement by basin.

Thus, we have a rapid approximation to the linear programming model. Given a new socio-economic and climate scenario this speed enables crop prices to be iteratively adjusted to meet a demand as well as the cost of water per basin to meet irrigation availability. The yield of irrigated crops is calculated as the optimum level of irrigation given the yield with and without irrigation and the crop and water price. Limited water increases the water price, reduces the gross margin and, hence, the water used by the crop. Note that given fixed demand it is likely to *increase* the area of the crop in total. Furthermore, if demand is very high (e.g. due to a large population increase) and large restrictions (e.g. a large percentage of set-aside and no imports), it is possible that no prices can be found which meet demand.



Note: The sugar beet area must be in the range 0-25% and zero error indicates cases where the fit and target was 0 or 25.

Figure 10.3: Comparison of the performance of the SFARMOD meta-model with the results for the full SFARMOD-LP for the percentage of the cluster allocated to (top) wheat, (middle) sugar beet and (bottom) potatoes.

10.3.3 *Validation of the SFARMOD meta-model*

The fit of the SFARMOD meta-model to the LP model is illustrated by the results in Figure 10.4. The fit of the meta-model to the Eurostat data is shown in Figure 10.5 for arable and Figure 10.6 for intensive grassland. The problem with grassland is that extensive and intensive are both grassland so the Eurostat estimate was based on the difference between agricultural area and arable area, which is some cases is negative.

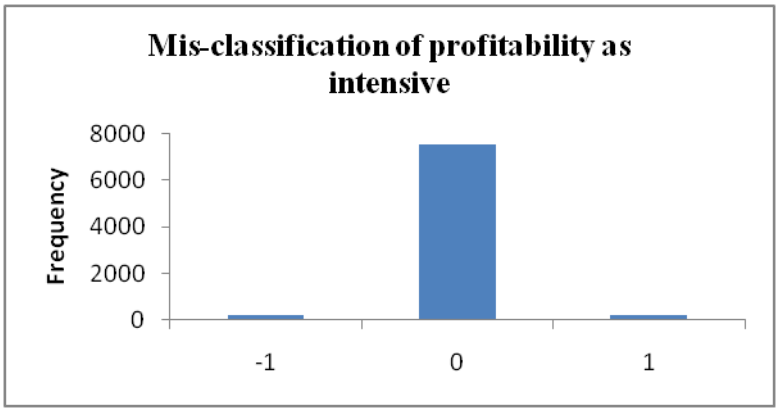


Figure 10.4: Mis-classification of land by the SFARMOD meta-model versus the full SFARMOD-LP model.

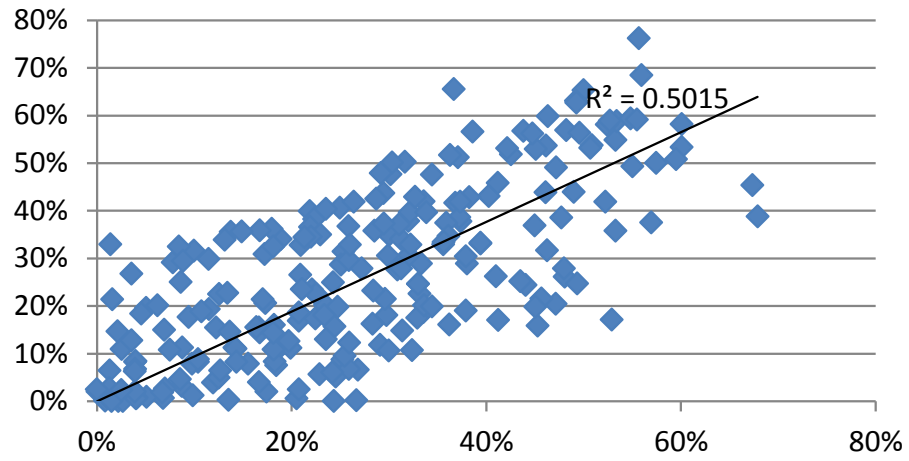


Figure 10.5: Fit of SFARMOD meta-model arable land classification to Eurostat data (% grid).

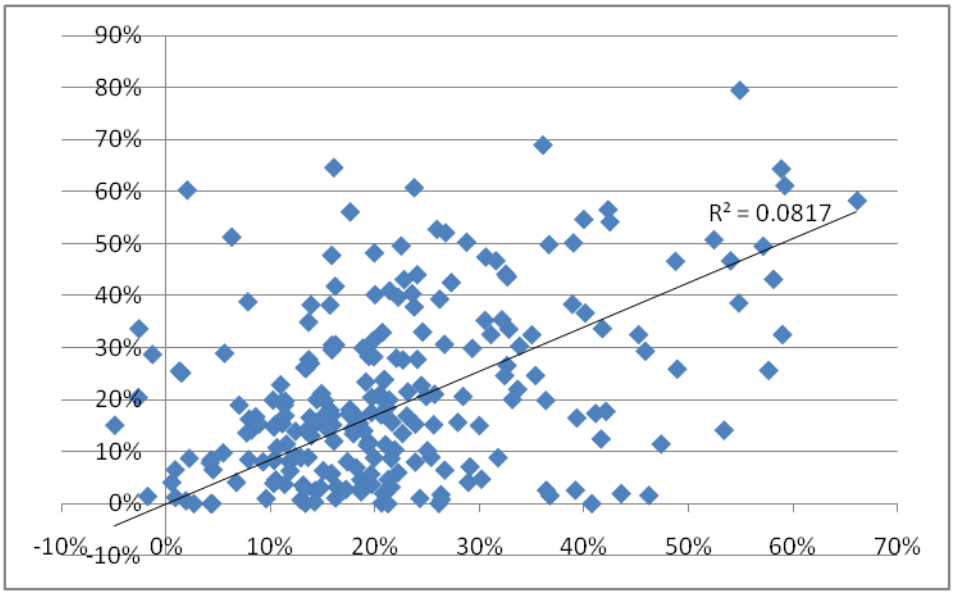


Figure 10.6: Fit of SFARMOD meta-model grassland classification to Eurostat data (% of grid cell).

10.4 Application of the SFARMOD meta-model within the IAP and integration with the other sectoral meta-models

There are a series of steps to the running of the SFARMOD meta-model:

1. The potential agricultural area of the grid cells is reduced pro-rata by the proportion of the increase in urbanisation in the grid given by the RUG meta-model (Section 5).
2. Land which is flooded (input from the CFFlood meta-model; Section 7) has two categories: land which is flooded frequently and unusable for agriculture and which is removed pro-rata from each grid cell, and land which is flooded infrequently and is only suitable for grassland.
3. For each soil-climate cluster:
   a. Calculate crop yields and suitability from the Yield meta-model (section 9).
   b. Calculate the gross margin of each crop from yields and scenario inputs including irrigation; G = (P*Y*F) – (C*M) – (I*E*W)
      where
      P is the price of the commodity, e.g. cereal; Y is the yield; F is the NUTS2 scaling factor;
      C is the input costs of the crop which varies with yield (low yields require low nitrogen input); M is the scenario factor for input costs (e.g. fossil fuel prices give higher fertilizer prices) (crop invariant); I is the amount of irrigation required by the crop (determined as the optimum); E is the scenario factor for efficiency of irrigation (crop invariant); and
      W is the price of irrigation water for the basin.
4. Apply the meta-model to calculate the percentage under each crop.
5. Apply the meta-model to calculate the profit of the cluster.
6. Compare the profit to the intensive threshold €350/ha to define land use and crop allocation.
7. If not intensive: Apply the meta-model to determine profit under grass.
8. Apply the forest meta-model (section 6) to determine profit under managed forest.
9. Compare the best of these profits to the extensive threshold €120/ha to define land use.
10. If not extensive: apply the forest meta-model (section 6) to determine NPP under unmanaged forest.
11. If NPP is positive then define land use as forest, else land use is abandoned.
12. Apply clusters to grids. A grid can have intensive, extensive, managed and unmanaged forest and abandoned (as well as urban).
13. The percentage of a grid cell which is classified as a Protected Areas is defined by the SPECIES model (Section 13) for each land use type. Where these proportions are not satisfied (plus the forced flooded grassland), land is moved down from the next available higher land use (e.g. intensive is forced to be forest). The land with the worst soils (defined by Available Water Capacity) is assumed to be the protected area.
14. Calculate total production of commodities, and total irrigation.
15. Compare the supply of commodities (cereal, carbohydrate, oil, soya, cotton, milk and meat) with demand, and compare the irrigation required by the basin with water availability provided by the WGMM meta-model (Section 4). Iterate prices to meet demand and irrigation to meet water supply.

10.4 Application of the SFARMOD meta-model

Figures 10.7-10.10 show example spatial output from the SFARMOD meta-model. Figure 10.7 shows the intensive agriculture and arable land in Europe for the baseline conditions. Figure 10.8 shows the impact of a climate and socio-economic scenario for the 2050s on these. There has been a visible shift of intensive and arable cropping to the north. Figure 10.9 shows the effect of using a different climate model (GFCM21 and HADGEM instead of CSMK3). Finally, Figure 10.10 shows the baseline for Scotland and the MacTopia socio-economic scenario for the 2050s for intensive agriculture (the majority of Scotland is extensive, i.e. beef and sheep grazing). The large reduction is due to the reduced demand and large increase in productivity.

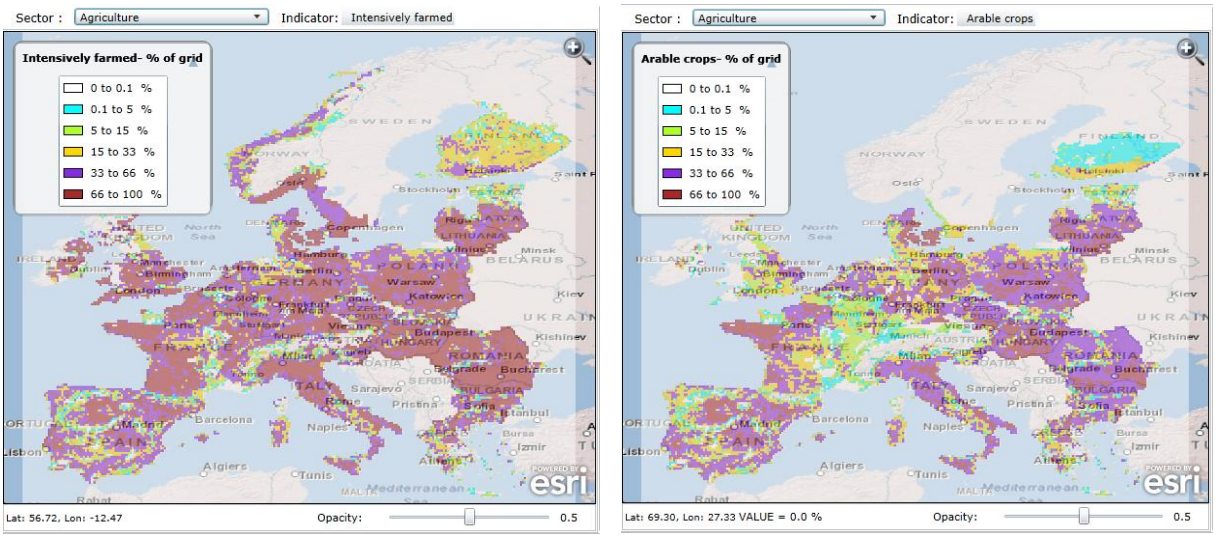


Figure 10.7: Intensive agriculture and arable land in Europe for the baseline conditions.

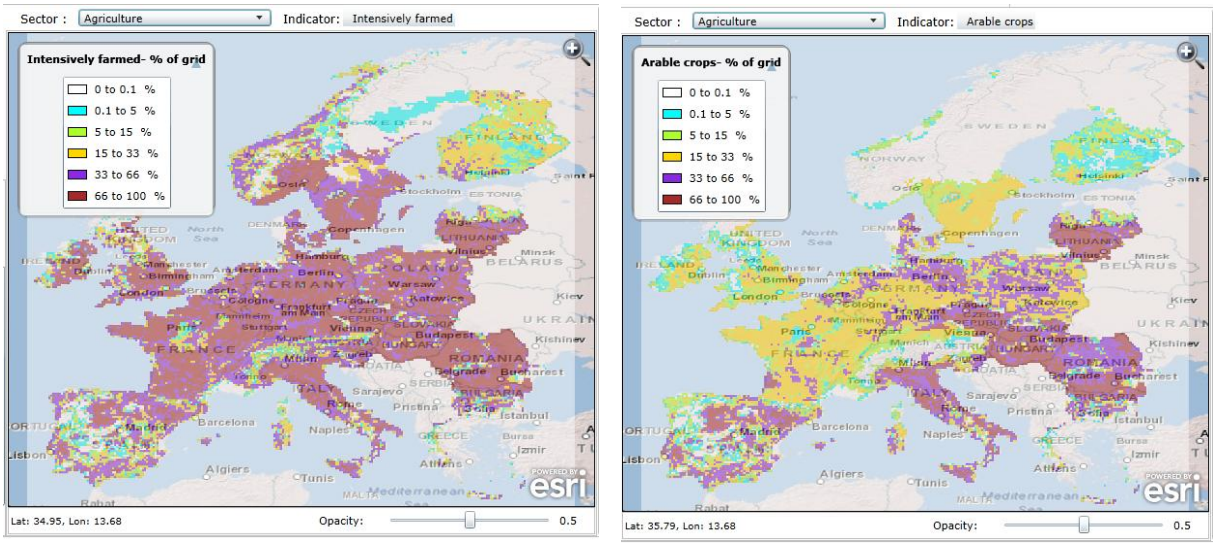


Figure 10.8: Intensive agriculture and arable land in Europe for the 2050s under A1 emissions scenario, CSMK3 climate model and the We Are the World socio-economic scenario.

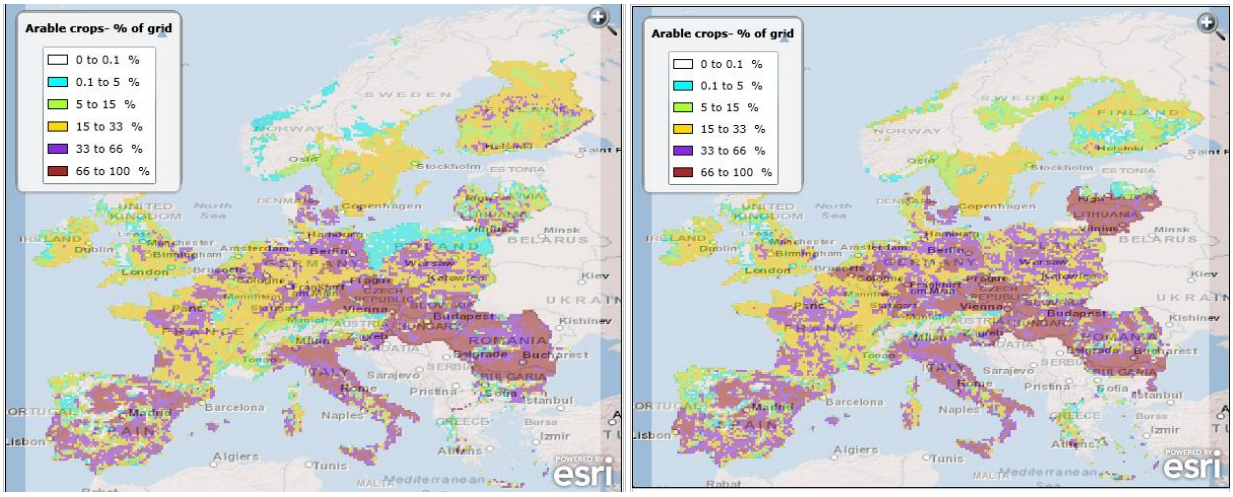


Figure 10.9: The effect of using a different climate model on simulated arable crop proportions in the 2050s using (left) HadGEM and (right) GFCM21 (using the same emissions, climate sensitivity and socio-economic scenario as Figure 10.8).

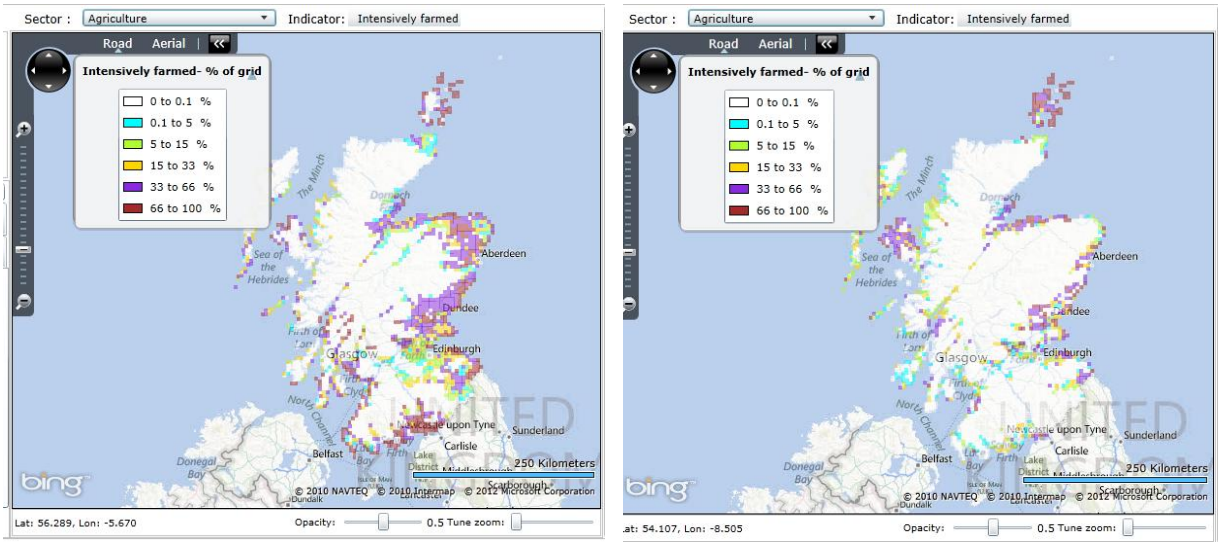


Figure 10.10: Intensive agriculture (the majority of Scotland is extensive, i.e. beef and sheep grazing) in Scotland for the (left) baseline and (right) 2050s under the MacTopia socio-economic scenario combined with a medium climate and low emissions scenario.

10.5 References

Annetts, J.E. & Audsley, E. (2002). Multiple Objective Linear Programming for Environmental Farm Planning. The Journal of the Operational Research Society. 53(9) 933-943.

Audsley, E. (1981). An arable farm model to evaluate the commercial viability of new machines or techniques. Journal of Agricultural Engineering Research, 26 (2) 135-149.

Audsley E., Pearn, K.R., Simota, C., Cojocaru, G., Koutsidou, E., Rounsevell, M.D.A., Trnka, M. & Alexandrov, V. (2006). What can scenario modelling tell us about future European scale land use, and what not? Environmental science & policy, 9, 148-162.

Holman, I.P., Rounsevell, M.D.A., Shackley, S., Harrison, P.A., Nicholls, R.J., Berry, P.M. & Audsley, E. (2005). A regional, multi-sectoral and integrated assessment of the impacts of climate and socio-economic change in the UK. Climatic Change, 71(1), 9-41.

Development and validation of the pest meta-models

Miroslav Trnka
Institute of Agrosystems and Bioclimatology, Mendel University, Brno, Czech Republic

11.1 Pest occurrence model description

The pest meta-models were designed based on the outputs of the climate-matching software program CLIMEX that estimates the geographical distribution of a species based on the climate conditions of a given location. CLIMEX is based on the assumption that the climate suitability for a given species can be derived from knowing its present area of occurrence. In other words, CLIMEX attempts to mimic the mechanisms that limit species' geographical distributions and determine their seasonal phenology. CLIMEX is a climate-rather than weather-driven modelling program that is designed to provide insights into species' requirements for climate, as expressed by their geographical distribution, seasonal phenology and relative abundance. This approach suits the aim of showing climate induced and robust shifts in pest species' distributions under future climate(s). CLIMEX is based on the premise that it is possible to define climates that are conducive to the generation of particular weather patterns, which directly affect populations on a short time-scale (Sutherst *et al.*, 2000). The software has been used extensively in the fields of biological control, climate change and pest risk assessment with positive results in many countries.

Knowing the climatological requirements of a given species allows assessment of the suitability of a particular area for population growth and to determine the stress induced by unsuitable climate conditions. These are expressed in terms of the ecoclimatic index (EI), which describes the overall suitability of climate conditions for the establishment and long-term presence of a pest population at a given location:

$$EI = GIA \times SI \times SX,$$

where GIA is the annual growth index describing population growth under favourable conditions, SI is the annual stress index describing survival during unfavourable periods, and SX represents stress interactions. The calculation of GIA and the stress indices are based on the ranges of threshold parameters for species development adjusted by the user. Temperature parameters include the lower and upper thresholds and optimal range of air temperature for development, and similar parameters are used for soil moisture. In addition to temperature and moisture limitations, CLIMEX also takes into account the process of diapause, which is driven by temperature (initiation and termination temperature) and day-length thresholds. The number of generations is calculated based on the number of degree-days above the lower temperature threshold per generation.

Generally EI ranges from 0 to 100, where EI = 0 indicates climate conditions unfavourable for long-term species occurrence and EI > 30 represents very suitable climate conditions for species occurrence (Sutherst & Myawald, 1985; Sutherst *et al.*, 2001). Hoddle (2003) considers locations with EI > 25 as very favourable for species occurrence, 10 < EI < 25 as favourable and EI < 10 as limiting for species survival and occurrence. CLIMEX models use monthly input data of minimum and maximum temperature, relative humidity at 9 am and 3 pm and precipitation. In the CLIMSAVE project mean daily relative humidity is used to

approximate required inputs as sensitivity analysis showed a negligible effect on the model outputs when observed or estimated relative humidity values are used to calculate EI.

11.2 Pest occurrence model validation

In the first stage of assessing meta-model performance, the results from CLIMEX were compared with reports in the CAB International database and Fauna Europea. Figure 11.1 shows the evaluation of the presence/absence of seven species according to these databases. The figure shows the number of countries in which the results from CLIMEX and the observed databases agree (both present – light green; both absent – dark green) or disagree (orange and red). However, these data are only provided at the national level and hence have limited value for direct model validation and cannot be used to derive validation statistics, such as Kappa.

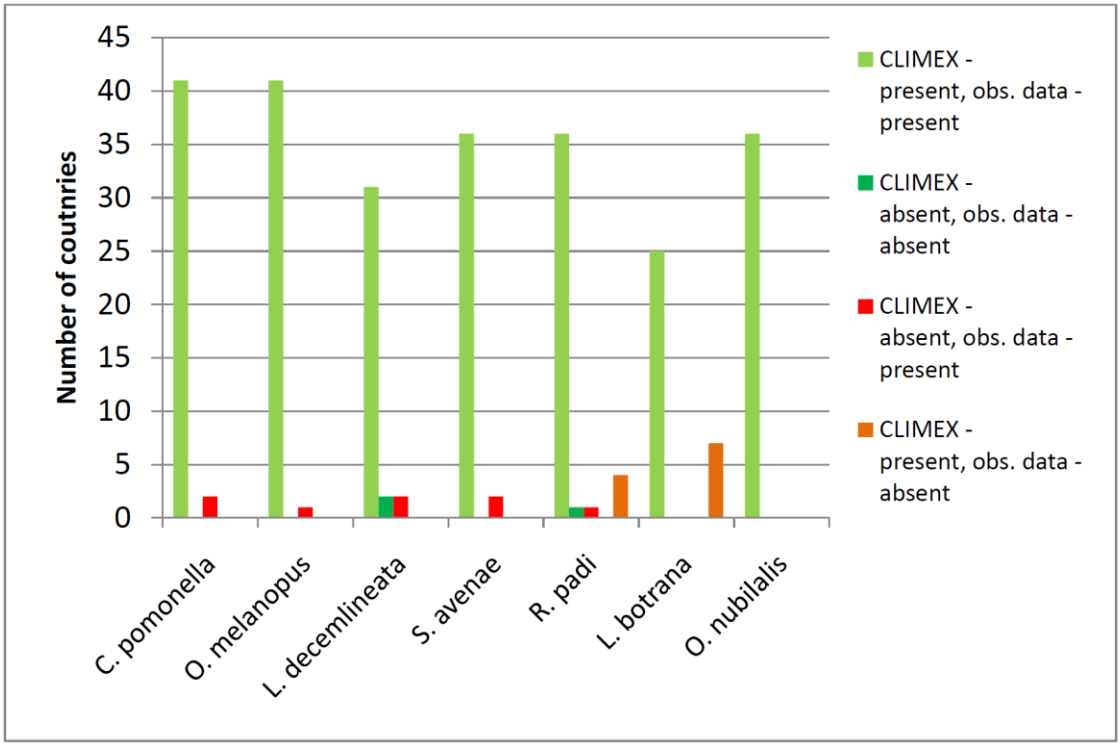


Figure 11.1: Number of European countries with presence or absence of seven pests according to CLIMEX vs. records available in databases of the observed pests' occurrences (CABI and Fauna Europea): light green – CLIMEX and databases agree on presence; dark green - CLIMEX and database agree on absence; red and orange - CLIMEX and databases disagree on presence and absence.

The performance of CLIMEX was therefore compared against a range of available data originating from published studies and the results are summarised in the following sections.

11.2.1 *Ostrinia nubilalis*

The European corn borer (*O. nubilalis*) is the most important native pest of grain maize and is widespread in Europe. Under current climate conditions, *O. nubilalis* has between one and three generations per year, depending on latitude and temperature conditions. In northern areas it has one or a partial generation. In central Europe, it has one generation in north-

western Hungary and two in the southern part of the country (Keszthelyi & Lengyel, 2002). In warmer areas the pest can complete two or three generations (southwest France and Italy).

To provide increased confidence in CLIMEX, the performance of CLIMEX was compared with that of the process-based ECAMON model (Trnka *et al*., 2007) using the reported distribution of the European corn borer (ECB) in the Czech Republic. The database, consisting of almost 900 reports of *O. nubilalis* occurrence from more than 200 sites spanning the entire Czech Republic, was derived through personal contacts with individual research stations and from farmers for the period 1961 to 2003. Figure 11.2 shows that both CLIMEX and ECAMON show very good agreement with observations during two different model periods (1961-1990 and 1991-2000). Both models also properly recorded the pest expansion based on the higher temperatures of the last decade of the 20th century, which seems to support the hypothesis that this expansion was at least partly climate driven. The slight superiority of ECAMON over CLIMEX is due to the very detailed developmental module and the use of a daily time step compared to the simpler climatology used in CLIMEX. However, CLIMEX gives reliable results whilst having far lower input data requirements, and so demonstrates its applicability for incorporation within the IAP.

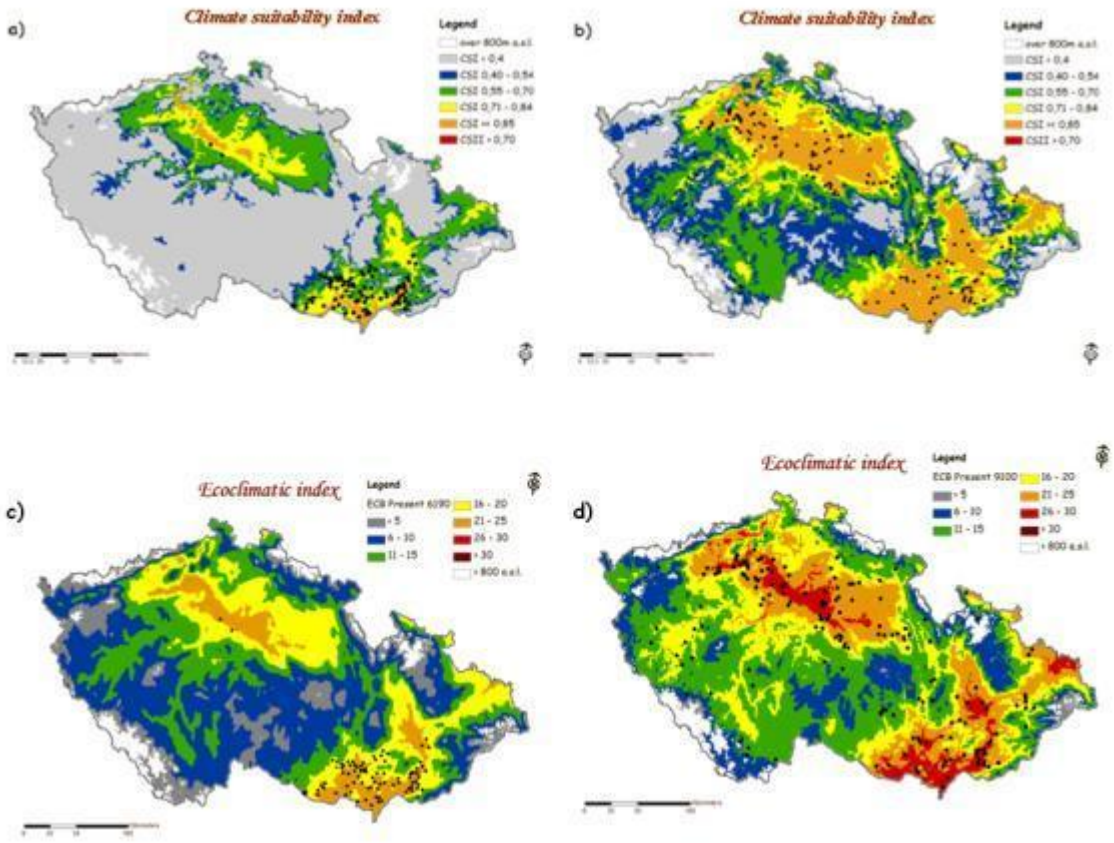


Figure 11.2: Validation of European corn borer (*Ostrinia nubilalis*) occurrence in the Czech Republic according to the detailed model ECAMON (a, b) and CLIMEX (c, d) that has been used to develop the meta-models for the IAP. Figures a) and c) correspond to the estimated range for the 1961-1990 climate, whilst b) and d) correspond to the estimated range for the 1991-2000 climate.

Validation of the CLIMEX model for *O. nubilalis* was carried out coupled with the model of *L. decemlineata* in the domain of central Europe (Kocmánková *et al*., 2001), where the predicted number of generations was in accordance with observed records. Across the whole of Europe the model correctly simulated the higher number of generations in Italy and France, and a single generation of *O. nubilalis* in northern countries such as Norway, Sweden, Finland, Denmark, United Kingdom and Ireland (Figure 11.3), where it is long-established.

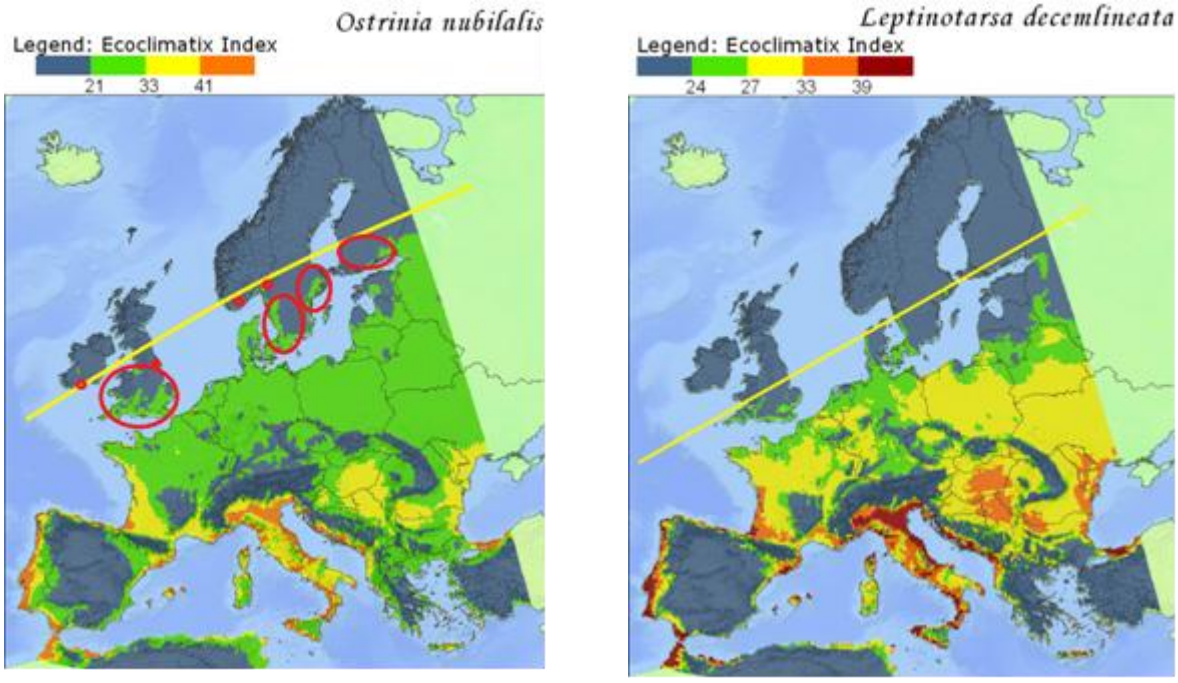


Figure 11.3: CLIMEX simulation of the Ecoclimatic Index representing the climate conditions favourable for the establishment of one (green), two (yellow), three (orange), and four (red) generations of (left) *O. nubilalis* and (right) *L. decemlineata*. The yellow line constitutes the northern range limit from CLIMEX. Red circles mark observed occurrences of the pest available in the Global Biodiversity Information Facility database (http://data.gbif.org/).

11.2.2 *Leptinotarsa decemlineata*

The Colorado potato beetle (*L. decemlineata* (Say)) is one of the most destructive potato pests. The beetle is present throughout Europe except for Britain, Ireland and Scandinavia, having its northern range limit in Russia (60°N) (EPPO, 2006). The number of Colorado beetle generations is largely a function of temperature, varying between about four in the hottest areas to one full and one partial generation near the colder extremes (Hiiesaar *et al*., 2006).

The CLIMEX model for *L. decemlineata* has been validated in previous studies, with the spatial distribution of the number of generations corresponding with observations across central Europe (Kocmánková *et al*., 2001). Within the wider spatial extent of CLIMSAVE, CLIMEX correctly indicates that the climate of southern and south-eastern England and the southern border of Sweden as potentially suitable for the establishment of the *L. decemlineata* (Figure 11.3). It is only pest management in these countries that has successfully avoided the long-term survival of the pest to date.

11.2.3 *Cydia pomonella*

Codling moth (*Cydia pomonella*) is the oldest known and the most widely distributed pest of deciduous pome fruit (Ferro and Harwood, 1973). The native home of the codling moth is considered to be south-eastern Europe from where it has spread to wherever the climate is suitable for commercial production of apple and pear trees. The present distribution of codling moth is related to climatic factors as well as to food conditions (Wearing *et al.*, 2001), with temperature considered to be the determining factor of the life-cycle length and consequently of the number of completed generations.

Records regarding the number of generations of the moth across the European area were used for validation. Codling moth develops one generation in the coldest regions, four or five generations in the hottest regions, generally three generations are present in Spain (Gonzales, 2007), two generations in Romania (Neamtu *et al.*, 2008), three generations in Italy (Reggiany *et al.*, 2006), and a maximum of two generations in the Czech Republic (SRS, 2007). The number of generations predicted by the CLIMEX simulation are in the agreement with these records (Figure 11.4), and the simulated northern boundary of the pest occurrence area corresponds with the Global Biodiversity Information Facility (http://data.gbif.org/occurrences/).

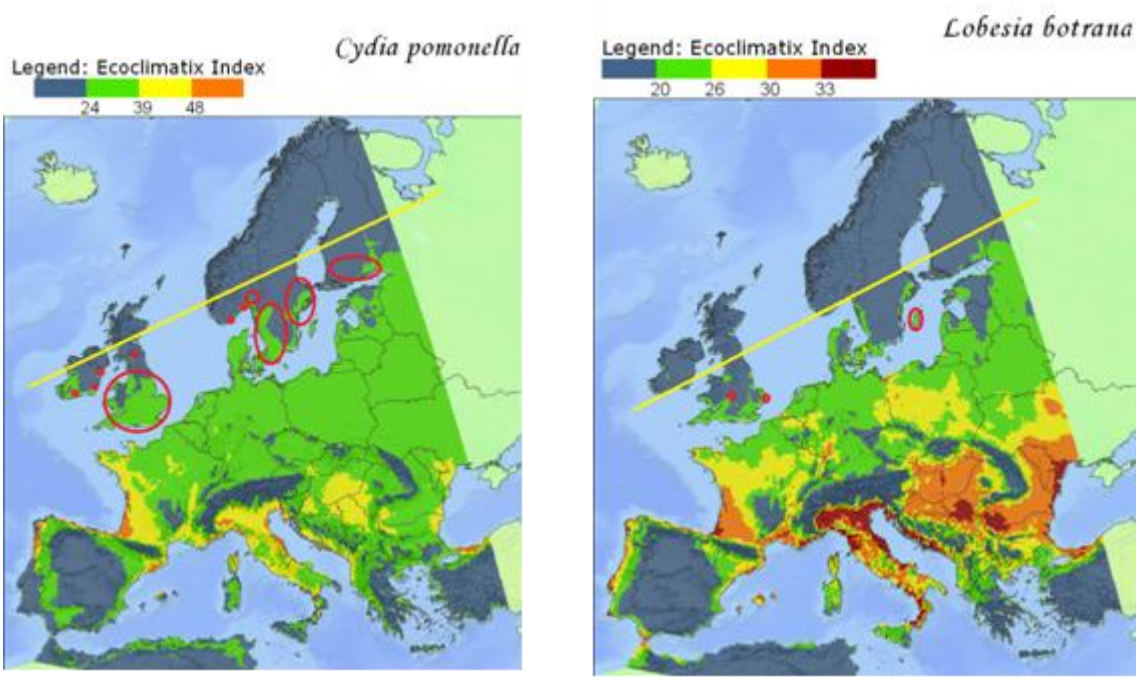


Figure 11.4: CLIMEX results of the Ecoclimatic Index representing climate conditions favourable for the establishment of one (green), two (yellow), three (orange) and four (red) generations of (left) *C. pomonella* and (right) *L. botrana*. Red circles mark observed occurrences of the pest available in the database of the Global Biodiversity Information Facility (*C. pomonella*) and Fauna Europea/Suffolk Moth Group (*L. botrana*). The yellow line constitutes the CLIMEX estimate of the potential northern range limit.

11.2.4 *Lobesia botrana*

The European grapevine moth (*L. botrana*) is a significant pest of berries and berry-like fruits in Europe and the Mediterranean. *L. botrana* is native to southern Italy but is now distributed in vineyards throughout Europe (CABI Distribution Maps of Plant Pests, www.cabi.org).

The number of generations is determined by several factors including photoperiod, temperature, humidity, latitude, food quality, and the effects of predators and diseases (Deseo *et al*., 1981). In response to differences in climate, the number of generations completed by *L. botrana* differs geographically. In general, more generations are completed in southern latitudes than in northern latitudes - up to four generations can be completed in warmer regions such as Greece (Moschos *et al*., 1998), whilst two or three generations are present in Germany (Lous *et al*., 2002).

CLIMEX simulations have successfully estimated the climate conditions which are favourable for completing the number of generations in the relevant European countries mentioned above (Figure 11.4). The CLIMEX results for Poland are consistent with Fauna Europea (www.faunaeur.org), which indicates the moth as surviving in this area. Although CLIMEX indicates that Denmark and the southern coastal areas of Sweden and Finland as the northern limit for *L. botrana* occurrence, in disagreement with both CABI and Fauna Europea, CLIMEX correctly indicates the climate suitability in Gotland (where the moth presence is recorded by Fauna Europea). *L. botrana* is a rare immigrant to Suffolk in the United Kingdom (www.suffolkmoths.org.uk), which CLIMEX has assessed as suitable.

11.2.5 *Oulema melanopus*

The Cereal leaf beetle (*O. melanopus*) is an invasive pest of small grain cereal crops, particularly of wheat, oats, and barley (CAB International, 2002). This species is now present throughout Europe. *O. melanopus* typically has one generation per year, but occasionally two years are necessary to complete the development of a single generation in more northern climates (NCSU, 2003).

The climate suitability for the establishment of *O. melanopus* was, due to the obligate univoltinism of the pest, evaluated at three levels: unsuitable, suitable and very suitable climate. The model correctly predicts pest presence in northern areas such as in Norway, Sweden, Finland and Denmark, and in the United Kingdom and Ireland where it is widespread. The moisture requirements of the pest in southern countries like Greece and Italy were also fulfilled (Figure 11.5).

11.2.6 *Rhopalosiphum padi and Sitobion avenae*

The Bird cherry-oat aphid (*R. padi*) and the English grain aphid (*S. avenae*) which are both cereal pests, are important vectors of plant viruses that may cause considerable damage, the most important among them being BYDV (Barley yellow dwarf virus). The distribution of cereal aphids is generally affected by climatic conditions and some biotic factors such as the quality of host plants, dispersal efficacy and natural enemies (Elliot and Kieckhefer, 2000).

The geographical distribution of both species is almost pan-European, including Scandinavia, UK, Ireland and also southern locations such as Italy or Sicily. However, the CABI and Fauna Europea databases do not provide more detailed specification of the pests' occurrence. The CLIMEX model matches the occurrence of infested areas such as in Norway where CLIMEX estimated suitable climate conditions for both species on the south-eastern coast only and for *S. avenae* on the southern coast of Finland (Figure 11.6). The verification of the number of generations is also rather problematic due to the complicated and variable reproduction cycle of aphids, but there are records in England where *S. avenae* can develop eighteen generations, in agreement with CLIMEX.

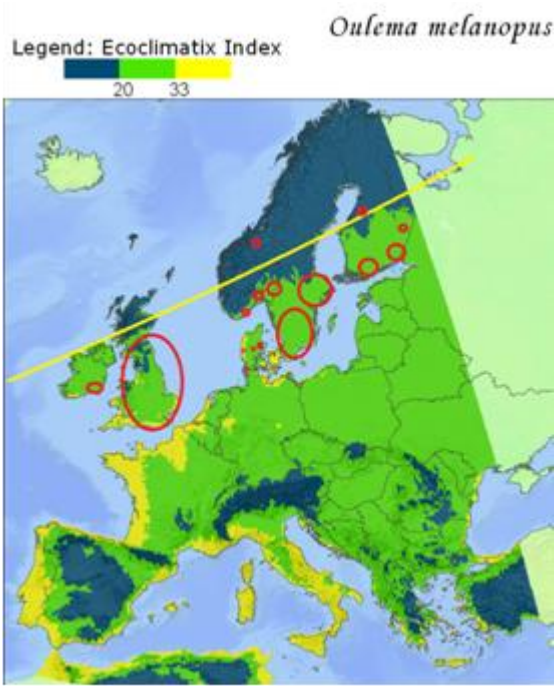


Figure 11.5: CLIMEX results of the Ecoclimatic Index representing Suitable (yellow) and Very Suitable (green) climate conditions for the establishment of *O. melanopus*. Red circles mark observed occurrences of the pest available in the Global Biodiversity Information Facility (http://data.gbif.org/) database. Yellow line constitutes the potential northern range as was estimated based on the CLIMEX results.

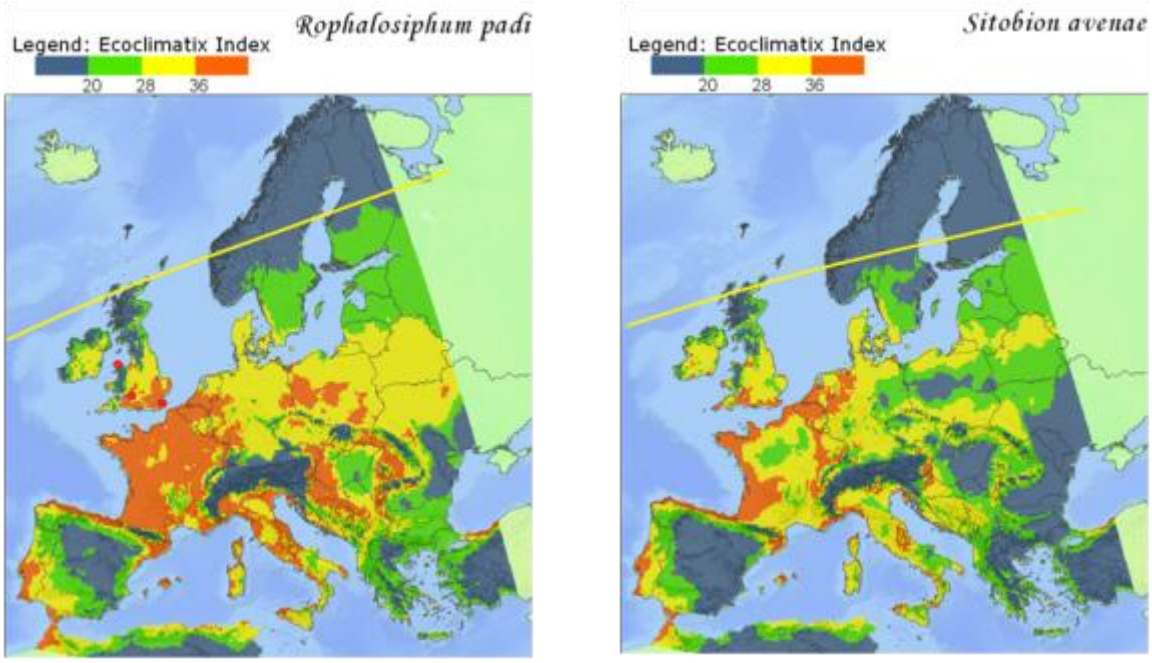


Figure 11.6: CLIMEX simulation of the Ecoclimatic Index representing climate conditions favourable for the establishment of eight (green), twelve (yellow), and sixteen (orange) generations of (left) *R. padi* and (right) *S. avenae*. Red circles mark observed occurrences of the pest available in the Global Biodiversity Information Facility database. Yellow line constitutes the potential northern range as estimated by CLIMEX.

11.3 Development and validation of the pest meta-model

Overall the CLIMEX model reproduces well the regional, as well as the local, presence/absence suitability for the seven pest species, and therefore has been used as the basis for the development of the pest meta-models. Preliminary pest meta-models based on step-wise regression models lacked precision and reliability, and so meta-models using artificial neural networks (ANNs) were developed to reproduce the behaviour of the CLIMEX model.

The procedure for developing the pest meta-models is summarised in Figure 11.7. In the first stage, the best performing ANN design (e.g. input parameter selection, number of layers and hidden layers) was determined by training the ANN on a calibration dataset and then validating it on an independent validation subset. The best design was then initiated using 100 different random seeds and the top five ANNs were selected based on their R^2, RMSE and MBE. The training and validation dataset included the whole CLIMSAVE 10´ European domain (1961-1990). For the five top ANNs, a higher number of iterations were used in order to obtain the best final meta-model ensemble. The run-time of the meta-models increases considerably with the number of ANNs in the ensemble and therefore it was decided to keep the number low (five) whilst maintaining good model performance.

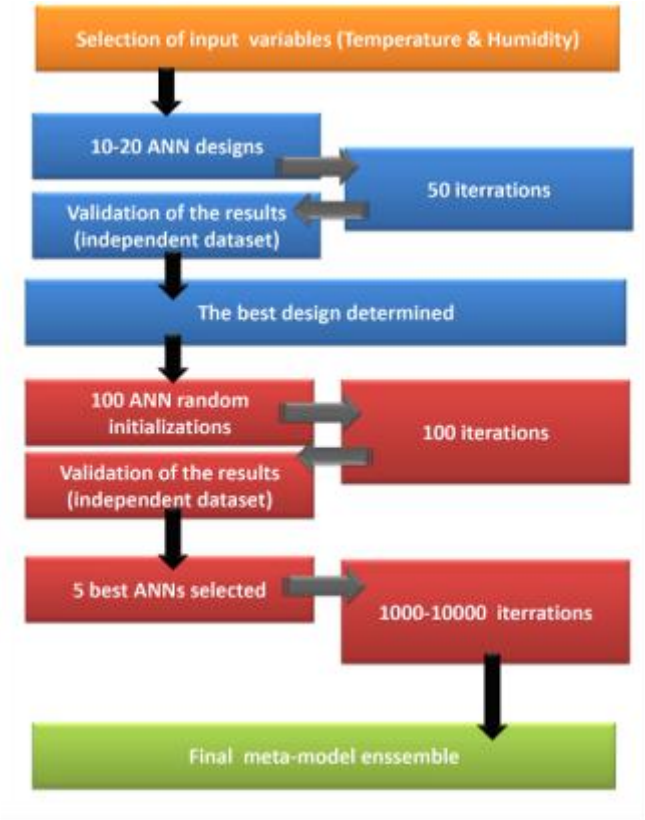


Figure 11.7: Overview of the development of the pest meta-models based on ANNs.

The results of 13 meta-models are summarised in Table 11.1, showing the ensemble mean performance and the range across the five constituent meta-models. It shows very good performance for the meta-models for all pest species for both the Ecoclimatic Index and the number of generations with at least 91% of the variability explained. Figure 11.8 also shows, as examples, the excellent spatial comparison between the results of CLIMEX and that of the

mean of the five constituent ANN meta-models for the Ecoclimatic Index for three species for the period 1961-2000.

Table 11.1: Results of the meta-model validation for the period 1980-1990 for the ensemble mean (range across the five best performing ANNs in parentheses).

	R^2		RMSE		MBE	
Pest species	Ecoclimatic Index	Number of Generations	Ecoclimatic Index	Number of Generations	Ecoclimatic Index	Number of Generations
Cydia pomonella	0.99 (0.98-0.98)	0.99 (0.99-0.99)	1.77 (1.81-1.89)	0.09 (0.09-0.1)	0.03 (-0.08- 0.05)	-0.002 (-0.005-0.001)
Leptinotarsa decemlineata	0.98 (0.98-0.98)	0.99 (0.99-0.99)	1.67 (1.67-1.68)	0.06 (0.06-0.06)	-0.011 (-0.089-0.002)	0.0004 (-0.002-0.004)
Lobesia botrana	0.98 (0.98-0.98)	0.99 (0.99-0.99)	1.41 (1.45-1.49)	0.05 (0.04-0.11)	-0.012 (-0.055-0.03)	0.004 (-0.003-0.004)
Ostrinia nubilalis	0.98 (0.98-0.98)	0.99 (0.99-0.99)	1.54 (1.56-1.60)	0.04 (0.04-0.04)	-0.012 (-0.055-0.03)	0.0007 (-0.0005-0.002)
Oulema melanopus	0.99 (0.98-0.98)	-	1.54 (1.64-1.72)	-	-0.03 (-0.06-0.001)	-
Rophalosiphum padi	0.95 (0.94-0.94)	0.99 (0.99-0.99)	2.47 (2.58-2.71)	0.317 (0.367-0.49)	-0.015 (-0.05-0.11)	0.003 (-0.028-0.03)
Sitobion avenae	0.92 (0.91-0.91)	0.99 (0.99-0.99)	2.74 (2.87-2.96)	0.34 (0.45-0.47)	0.092 (-0.047-0.24)	0.001 (-0.04-0.05)

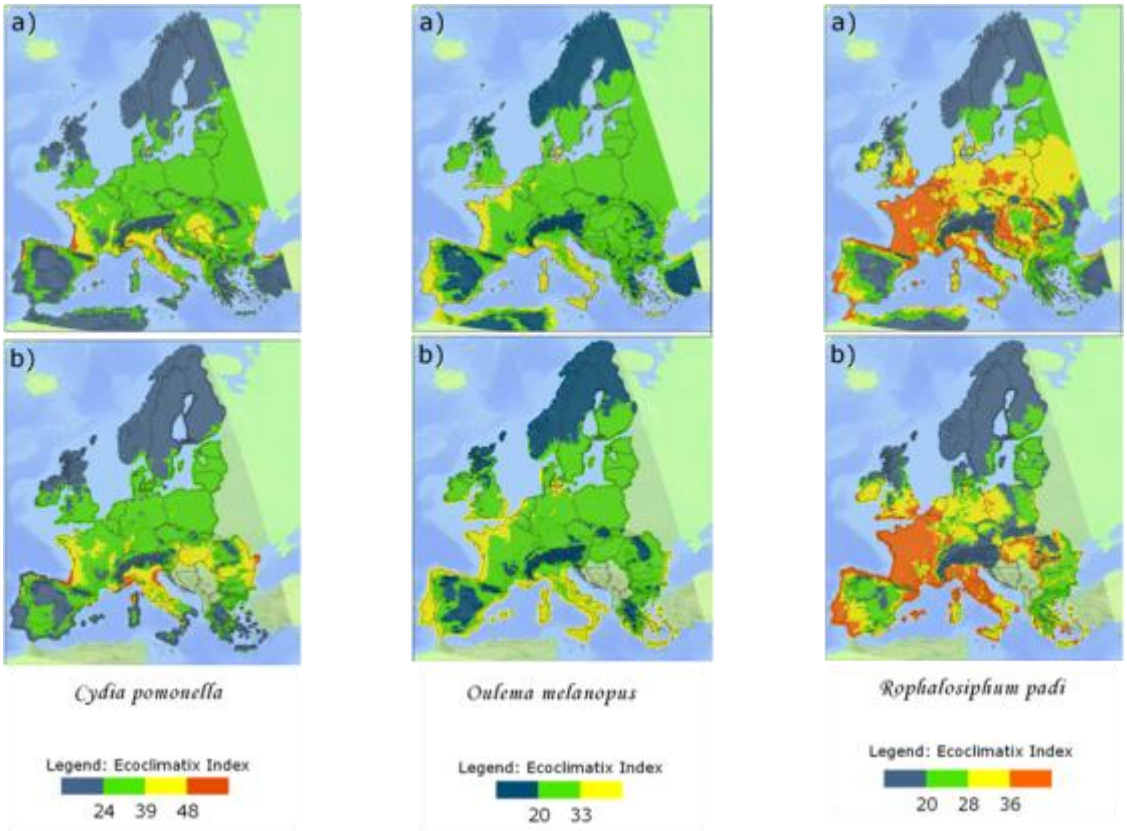
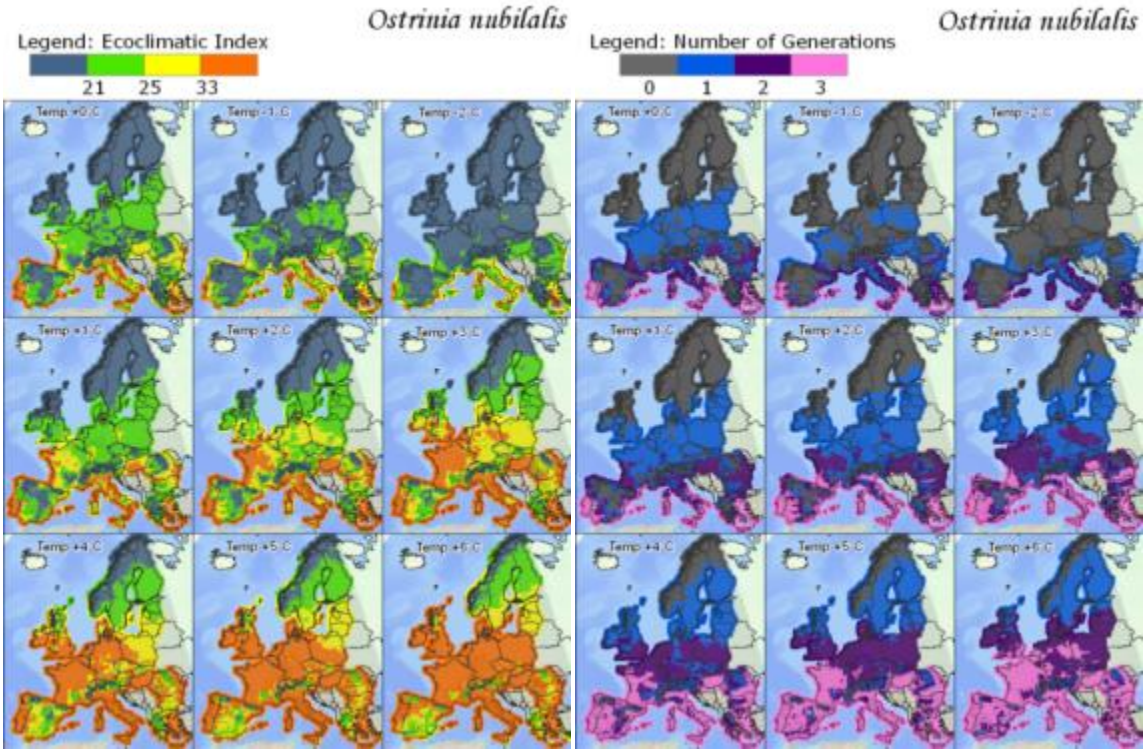


Figure 11.8: Comparison of the Ecoclimatic Index for three species (*Cydia pomonella*, *Oulema melanopus* and *Rophalosiphum padi*) according to (a) CLIMEX and (b) the mean of the five meta-models.

11.4 Pest meta-model illustrative results

Figure 11.9 shows the results of the meta-models for *Ostrinia nubilalis* across the CLIMSAVE 10´ European domain, including the sensitivity of Ecoclimatic Index and the number of generations to temperature changes across the range from -2°C to +6°C. This shows, for example, the increasing northward shift in the number of generations with increasing temperature.



a) b)

Figure 11.9: Illustrative results for the value of (a) Ecoclimatic index (EI) and (b) number of generations per season for *Ostrinia nubilalis* during the period 1961-1990 with the meta-model being tested for sensitivity to temperature change from -2°C to +6°C.

Figure 11.10 shows the expected number of generations for the Colorado potato beetle in the 2050s assuming CSMK3 climate model, A1 emission scenario with medium climate sensitivity. The results for the expected future climate conditions show:

1. A first generation in the Scandinavian region (Norway, Sweden, Finland) as far as 64° and in UK to 57 °N latitude for the first time compared to baseline.
2. Second generations cover the coherent region 46-60°N (46-54°N in UK) and southern areas surrounding the dryer regions or foothills of Spain, France, Italy, Romania and Bulgaria.
3. A third generation is present:
    a) especially in France, in the western lowlands of Germany, almost the whole of Poland, and the lowlands in Bohemia.
    b) The next core area of a third generation is in the central European region in southern Moravia, the lowlands of Austria, and the whole of Hungary

extending to Slovenia and Romania. In western and central Hungary the third generation progresses to the fourth generation. The minor regions with presence of third generation are in the Black Sea coast (Moldova and Bulgaria), isles of eastern Denmark, Italy, Sardinia, and the coastal and western part of Spain.

c) Negligible niches providing climate suitable for the third generation are visible in Greece, the southern coast of the UK and the eastern part of Sweden.

4. Four generations occur in the warmest areas – Portugal and western Spain, the east coast and lowlands of France, the lowlands of Italy, central Hungary, the south-western border of Romania and the Black Sea coast.

Differences between the baseline and climate change scenario (Fig. 11.10) determine the areas with increases/decreases in the number of generations across the European domain. An increase of about one generation is recorded in areas which currently have one or two generations. This growth is especially significant in areas north of 55°N where this represents the pest's new establishment (growth of about one generation) or the shift to second generations in primary univoltine regions. A similar trend is obvious in higher altitudes of the whole European domain. Decreases of about one or two generations are simulated in the Pannonian lowland.

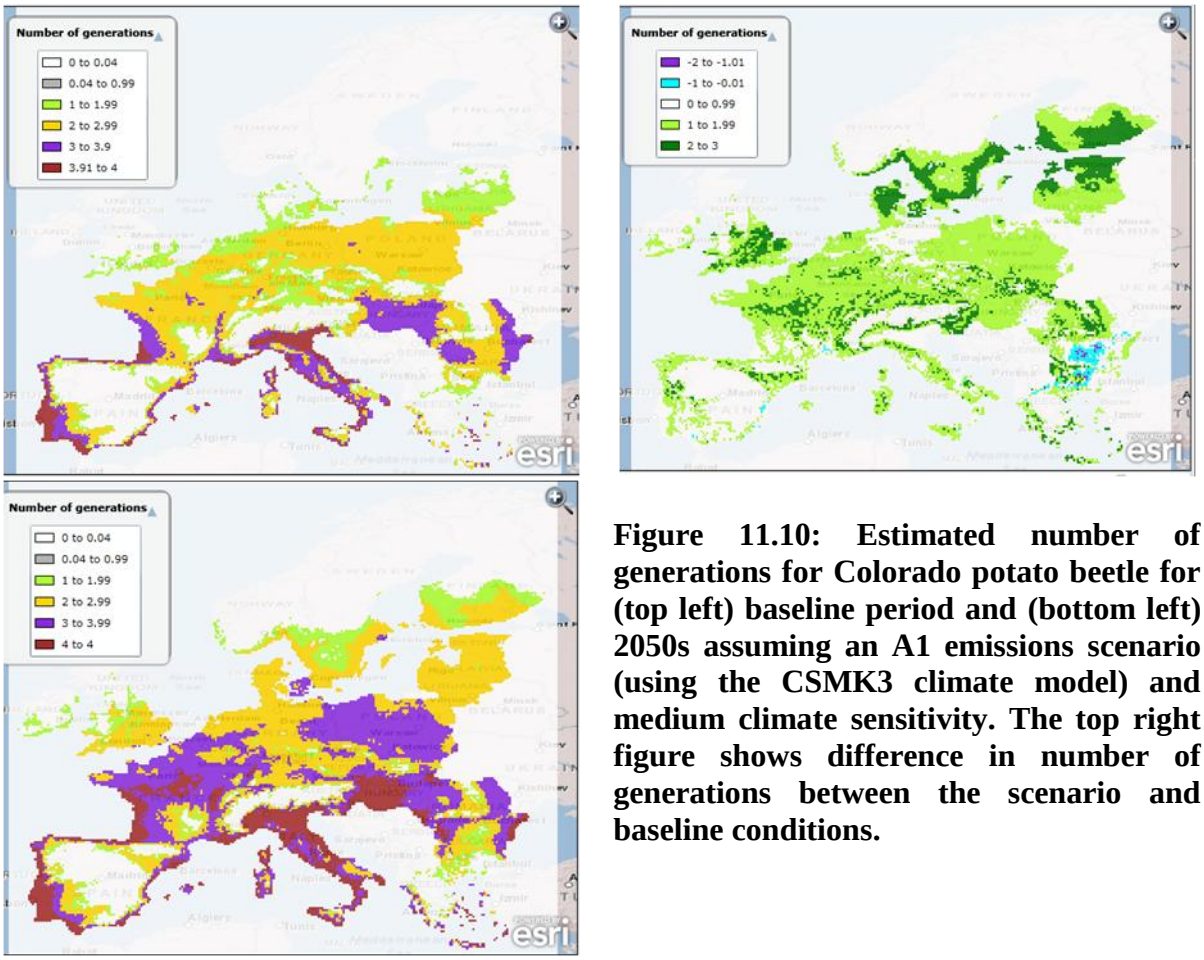


Figure 11.10: Estimated number of generations for Colorado potato beetle for (top left) baseline period and (bottom left) 2050s assuming an A1 emissions scenario (using the CSMK3 climate model) and medium climate sensitivity. The top right figure shows difference in number of generations between the scenario and baseline conditions.

11.6 References

CAB International (2002). Crop Protection Compendium, 2002 ed., *Oulema melanopus* L. Pest Module. CAB International (CD-ROM).

Elliott, N.C. & Kieckhefer, R.W. (2000). Response by coccinellids to spatial variation in cereal aphid density. Population Ecol., 42: 81–90.

EPPO (2006), Distribution Maps of Quarantine Pests for Europe, *Leptinotarsa decemlineata*. http://pqr.eppo.org/datas/LPTNDE/LPTNDE.pdf.

Ferro, D.N. & Harwood, R.F. (1973). Intraspecific larval competition by the codling moth, *Laspeyresia pomonella*. Environ. Entomol., 2: 783-789.

Gonzáles, D.C. (2007). Cydia pomonella behaviour and responses to host volatiles. Ph.D. thesis, Univesitat de Lleida, Department de Química.

Hiiesaar, K., Kuusik, A., Jõudu, J. *et al.* (2001). Laboratory experiments on cold acclimation in overwintering Colorado potato beetles, *Leptinotarsa decemlineata* (Say). Norwegian Journal of Entomology, 48, 87–90.

Hoddle, M.S. (2003). The potential adventive geographic range of glassy-winged sharpshooter, *Homalodisca coagulata* and the grape pathogen *Xylella fastidiosa*: implications for California and other grape growing regions of the world. Crop Protection, 23, 691-699.

Keszthelyi, S. & Lengyel, Z. (2002). Összefoglalás flight of the European corn borer (*Ostrinia nubilalis* Hbn.) as followed by light- and pheromone traps in Várda and Balatonmagyaród. Journal of Central 56 European Agriculture, Volume 4 (2002).

Kocmánková, E., Trnka, M., Eitzinger, J., Dubrovský, M., Štěpánek, P., Semerádová, D., Balek, J., Skalák, P., Farda, A., Juroch, J. & Žalud, Z. (2011). Estimating the impact of climate change on the occurrence of selected pests at a high spatial resolution - a novel approach. The Journal of Agricultural Science, 149(2), 185-195. ISSN 0021-8596.

Louis, F., Schmidt-Tiedemann, A. & Schirra K.J. (2002). Control of *Sparganothis pilleriana* Schiff. and *Lobesia botrana* Den. & Schiff. in German vineyards using sex pheromone-mediated mating disruption. Use of pheromones and other semiochemicals in integrated production, IOBC wprs, Bulletin Vol. 25.

Neamtu, M., Barbulescu, Petrisor, A.C., Stinga, A., Cornel, S., Drosu, S., Chireceanu, C., Ioan, R. & Cristian, B. (2008). Using a Phenology Model of Codling Moth (*Cydia pomonella*) to Improve Timing of Control Measures in Apple Orchards. Proc. VIIIth IS on Modelling in Fruit Research Ed.: J. Samietz, Acta Hort.,803, ISHS 2008.

Moschos, T., Broumas, T., Souliotis, C., Tsourgianni, A. & Kapothanassi, V. (1998). Experiments on the control of the European grapevine moth *Lobesia botrana* Den. et Schiff. (Lepidoptera, Tortricidae) with the mating disruption method in the area of Spata Attiki, Greece. Anales de l'Institut Phytopathologique Benaki, 18: 81-95.

NCSU, USDA-APHIS-PPQ-CPHST-PERAL (2003). Pest Assessment: Cereal Leaf Beetle, *Oulema melanopus* (L.), (Coleoptera: Chrysomelidae). http://www.nappfast.org/pest%20reports/oulema_melanopus.pdf.

Reggiany, A., Curto, G., Vergnani, S., Cornale, R., Benuzzi, M., Ladurner, E. & Caruso, S. (2006). Effectiveness of entomopathogenic nematodes in the control of *Cydia pomonella* overwintering larvae in Northern Italy. Summit workshop COST 850 "Biocontrol Symbiosis", 1-6 Juhne 2006, Salzau, Germany.

SRS (2007). Informace o výskytu hlavních škůdců a chorob na území ČR ve vztahu ke globálnímu oteplování. (Information on the occurrence of main species and diseases in relation to the global warming) http://www.srs.cz, Published on web: Dec. 10 2007.

Sutherst, R.W. (2000). Climate change and invasive species - a conceptual framework. 211-240. In Mooney HA & Hobbs RJ (eds) Invasive Species in a Changing World. Island

Press, Washington DC, 457pp
Sutherst, R.W. & Maywald, G.F. (1985). A computerised system for matching climates in ecology. Agriculture Ecosystems and Environment, 13, 281-99.
Sutherst, R.W., Maywald, G.F., Bottomley, W. & Bourne A. (2001). CLIMEX v2 User's Guide. CSIRO Entomology, 12-13.
Trnka, M., Muška, F., Semerádová, D., Dubrovský, M., Kocmánková, E. & Žalud, Z. (2007). European Corn Borer Life Stage Model: Regional Estimates of Pest Development and Spatial Distribution under Present and Expected Climate. Ecological modeling, 207, 61-84.
Wearing, C.H., Hansen, J.D., Whyte, C., Miller, C.E. & Brown, J. (2001). The potential for spread of codling moth (Lepidoptera: Tortricidae) via commercial sweet cherry fruit: a critical review and risk assessment. Crop Protection, 20, 465-488.

12. Development and validation of the LPJ-GUESS biodiversity meta-model

Dorte Lehsten, Florian Sallaba and Martin T. Sykes
Department of Earth and Ecosystems Sciences, University of Lund, Sweden

12.1 LPJ-GUESS model description

LPJ-GUESS, a complex dynamic global vegetation model, is a process-orientated ecosystem modelling framework (Smith *et al.*, 2001). It simulates successional vegetation dynamics on different scales (Schurgers *et al.*, 2009, Wania *et al.*, 2009) while modelling the atmosphere-vegetation carbon and water fluxes, plant physiology, establishment, mortality, and disturbance due to land use and fire (Sitch *et al.*, 2003). The model input parameters are climate variables on a daily or monthly basis, atmospheric CO_2 concentration on an annual basis, soil parameters as static values, and plant specific traits to distinguish between species. The vegetation is modelled on a so-called stand. Within a stand, input parameters are equal. Establishment, growth, mortality, as well as disturbance events are simulated for a number of replicate patches within a stand to reduce stochasticity.

12.2 LPJ-GUESS model calibration and validation

LPJ-GUESS has been applied successfully in studies of different ecosystems and their responses to changing climatic drivers (Hickler *et al.*, 2004, Schröter *et al.*, 2005, Gritti *et al.*, 2006, Morales *et al.*, 2007, Thomas *et al.*, 2008). In these studies LPJ-GUESS was tested and further developed through extensive calibration and validation work, although ecosystem models tend to over-estimate low- to mid-range net primary production at boreal and temperate sites (Cramer *et al.*, 1999). Given this extensive previous validation of the underlying model, the LPJ-GUESS meta-model was not further calibrated or validated.

12.3 Development of the LPJ-GUESS meta-model

The application of the LPJ-GUESS framework within the CLIMSAVE IAP is unfeasible since a simulation at the European scale would take several hours and would not satisfy the "just-in-time" demands of the CLIMSAVE modelling framework. The CLIMSAVE framework demands a simulation of a few seconds. Therefore it was necessary to develop a rapid meta-model of LPJ-GUESS, implemented in a DLL (dynamic link library), that simulates ecosystem parameters in an acceptable accuracy/error range in comparison to the entire LPJ-GUESS framework. The output ecosystem parameters produced by the meta-model are net primary production (NPP), leaf area index (LAI) and aboveground carbon mass (Cmass). Input drivers of the meta-model are temperature, winter and summer precipitation and atmospheric CO_2 concentrations. The meta-model produces outcomes for the baseline, each time slice and scenario, which the user can select on the IAP.

The meta-model input drivers such as temperature, precipitation and atmospheric CO_2 are selected by users directly on the IAP. The input parameter sources, besides the baseline, are from the IPCC emission scenarios (A1, B1, A2 and B2), time slices (2020s, 2050s) and GCMs (HadGEM, GFCM21, IPCM4, CSMK3 and MPEH5). A fourth input parameter which is provided by the SFARMOD land use model (Section 10) is the percentage of four land use types in each grid cell (unmanaged forest, intensive agriculture, extensive agriculture and

abandoned land). The LPJ meta-model is run within those grid cells containing some unmanaged forest and/or abandoned land for 22 species (see Table 12.1) parameterised according to Hickler *et al*. (2012). In abandoned land, it is assumed that the potential forest vegetation has 35 years to develop. For intensive and extensive land use classes, only cool C3 grass (Species 21) and C4 grass (Species 22) are simulated. Unlimited irrigation is assumed for intensive land use and, hence, there is no influence of precipitation on the species growing.

Table 12.1: Description of modelled species based on Hickler et al. (2012).

Species no.	***Description (Latin Name)***
1	Silver fir (*Abies alba*)
2	Cranberry (*Vaccinium*)
3	Silver Birch (*Betula pendula*)
4	White Birch (*Betula pubescens*)
5	Hornbeam (*Carpinus betulus*)
6	Hazel (*Corylus avellana*)
7	Beech (*Fagus sylvatica*)
8	Ash (*Fraxinus excelsior*)
9	Cade Juniper (*Juniperus oxycedrus*)
10	Rosemary (*Rosmarinus*)
11	Norway Spruce (*Picea abies*)
12	Scots Pine (*Pinus sylvestris*)
13	Aleppo Pine (*Pinus halepensis*)
14	Aspen (*Populus tremula*)
15	Kermes Oak (*Quercus coccifera*)
16	Holm Oak (*Quercus ilex*)
17	Downy Oak (*Quercus pubescens*)
18	English Oak (*Quercus robur*)
19	Lime (*Tilia cordata*)
20	Elm (*Ulmus glabra*)
21	Cool grass (C3 herbaceous)
22	Warm grass (C4 herbaceous)

The meta-model development is based on LPJ-GUESS simulations of 63 grid cells that are situated along two cross European transects as shown in Figure 12.1. Thus, the meta-model development captures several bio-geographical as well as different climatic zones in Europe. The baseline input parameter time-series are taken from CRU TS 3.0 data covering a period from 1900 to 2006 (University of East Anglia Climate Research Unit, 2008). Furthermore, the baseline climate of the 63 grid cells was adjusted in 500 different combinations using the same ranges as the sensitivity analysis (see Table 12.3). This led to 31,500 (63*500) LPJ-GUESS simulations with NPP, LAI and Cmass outcomes. We calculated the average temperature, average sum of yearly winter and summer precipitation of the last 20 years of the climate time-series. The simulated ecosystem parameter values were also averaged over the last 20 years of the respective climate time-series combination in order to smooth extreme event peaks, e.g. flooding and heat-wave events.

Transfer functions were developed to describe the relationship between temperature, winter and summer precipitation, and atmospheric CO_2 with the desired ecosystem outputs.

Functions were fitted empirically for each species of each land-use type ecosystem parameter and the respective climate driver. These functions of the different climatic drivers were combined into a transfer function that describes the species ecosystem parameter value for the respective climate conditions (Sallaba *et al.*, in prep).

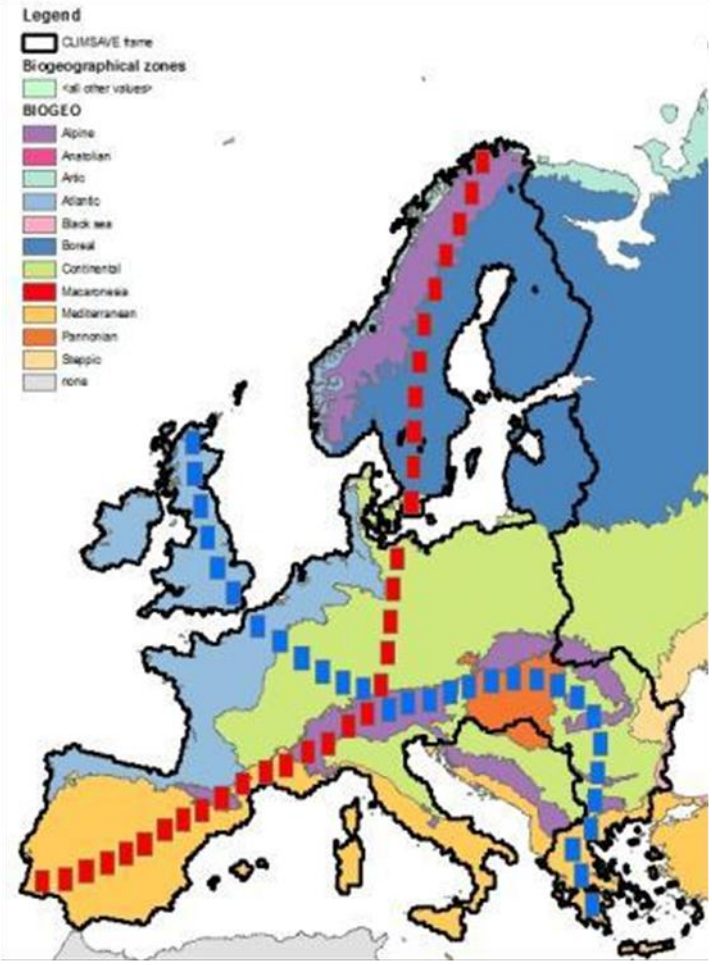


Figure 12.1: Cross European transects capturing north to south-west and north-west to south-east climatic transitions. The area of a cell is 1° and is based on the 10' CLIMSAVE grids. The extent of a grid is ~60km - ~100km depending on its location – since Lambertian equal area projection is not isogonic or isometric it leads to distortion towards the edges.

Additionally, transfer functions were established for the total values of the ecosystem parameters. A total ecosystem parameter represents the sum of all simulated species but is based on a different transfer function (Sallaba *et al.*, in prep). The total ecosystem parameters are used to scale down the estimated individual species ecosystem parameters. Thus, the sum of the individual species ecosystem parameter values will not exceed the value of the total. This step was necessary since the individual species values exceed the total ecosystem parameter values due to missing competition and other vegetation dynamics, which are difficult to implement in an empirical model.

The information on the user choice is provided by the Running Module to the DLL. Within the DLL the appropriate transfer functions are applied to calculate the ecosystem parameters of each species and to provide the relevant indicators to the IAP. Species are grouped into plant functional types (PFT) (such as coniferous trees) and the DLL sends back only those

species/PFTs that the user has requested. Timber is calculated according to a percentage of biomass, cover ratio is calculated according to LAI, and productivity is given by the net primary production. Scenic potential is a measure for landscape diversity, in this case biodiversity. Scenic potential is calculated as the sum of the squared quotients of $LAI_{species}$ to LAI_{total}. For each land use type the scenic potential is calculated and multiplied with the ratio of this land use type. Finally, the Simpson's diversity index (Simpson, 1949) is calculated, using leaf area index as a measure for species abundance, as an aggregate indicator of climate change impacts on biodiversity.

When the baseline climate is selected in the IAP, the user can change a number of sliders related to annual temperature change, summer and winter precipitation change and CO_2 concentration. The number of combinations of slider changes is too great to create look-up tables for every combination. Hence, a sensitivity analysis (see Section 12.5) of LPJ-GUESS has been undertaken to define relationships between the altered climatologies and the outputs of LPJ-GUESS.

12.4 Calibration and validation of the LPJ-GUESS meta-model

The dataset was randomly divided into calibration (50%) and validation (50%) parts. The analysis dataset (31*500) and the respective ecosystem parameter values of each species were further examined for relationships. This resulted in the ecosystem parameters of each species being described empirically using maximum functions of temperature, winter and summer precipitation and atmospheric CO_2. Each function was determined using an optimised fitting and calibration approach; and then combined to describe each ecosystem parameter of each species. The same methodology was applied to find empirical relationships for total NPP, LAI and Cmass values for the respective climate of a grid cell.

The combined functions were tested on the validation dataset (32*500). We compared LPJ-GUESS with the meta-model outcomes and applied R^2 as the measure of accuracy. Since C4-grass (Species 22) did not occur in our LPJ-GUESS simulations, it could not be included in the meta-model development due to missing data. The climate demands of warm C4-grass were probably not met in the applied climate combinations. The accuracy assessment of intensive and extensive is given in Table 12.1. The individual species validation for unmanaged forest and abandoned land is shown in Figure 12.2.

Table 12.1: Validation of the intensive and extensive land-use ecosystem parameters.

R^2 values	***NPP***	***LAI***	***CMASS***
Intensive (C3 grass)	0.33	0.47	0.46
Extensive (C3 grass)	0.75	0.79	0.80

Table 12.1 reveals a difference in accuracy between intensive (C3-grass) and extensive (C3-grass) land-use. The lower R^2 values of intensive can be explained by the difficulties to describe intensive land use using inputs of only temperature and atmospheric CO_2. The implementation of unlimited irrigation into the meta-model may be further developed using a correction factor. However, the accuracy of extensive is good compared to the individual species performances in the other land use types (i.e. compared with Figure 12.2) and supports the chosen approach of transfer functions.

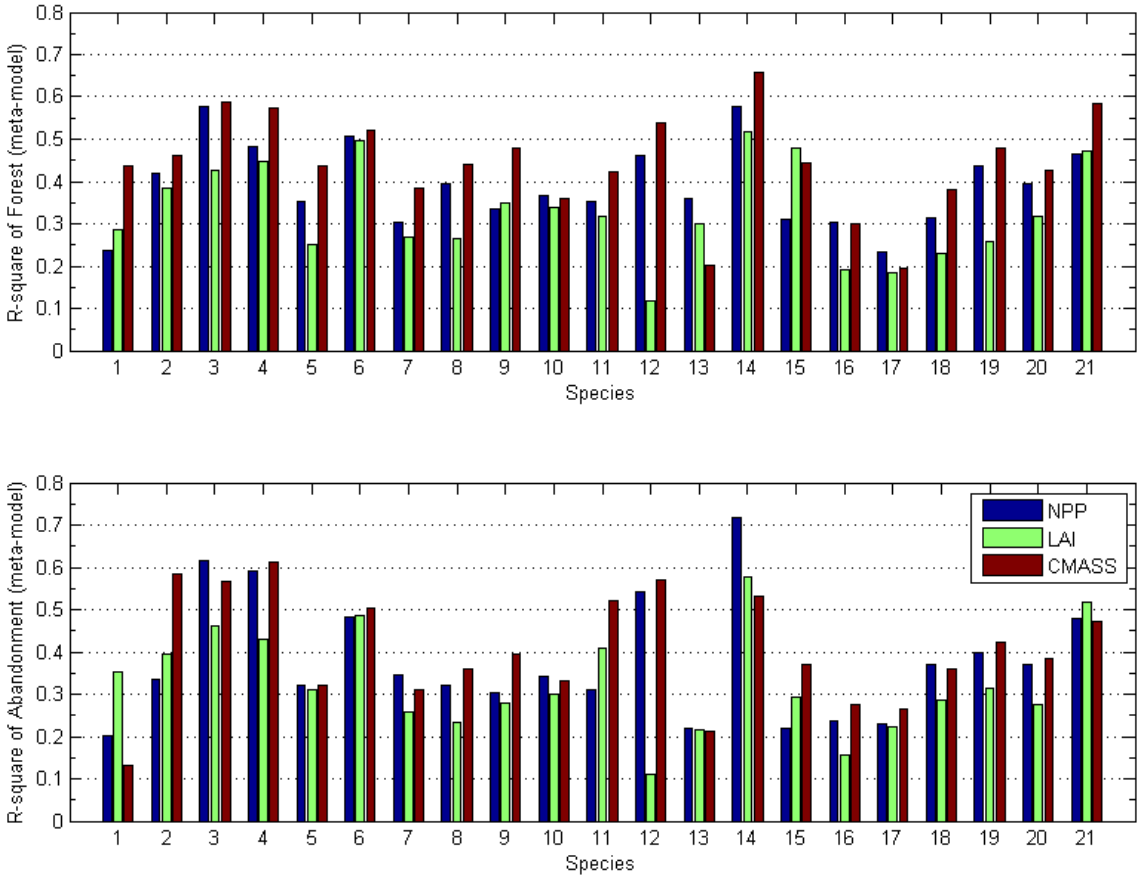


Figure 12.2: Individual species validation of the ecosystem parameters of unmanaged forest (top plot) and abandoned land (bottom plot) use types. The validation was done by comparing the LPJ-GUESS and meta-model results for individual species using R^2 as a measure of accuracy.

The validation for unmanaged forest land use shows that the majority of species have R^2 values less than 0.5 for all the ecosystem parameters. The best performing (R^2 above 0.5) species are *betula pendula* (Species 3), *betula pubescens* (Species 4), *corylus avellana* (Species 6), *populus tremula* (Species 14) and C3-grasses. The lowest accuracies (R^2 below 0.2) are found in *abies alba* (NPP & Cmass - Species 1), *pinus sylvestris* (LAI - Species 12), *quercus ilex* (LAI – Species 16) and *quercus pubescens* (Species 17).

Abandoned land-use species, as these represent an early stage of the unmanaged forest development, follow the trend for forest. The modelled species NPP values have higher accuracies than in forest. The NPP of p*opulus tremula* (Species 14) has a good R^2 of 0.7 but Cmass is lower for abandoned land. Cmass species values have somewhat lower R^2 values in comparison to forest. In particular, *abies alba* (Cmass - Species 1) is different with an R^2 value of approximately 0.1.

The general performance of the meta-model is acceptable since important vegetation dynamics (i.e. competition, mortality and disturbance) cannot be implemented into the empirically-based meta-model.

12.5 Sensitivity analysis of the LPJ-GUESS meta-model

The sensitivity analysis of LPJ-GUESS has been undertaken for the 63 grid cells situated along two transects that cover several bio-geographical zones as shown in Figure 12.1. The sensitivity of vegetation to climate parameters (annual temperature, winter and summer precipitation) and atmospheric CO_2 concentration was analysed in order to estimate their influence on net primary production, biomass accumulation and leaf area index across the Arctic, Boreal, Continental, Atlantic, Alpine, Mediterranean and Pannonian climatic and bio-geographical zones.

The LPJ-GUESS simulations were performed using CRU data from 1900 – 2006 (University of East Anglia Climate Research Unit, 2008) as the baseline climate. The sensitivity analysis is divided into two parts: (i) independent changes; and (ii) combined changes in climatic drivers.

12.5.1 *Independent changes in climatic drivers*

This part of the sensitivity analysis involves the adjustment of one climatic driver in a LPJ-GUESS simulation. Table 12.2 shows the agreed minimum and maximum values for each climatic driver and atmospheric CO_2 concentration which were adjusted during the independent sensitivity analyses by the stated increments. The climatic driver values were added to the CRU data (Mitchell and Jones, 2005) in each of the 43 simulations.

Table 12.2: Minimum and maximum values of each climatic drivers and their individual increments.

Climatic Driver	Temperature [°C]	Winter Precipitation [%]	Summer Precipitation [%]	Atmospheric CO_2 [ppm]
Min. Value	0	50	50	350
Max. Value	6	150	150	700
Increment Value	0.5	10	10	50
Sum of Steps	13	11	11	8

12.5.2 *Combined changes in climatic drivers*

The second part of the analysis is based on combined changes in multiple climatic drivers which influence vegetation dynamics in different dimensions. The climatic driver adjustments are constrained to the climatic driver values as stated in Table 12.3. The constrained climatic driver values were added to the CRU data (Mitchell and Jones, 2005) and led to 500 different simulations.

Table 12.3: Definition of the climatic drivers used within the multivariate sensitivity analysis.

Temperature change [°C]	Winter Precipitation [% of baseline]	Summer Precipitation [%of baseline]	Atmospheric CO_2 [ppm]
0	50	50	350
2	75	75	437.5
4	100	100	525
6	125	125	612.5
-	150	150	700
4	5	5	5

12.5.3 *Results of the sensitivity analysis*

The **independent changes** in climatic drivers along the transects led to a wide range of effects on NPP. Figure 12.3 visualises the range of NPP values (maximum minus minimum NPP) for each grid cell along the transects. Considering the difference between the minimum and maximum NPP values, the first transect (red line in Figure 12.3) has small NPP variations in the Norwegian Alpine and Boreal bio-geographical zones (0.1 – 0.2 kgC*m^{-2}). Increasing NPP variations up to ~0.3 kgC*m^{-2} occur from southern Sweden to southern Germany in the Continental bio-geographical zone. In the Alpine zone the NPP variations are less elevated at ~ 0.2 kgC*m^{-2}. High NPP variations (up to 0.4 kgC*m^{-2}) are modelled in the Mediterranean zone of France and northern Spain. In central Spain NPP variations are stable on an elevated level and increase strongly (0.4 kgC*m^{-2}) towards the Atlantic coast.

The second transect (blue line in Figure 12.3) shows a slight increase in NPP variations from Scotland (~0.2 kgC*m^{-2}) through England (~0.25 kgC*m^{-2}) and northern France (0.3 kgC*m^{-2}) in the Atlantic bio-geographical zone. In the Continental zone the variations are stable at ~0.3 kgC*m^{-2} NPP towards the Alpine zone. The model results show a decline of ~0.2 kgC*m^{-2} NPP variations in the Alpine zone. From the eastern Alpine to the Pannonian zone NPP variations increase up to ~0.35 kgC*m^{-2}. In the Romanian Alpine (Carpathian Mountains) zone the variations show a dip. Then the NPP variations increase to ~0.35 kgC*m^{-2} in the Continental bio-geographical zone. Towards Bulgaria and the Mediterranean, NPP variations decrease down to ~0.25 kgC*m^{-2}.

The effect of independently changing climatic drivers leads to elevated NPP variations in the Continental and Mediterranean bio-geographical zones. The Alpine and Boreal zones as well as areas of higher altitude seem to be less sensitive towards the independent changes.

The main driver of maximum total NPP values is high atmospheric CO_2 concentrations (600 to 700 ppm) in 89% of the grid cells. This is due to the effects of CO_2 fertilization as a main driver of NPP as reported by Cramer *et al.* (2001). In the Norwegian Alpine zone elevated temperature (2 - 3.5°C) leads to maximum total NPP values in ~10% of the grid cells.

Minimum NPP values are caused by rising temperature (5 - 6°C) in 55 % of the grid cells, mostly in the Mediterranean and Continental bio-geographic zone due to water limitation caused by high temperatures. Elevated summer (110 – 150%) and winter precipitation (140 – 150%) led to minimum NPP values in 12 % of the grid cells, mostly in the Alpine zone.

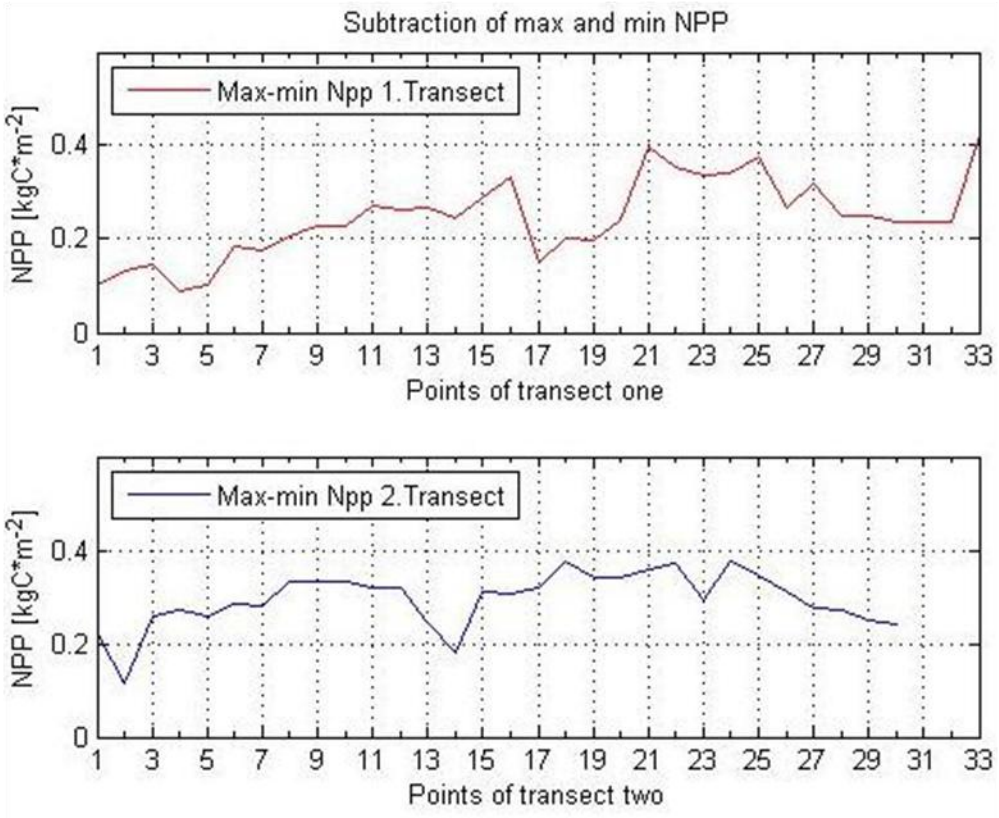


Figure 12.3: Independent changes in climatic drivers that cause minimum and maximum values of NPP in each cell. Showing the range of minimum and maximum annual total NPP for each grid cell along the two transects: [upper] north to south-west transect; [lower] north-west to south-east transect.

The **combined changes** in climatic drivers along the transect led to higher dynamics of NPP compared to the independent changes. Figure 12.4 illustrates the range of NPP values (maximum minus minimum NPP) for each grid cell. The first transect (red line in Figure 12.4) has small NPP variations in the Norwegian Alpine bio-geographical zones (~0.2 kgC*m^{-2}). From the Boreal zone in northern Sweden to the Continental zone in southern Germany, NPP variations increase from ~0.2 kgC*m^{-2} to ~0.5 kgC*m^{-2}. In the Alpine zone, NPP variations decline to ~0.3 kgC*m^{-2}. High NPP variations (up to ~ 0.6 kgC*m^{-2}) are modelled in the Mediterranean zone of France and northern Spain. In central Spain NPP values vary at an elevated level of ~0.35 kgC*m^{-2} and increase strongly (0.4 kgC*m^{-2}) towards the Atlantic coast.

The second transect (blue line in Figure 12.4) shows fluctuating NPP variations from Scotland (~0.3 kgC*m^{-2}) through England and northern France (~0.4 kgC*m^{-2}) in the Atlantic bio-geographical zone. However, there might be a slight increase of NPP to the east. In the Continental zone the variations show a fluctuation of NPP values at a low level along the transect (~0.45 kgC*m^{-2}). The NPP variations decrease in the Alpine zone to ~0.3 kgC*m^{-2}. From the eastern Alpine to the Pannonian zone, NPP variations increase by up to ~0.5 kgC*m^{-2}. In the Romanian Alpine (Carpathian Mountains) zone, the variations show a dip followed by a strong increase (~0.6 kgC*m^{-2}). Towards Bulgaria and the Mediterranean, NPP variations decrease slightly to ~0.5 kgC*m^{-2}. In southern Greece NPP variations fall to ~0.35 kgC*m^{-2}.

The climatic driver combinations of high atmospheric CO_2 concentrations (612.5-700 ppm) and increased summer (125 - 150%) and winter (125 - 150%) precipitation led to maximum NPP values in 52 % of the grid cells. In the Alpine zones increased temperature (+2-4°C), high atmospheric CO_2 concentrations (612.5-700 ppm) and elevated summer precipitation led to maximum NPP. Minimum NPP is caused mainly by high temperature (+6°C) and decreased summer precipitation (50%) in 71% of the grid cells. The effect of changes in winter precipitation vary, but it generally has less influence on minimum NPP compared to the other drivers.

In the Norwegian Alpine zone minimum NPP is caused by increased summer and winter precipitation. Temperature does not limit NPP production. The other European Alpine zones are sensitive to increased temperature (+6°C) and decreased precipitation patterns, resulting in minimum NPP values.

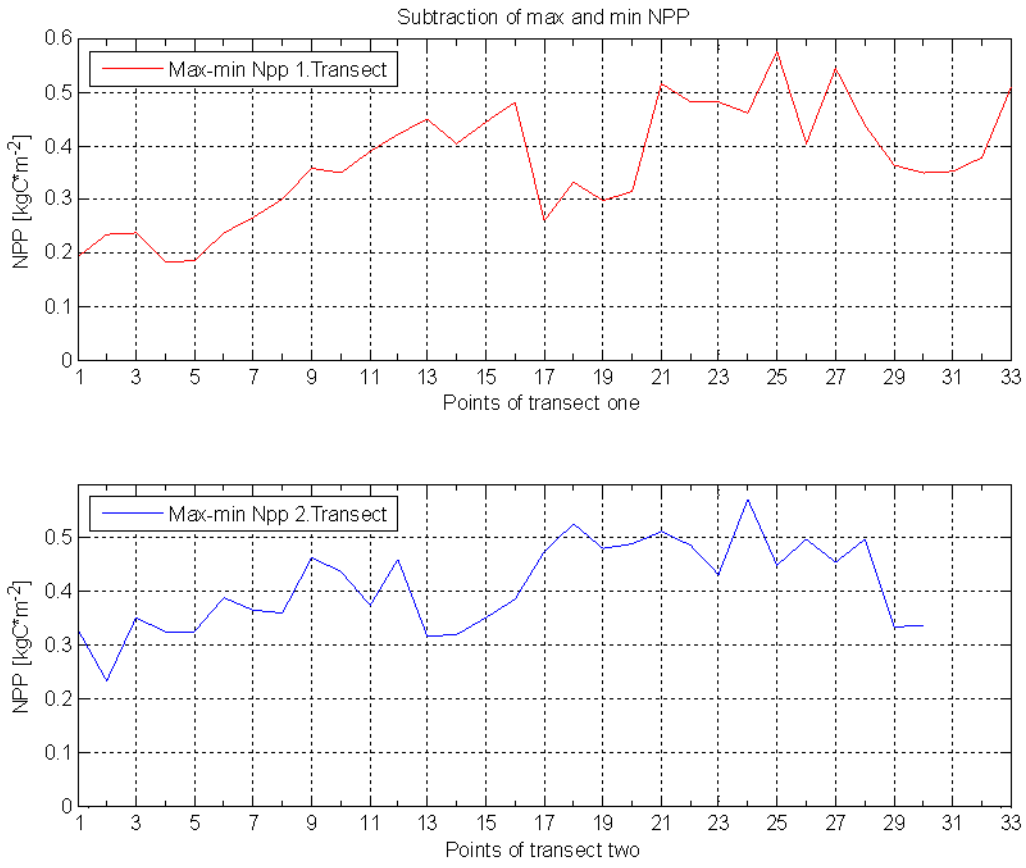


Figure 12.4: Combined changes in climatic drivers that cause minimum and maximum NPP. Showing the range of maximum and minimum annual total NPP for each grid cell along the two transects: [upper] north to south-west transect; [lower] north-west to south-east transect.

12.6 LPJ-GUESS meta-model illustrative results

The LPJ-GUESS meta-model produces outputs for NPP, LAI and potential biomass for the species within the Plant Functional Type (PFT) selected by the user on the IAP. Figure 12.5 shows illustrative output for total potential NPP for the Boreal needle leaved evergreen tree

and the Temperate broadleaved evergreen tree PFTs under a 2050s climate scenario based on the CSMK3 climate model, the A1 emissions scenario and middle climate sensitivity.

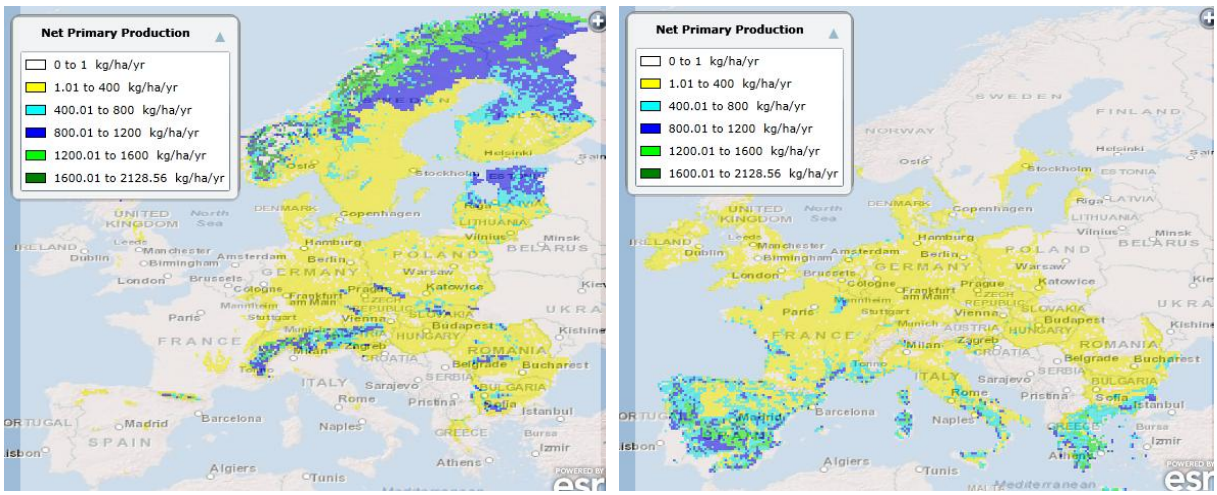


Figure 12.5: Illustrative results from the IAP for total potential NPP for two plant functional types under the CSMK3 climate model, the A1 emissions scenario and middle climate sensitivity in the 2050s: Boreal needle leaved evergreen tree (left) and ??? (right).

12.7 References

Cramer, W., Bondeau, A., Woodward, F. I., Prentice, I. C., Betts, R. A., Brovkin, V., Cox, P. M., Fisher, V., Foley, J. A., Friend, A. D., Kucharik, C., Lomas, M. R., Ramankutty, N., Sitch, S., Smith, B., White, A. & Young-Molling, C. (2001) Global response of terrestrial ecosystem structure and function to CO2 and climate change: results from six dynamic global vegetation models. *Global Change Biology*, **7,** 357-373.

Cramer, W., Kicklighter, D. W., Bondeau, A., Iii, B. M., Churkina, G., Nemry, B., Ruimy, A., Schloss, A. L. & Intercomparison, T. P. O. F. T. P. N. M. (1999) Comparing global models of terrestrial net primary productivity (NPP): overview and key results. *Global Change Biology*, **5,** 1-15.

Gritti, E. S., Smith, B. & Sykes, M. T. (2006) Vulnerability of Mediterranean Basin ecosystems to climate change and invasion by exotic plant species. *Journal of Biogeography*, **33,** 145-157.

Hickler, T., Smith, B., Sykes, M. T., Davis, M. B., Sugita, S. & Walker, K. (2004) USING A GENERALIZED VEGETATION MODEL TO SIMULATE VEGETATION DYNAMICS IN NORTHEASTERN USA. *Ecology*, **85,** 519-530.

Hickler, T., Vohland, K., Feehan, J., Miller, P. A., Smith, B., Costa, L., Giesecke, T., Fronzek, S., Carter, T. R., Cramer, W., Kühn, I. & Sykes, M. T. (2012) Projecting the future distribution of European potential natural vegetation zones with a generalized, tree species-based dynamic vegetation model. *Global Ecology and Biogeography*, **21,** 50-63.

Mitchell, T. D. & Jones, P. D. (2005) An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *International Journal of Climatology*, **25,** 693-712.

Morales, P., Hickler, T., Rowell, D. P., Smith, B. & Sykes, M. T. (2007) Changes in European ecosystem productivity and carbon balance driven by regional climate model output. *Global Change Biology*, **13,** 108-122.

Schröter, D., Cramer, W., Leemans, R., Prentice, I. C., Araújo, M. B., Arnell, N. W., Bondeau, A., Bugmann, H., Carter, T. R., Gracia, C. A., de la Vega-Leinert, A. C., Erhard, M., Ewert, F., Glendining, M., House, J. I., Kankaanpää, S., Klein, R. J. T., Lavorel, S., Lindner, M., Metzger, M. J., Meyer, J., Mitchell, T. D., Reginster, I., Rounsevell, M., Sabaté, S., Sitch, S., Smith, B., Smith, J., Smith, P., Sykes, M. T., Thonicke, K., Thuiller, W., Tuck, G., Zaehle, S. & Zierl, B. (2005) Ecosystem Service Supply and Vulnerability to Global Change in Europe. *Science,* **310,** 1333-1337.

Schurgers, G., Hickler, T., Miller, P. A. & Arneth, A. (2009) European emissions of isoprene and monoterpenes from the Last Glacial Maximum to present. *Biogeosciences,* **6,** 2779-2797.

Simpson, E. H. (1949) MEASUREMENT OF DIVERSITY. *Nature,* **163,** 688-688.

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T., Thonicke, K. & Venevsky, S. (2003) Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biology,* **9,** 161-185.

Thomas, C. D., Ohlemuller, R., Anderson, B., Hickler, T., Miller, P. A., Sykes, M. T. & Williams, J. W. (2008) Exporting the ecological effects of climate change. *EMBO Rep,* **9,** S28-S33.

Wania, R., Ross, I. & Prentice, I. C. (2009) Integrating peatlands and permafrost into a dynamic global vegetation model: 1. Evaluation and sensitivity of physical land surface processes. *Global Biogeochemical Cycles,* **23**.

University of East Anglia Climate Research Unit (CRU). [Phil Jones, Ian Harris]. CRU TS3.0, [Internet]. NCAS British Atmospheric Data Centre, 2008, 2011/02/28. Available from http://badc.nerc.ac.uk/view/badc.nerc.ac.uk__ATOM__dataent_1256223773328276.

13. Development and validation of the SPECIES biodiversity meta-model

Paula A Harrison, Robert Dunford and Pam M Berry
Environmental Change Institute, University of Oxford, UK

13.1 SPECIES model description

The SPECIES model (Spatial Estimator of the Climate Impacts on the Envelope of Species; Pearson et al., 2002) is used in the IA Platform to simulate the impacts of climate change on the suitable climate space of over 100 species. The species were selected to interact with the agricultural, forest, coastal and water sectors and to indicate a range of ecosystem services (pollination, berries for food from wild plants, charismatic or iconic wildlife for aesthetic enjoyment, and species for hunting; see Section 14, Table 14.1).

SPECIES is based on ensembles of artificial neural networks (ANN), which integrate bioclimatic variables for projecting the distribution of species through the characterisation of bioclimatic envelopes. Integrated algorithms, including a soil water balance model, are used to pre-process climate (temperature, precipitation, solar radiation and wind speed) and soils (AWC – available water holding capacity) data to derive relevant bioclimatic variables for input into the ANN. Those variables found to be most successful for projecting the distributions of birds (Harrison *et al.*, 2003) and other taxa (Berry *et al.*, 2003) are given in Table 13.1.

Table 13.1: Bioclimatic input variables used for birds and other taxa in the SPECIES model (from Harrison *et al.*, 2006).

Birds	Other taxa
Growing degree days > 5°C	Growing degree days > 5°C
Absolute minimum temperature expected over a 20-year period	Absolute minimum temperature expected over a 20-year period
Mean summer temperature (MJJ)[a]	Annual maximum temperature
Mean summer precipitation (MJJ) [a]	Accumulated annual soil water deficit
Mean winter precipitation (DJF)[b]	Accumulated annual soil water surplus
Mean summer water availability (MJJ) [a]	

[a] May, June, July
[b] December, January, February

The model is trained using existing empirical data on the European and North African (north of 15^{o}N) distributions of species to enable the full climate space of a species to be characterised and to ensure that the model does not extrapolate outside its training dataset when used to project the distribution of species under potential future climates in Europe. To improve performance, these variables, which can vary by several orders of magnitude, are first normalised to the range 0 to 1 using the minimum and maximum values for the European and North African region (Tarassenko, 1998) before proceeding with model training.

13.2 SPECIES model calibration and validation

The SPECIES model ANNs are calibrated and tested using an ensemble forecasting approach whereby projections are derived by constructing and training *multiple* ANNs for a single species (O'Hanley, 2007; 2009). The outputs from each of these models are then combined together in order to generate a final composite projection. There is a large body of statistical theory and practical work showing the superiority of ensembles over the use of any single model (Naftaly *et al*., 1997; Sharkey, 1999; Granitto *et al*., 2005).

Ensemble forecasting in SPECIES has been carried out using an aggregate *k*-fold cross-validation. This involves randomly subdividing the available data into training (70%) and validation (30%) sets *k* times in order to construct *k* different ANN sub-models. Each sub-model is calibrated on one of the training sets and then independently tested on the complementary validation set in order to calculate statistics indicating its performance accuracy. Each training and validation set is constructed so that it contains the same presence-to-absence ratio as seen in the full dataset, thus eliminating any representational bias in the datasets between presence and absence points. Bootstrapping is then used to construct training and validation sets which are the same size as the full dataset and containing equal numbers of presence and absence points. This ensures that the datasets have a 50/50 prevalence between presence and absence points, thereby reducing any sensitivity bias in the trained models towards projections of overly high or low suitability values.

A value of ten for *k* was chosen based on preliminary tests showing this value as giving a good trade-off between greater model stability / reduced spatial variance and longer model running times. An ensemble model output is then formed by combining the simulations from the ten ANN sub-models based on whether presence or absence is most commonly projected for a grid cell.

The performance of each ANN sub-model is statistically evaluated using Cohen's Kappa statistic of similarity (*K*) and the Area Under the Receiver Operating Characteristic Curve (AUC). Kappa is a commonly used statistic that provides a measure of similarity between spatial patterns, adjusted for chance agreement (Cohen, 1960). Kappa values vary from 0, indicating no agreement between observed and projected distributions, to 1 for perfect agreement and are dependent on the particular classification threshold being applied for determining whether simulated results are treated as presence or absence points. Maximum agreement for Kappa is calculated by iteratively adjusting this threshold from 0 to 1 in increments of $1x10^{-4}$. AUC is calculated from plots of the Receiver Operating Characteristic (ROC) curve. ROC curves measure the trade-off between a model's sensitivity (the proportion of true presences to the actual number of projected presences) and its false positive fraction (the proportion of false presences to the actual number of projected absences) as a function of all possible classification thresholds. This index is an unbiased measure of a model's predictive accuracy and is independent of both species prevalence in the validation dataset and classification threshold (Fielding & Bell, 1997). AUC ranges from 0.5 for models with no discrimination ability, to 1 for models with perfect discrimination.

The accuracy of the ensemble model, as measured by AUC and Kappa, is approximated by the average performance of the ten individual sub-models. This provides a conservative estimate of the ensemble's accuracy as its performance is usually at least as good as this and usually even better (Bishop, 1995). There are several rules-of-thumb available to help interpret measures of agreement between observed and simulated distributions. For example,

Monserud and Leemans (1992) suggest the following ranges of agreement for Kappa: excellent $K>0.85$; very good $0.7<K<0.85$; good $0.55<K<0.7$; fair $0.4<K<0.55$; and poor $K<0.4$. For AUC, Swets (1988) recommends interpreting values using the ranges: excellent AUC>0.90; good 0.80<AUC<0.90; fair 0.70<AUC<0.80; poor 0.60<AUC<0.70; fail AUC<0.60.

Models have been trained and validated for all 111 species and all show AUC statistics greater than 0.8, indicating good discrimination ability and 84% has AUC statistics greater than 0.9, indicating excellent model performance. Kappa values are slightly lower, but this is to be expected as the index ranges from 0 to 1. Values were greater than 0.7 for 47% of species indicating very good agreement between observed and simulated distributions, and between 0.4 and 0.7 for 35% of species indicating reasonable agreement. Further visual comparison between observed and simulated distributions was being undertaken (e.g. Figure 13.1) and any models that were unable to capture the core observed distribution were removed from the IAP.

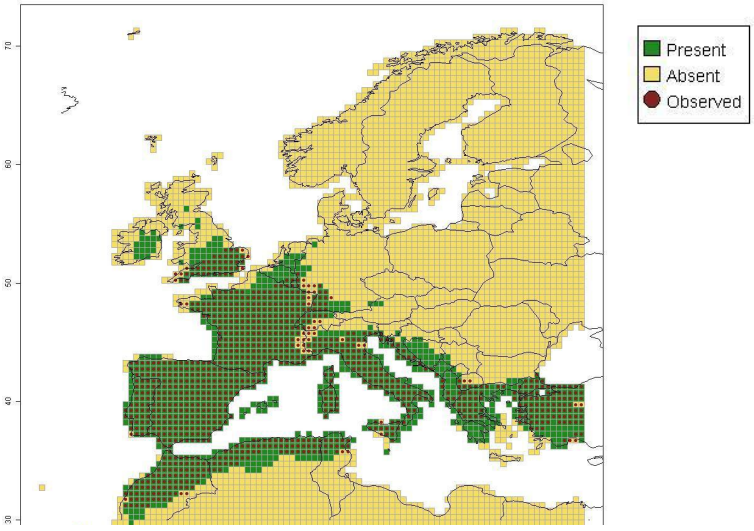


Figure 13.1: Comparison between the observed (red dots) and simulated (green grid cells) distribution of *Filago pyramidata* (broadleaved cudweed) for the baseline climate (1961-90).

Once a network is trained and validated for the European and North African region, it is then applied across the CLIMSAVE 10' European grid to produce a climate suitability surface. This is converted into a presence/absence distribution (see Figure 13.2) by applying the decision threshold which maximises agreement between observed and simulated distributions derived from the ROC curve. Further details concerning the definition of decision thresholds are provided in Pearson *et al.* (2002).

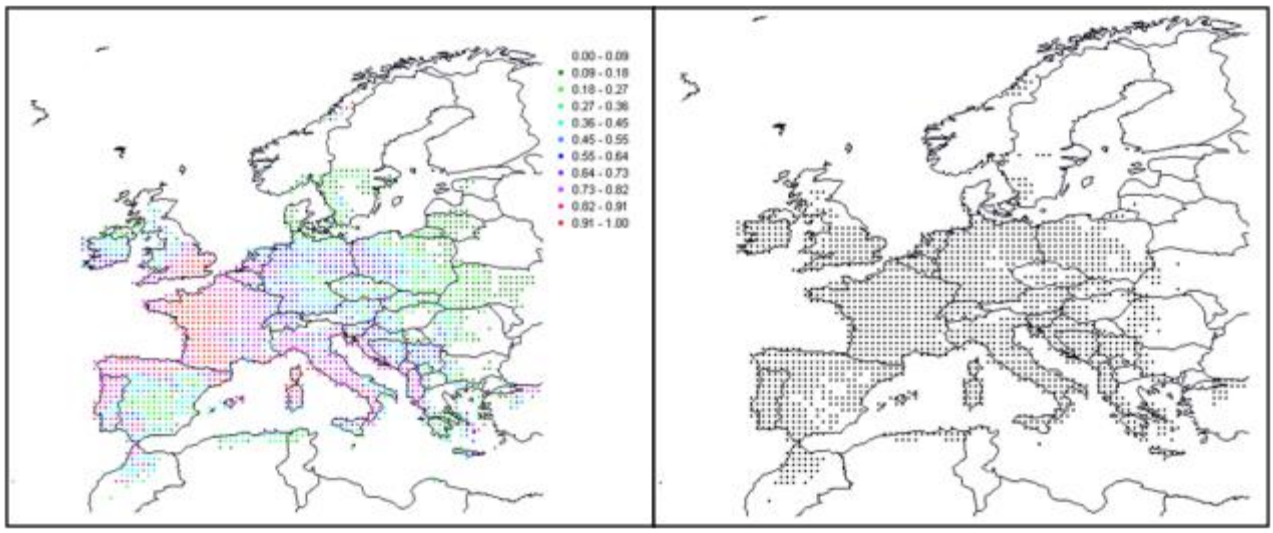


Figure 13.2: Illustrative results for *Silene gallica* (small-flowered catchfly) for Europe: (left) simulated climate suitability surface; (right) presence/absence distribution. Presence/absence ROC threshold is 0.22, AUC = 0.943, maximum Kappa = 0.64.

13.3 SPECIES model illustrative results

The user selects in the IAP interface a group of species, such as agricultural plant species, agricultural non-plant species, saltmarsh species or hunting species, which they would like to run. This cuts down the overall runtime for the IAP as running all 111 species at once would take several minutes. Each group contains between 3 and 10 species which enables the model to run more quickly.

13.3.1 *Species suitable climate space*

Figure 13.3 shows illustrative results for changes in the suitable climate space for *Filago pyramidata* (broadleaved cudweed; the baseline results are shown in Figure 13.1). With its predominantly Mediterranean and North African range, it is not surprising that current climate space is not threatened. The modelled polewards expansion of suitable climate space in the UK and into Germany and Scandinavia may offer potential for the cudweed's expansion, but it depends on the management of agricultural land and on whether it can disperse to suitable chalky or calcareous sites. The loss of climate suitability in Spain and parts of the eastern Mediterranean are likely to be due to increased temperatures and associated reductions in water availability, as while it is found on well-drained sites, moisture is important for seed germination.

13.3.2 *Species vulnerability*

Two species vulnerability indices are calculated from the outputs of the SPECIES model for the European region as a whole and for individual EU countries: vulnerability assuming no use of new climate space and vulnerability assuming full use of new climate space (Berry *et al.*, 2006). Each index is a function of the amount of change in suitable climate space, which is measured in terms of four species' indicators: new climate space; lost climate space;

overlap between present and future climate space; and size of the future distribution. Lost climate space indicates the sensitivity or degree of change. In reality, it is more likely that losses will be realised, as the species becomes stressed, less competitive and ultimately exhibits a mortality response. Gained or new climate space indicates the degree of opportunity for species to disperse and increase its range and thereby decrease its vulnerability. Overlap between present and future climate space indicates the continuity of suitable climate space. This measure also indicates the degree of threat to a species, as where there is little overlap between a species' current and potential future climate space, there could be a small population remaining *in situ* and the species will be forced to disperse if it is to realise much of its future climate space. Dispersal for some species is difficult and slow, thus they will become vulnerable. The size of the future distribution indicates the future rarity of species, as rarity is one factor thought to confer vulnerability to climate change (Berry, 2004).

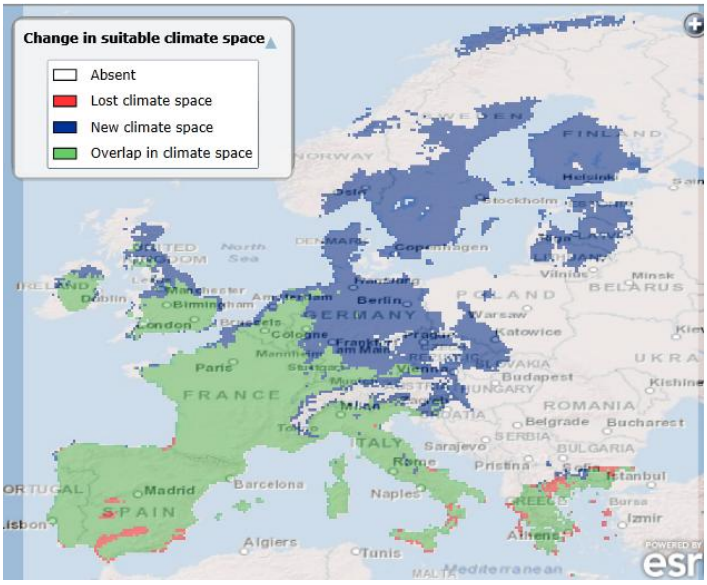


Figure 13.3: Illustrative results from the IAP for changes in suitable climate space for *Filago pyramidata* (broadleaved cudweed) under the climate scenario: CSMK3 climate model, A1 emissions scenario and middle climate sensitivity for the 2050s.

The Vulnerability Index with no use of new climate space assumes that autonomous adaptation is restricted to within the boundaries of the 10' grid cells which the species currently occupies, due to limited dispersal and there is no new planned adaptation. The Vulnerability Index with full use of new climate space assumes that autonomous and planned adaptation will take place to help species disperse into new areas. The degree to which planned adaptation can be implemented is assumed to be a function of the extent of new climate space, as this indicates the limit of the species' potential future distribution. Both indices range from 0 for no vulnerability to 20 for high vulnerability to climate change (Figure 13.4).

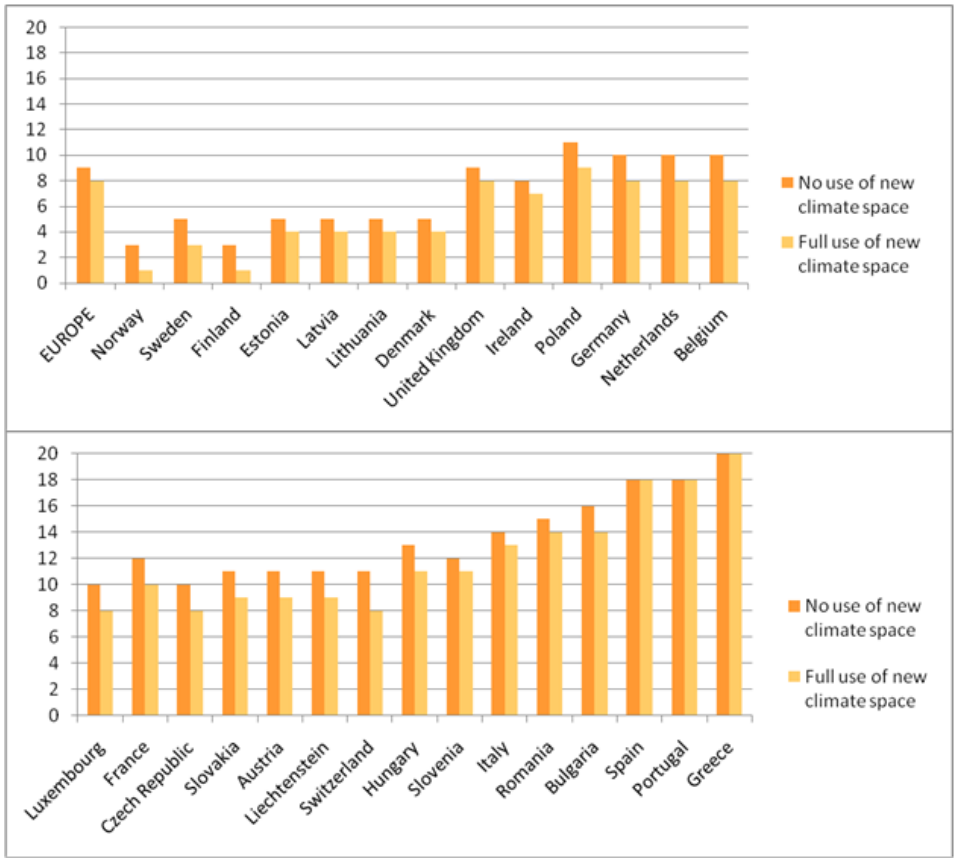


Figure 13.4: Illustrative output for the SPECIES vulnerability indices for Europe.

13.4 Integrating the SPECIES model with the other sectoral meta-models

13.4.1 *Agricultural and forest meta-models*

Predictions of potential climate space from the SPECIES model are combined with output on the area of arable and forest land, nitrogen and pesticide inputs and overwinter stubble from the SFARMOD land use model (Section 10) to simulate the impacts of climate and socio-economic changes on species' suitability in agricultural and forest habitats. The area of arable and forest land is used to create a habitat mask, which can optionally be applied to the species suitability maps. This habitat mask therefore alters with the climate and socio-economic scenarios depending on the spatial distribution of arable agriculture and forestry within the land use model.

The effects of nitrogen inputs on plant species are simulated by applying thresholds based on an individual species' sensitivity to nitrogen derived from the Ellenberg indicator values for Europe (Ellenberg, 1974; Ellenberg *et al.*, 1991). The various values were divided into classes indicating low, medium or high tolerance to nitrogen increases, as the Ellenberg values are on an arbitrary scale and species' ecological requirements may vary in different parts of their range and according to local conditions, thus a broad classification was appropriate. The species' nitrogen tolerances were linked to data on nitrogen inputs from the agricultural land use model based on results from Audsley *et al.* (2008) which attributed thresholds to the plant tolerance classes. Illustrative output on combining the effects of

nitrogen inputs with the SPECIES climate space outputs for the region of East Anglia in the UK is shown in Figure 13.3.

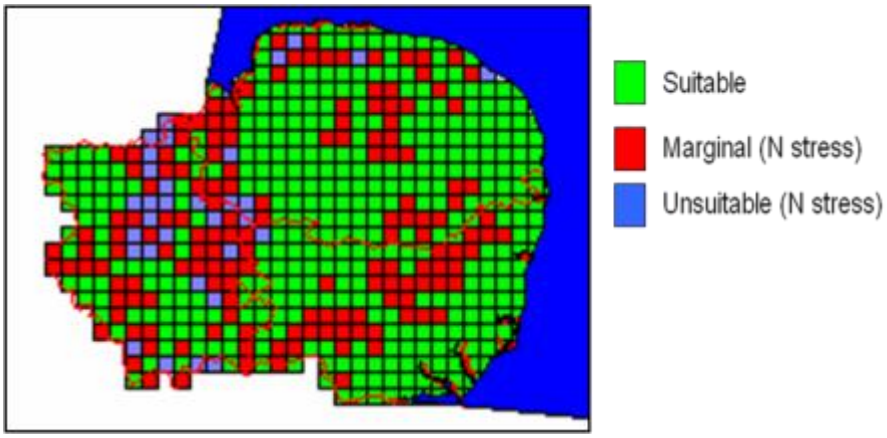


Figure 13.3: Sensitivity of *Scandix pecten-veneris* (shepard's needle) in East Anglia, UK derived from combining results from the SPECIES model with nitrogen fertilizer values from the SFARMOD agricultural land use model. Source: Audsley *et al.* (2008).

The effects of pesticide inputs on plants and pollinators are simulated by applying thresholds based on an individual species' sensitivity to pesticide derived from a literature review. The various values were divided into classes indicating low, medium or high tolerance to pesticides, as the species' tolerance may vary according to the pesticide type and thus a broad classification was appropriate. In each case, evidence for the highest level of sensitivity was used.

Overwinter stubble provides important habitat for ground nesting birds and can be an important food source during the winter. The effects of overwinter stubble on birds are simulated by applying thresholds based on an individual species' sensitivity to percentage changes from base in the amount of overwinter stubble per 10' grid cell from the agricultural land use model (Section 10).

13.4.2 *Water meta-model*

Predictions of potential climate space from the SPECIES model are combined with output on low and high river flows (Q95 and Q5 values, respectively) from the water model (WGMM- Section 8) and habitat data on wetlands from the flooding model (CFFLOOD- Section 7) to simulate the impacts of climate and socio-economic changes on species' suitability in wetland habitats. The area of inland wetlands is used to create a habitat mask, which can optionally be applied to the species suitability maps. This habitat mask therefore alters with the climate and socio-economic scenarios depending on the spatial distribution of wetlands within the CFFLOOD model.

The effects of low and high river flows on wetland species are simulated by applying thresholds based on an individual species' sensitivity to drought and waterlogging derived from Ellenberg indicator values (Ellenberg, 1974; Ellenberg et al., 1991). The various values were divided into classes indicating low, medium or high drought or flooding tolerance, as the indicator values are on an arbitrary scale and species' ecological requirements may vary in different parts of their range and according to local conditions and thus a broad

classification was appropriate. The species' water level requirements were linked to the outputs from the water model based on results from Harrison *et al.* (2008) which attributed thresholds to the plant tolerance classes. Illustrative output on combining the effects of water stress with the SPECIES climate space outputs for north-west England is shown in Figure 13.4.

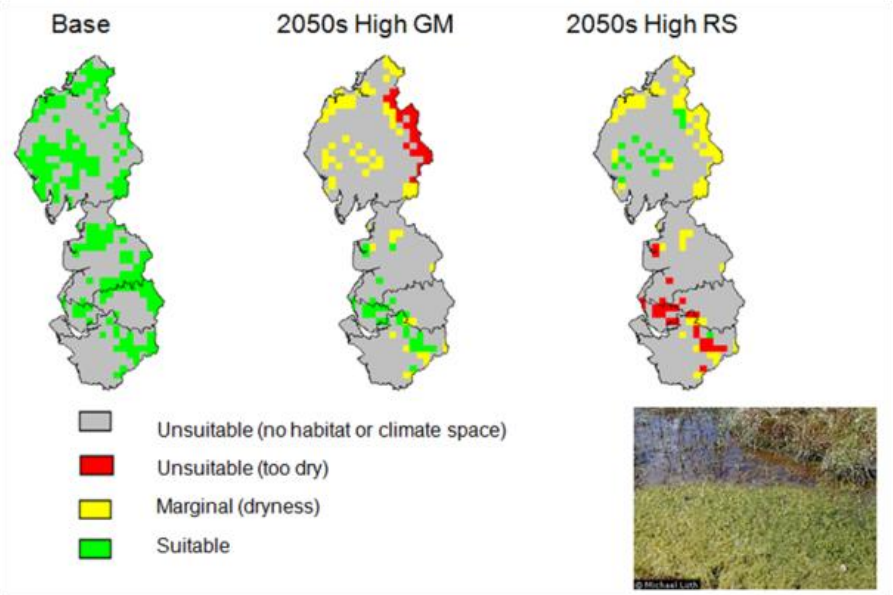


Figure 13.4: Sensitivity of *Sphagnum cuspidatum* in northwest England derived from combining results from the SPECIES model with Q95 values from a hydrological model. Source: Harrison et al. (2008).

13.4.3 *Coastal flooding meta-model*

Predictions of potential climate space from the SPECIES model are combined with output on the area of salt marsh and coastal and floodplain grazing marsh from the coastal model (CFFLOOD – Section 7) to simulate the impacts of climate and socio-economic changes on species' suitability. Changes in the area of salt marsh and coastal and floodplain grazing marsh simulated by the coastal model are directly overlaid onto the climate space simulations to create a habitat mask, which can optionally be applied to the suitability maps. This habitat mask therefore alters with the climate and socio-economic scenarios depending on the spatial distribution of these habitats within the coastal model.

13.4.4 *Habitat re-creation*

A habitat re-creation slider on the adaptation screen of the IAP allows the user to increase the percentage of protected areas (Natura 2000 sites). The percentage of land within a grid cell classified as a protected area is used as an input into the SFARMOD land use model (see Section 10), and thus affects the land use allocation and, hence, the habitat available for different species. There are three protected area sliders and 4 buttons on the IAP:

1. Protected Area changed (%): This determines how much protected area, relative to current day protected area, is allocated.
2. Change in protected area for forests.
3. Change in protected area for agriculture.
    a. Sliders (2) and (3) are used to determine to which land use the protected area is allocated (either forestry, agriculture or non-productive land, such as semi-natural

grassland, moors, bogs, etc). If the forest and agricultural sliders are less than 100% the remaining proportion is allocated to non-agriculturally productive land.

4. The buttons determine *how* protected area (PA) is allocated:
   a. Connectivity – allocates preferentially to areas with no existing PA (creates new sites).
   b. Buffering – allocates preferentially to areas with existing PA (enlarges existing sites).
   c. Buffering then connectivity – allocates all possible PA by buffering method, and any remaining PA by connectivity.
   d. Connectivity then buffering – allocates all possible PA by connectivity, and any remaining PA by buffering.

13.5 References

Audsley, E., Pearn, K.R., Harrison, P.A. and Berry, P.M. (2008). The impact of future socio-economic and climate changes on agricultural land use and the wider environment in East Anglia and North West England using a metamodel system. Climatic Change, 90, 57-88.

Berry, P.M. Rounsevell, M.D.A., Harrison, P.A. and Audsley, E. (2006). Assessing the vulnerability of agricultural land use and species to climate change and the role of policy in facilitating adaptation. *Environmental Science and Policy*, 9, 189-204.

Berry P.M. (2004). Plant vulnerability to climate change. In McGraw-Hill *Yearbook of Science and Technology*, New York, pp. 259-261.

Berry, P.M., Dawson, T.P., Harrison, P.A, Pearson, R.G. and Butt, N. (2003). The sensitivity and vulnerability of terrestrial habitats and species in Britain and Ireland to climate change. *Journal for Nature Conservation*, 11, 15-23.

Bishop, C.M. (1995). *Neural networks for pattern recognition*. Oxford University Press, Oxford.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37-46.

Ellenberg, H. (1974). Zeigerwerte der Gefässpflanzen Mitteleuropas. Scripta geobotanica, Vol. 9. Gottingen, 197 p.

Ellenberg, H., Weber, H.E., Dull, R., Wirth, V., Werner, W. and Paulisen, D. (1991). Zeigerwerte von Pflanzen in Mitteleuropa [Indicator values of plants in Central Europe]. Scripta Geobotanics, Vol. 18. Verlag Erich Goltze KG, Gottingen, 248 p.

Fielding, A.H., and Bell, J.F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24, 38-49.

Granitto, P.M., Verdes, P.F., and Ceccatto, H.A. (2005). Neural network ensembles, evaluation of aggregation algorithms. *Artificial Intellligence*, 163, 139-162.

Harrison, P.A., Berry, P.M., Henriques, C. and Holman, I.P. (2008). Impacts of socio-economic and climate change scenarios on wetlands: linking water resource and biodiversity meta-models. *Climatic Change*, 90, 113-139.

Harrison, P.A., Berry, P.M., Butt, N. And New, M. (2006). Modelling climate change impacts on species' distributions at the European scale: implications for conservation policy. Environmental Science and Policy, 9, 116-128.

Harrison, P.A., Vanhinsbergh, D.P., Fuller, R.J., and Berry, P.M. (2003). Modelling climate change impacts on the distribution of breeding birds in Britain and Ireland. *Journal for Nature Conservation*, 11, 31-42.

Monserud, R.A., and Leemans, R. (1992). Comparing global vegetation maps with the kappa statistic. *Ecological Modelling*, 62, 275-293.

Naftaly, U., Intrator, I. and Horn, D. (1997). Optimal ensemble averaging of neural networks. *Network: Computation in Neural Systems*, 8, 283-296.

O'Hanley, J.R. (2007). Automisation of the SPECIES model. In: P.M. Berry, J.R. O'Hanley, C.L. Thomson, P.A. Harrison, G.J. Masters and T.P. Dawson (Eds.) MONARCH (Modelling Natural Resource Responses to Climate Change) 3 Technical Report, Environmental Change Institute, University of Oxford, pp. 9-24.

O'Hanley, J.R. (2009). NeuralEnsembles: a neural network based ensemble forecasting program for habitat and bioclimatic suitability analysis. *Ecography*, 32, 89-93

Pearson, R.G., Dawson, T.P., Berry, P.M. and Harrison, P.A. (2002). SPECIES: a spatial evaluation of climate impact on the envelope of species. *Ecological Modelling*, 154, 289-300.

Sharkey, A.J.C. (Ed.). (1999). *Combining Artificial Neural Nets*, Springer, London.

Swets, K.A. (1988). Measuring the accuracy of diagnostic systems. *Science*, 240, 1285-1293.

Tarassenko, L. (1998). *A Guide to Neural Computing Applications*. Arnold, London, 139 pp.

14. Concluding remarks

Table 14.1 summarises the range of stakeholder-relevant impact indicators and indicators which translate the outputs from the integrated sectoral models into ecosystem services indicators which the meta-models each simulate. The impact and ecosystem service indicators listed in Table 14.1 may be subject to change, depending on the feedback received from stakeholders within the WP1 workshops.

The focus of activity within the next phase is to complete the implementation of the meta-models within the Platform (D2.3), ready for testing of the prototype Platform (M2.2).

Table 14.1: Summary of current sectoral output and potential ecosystem service indicators produced by the meta-model DLLs.

Sector	Meta-model DLL	Sectoral output indicators	Ecosystem Service indicators
Urban	RUG	• Artificial surfaces (area and % change) • Area of residential and non-residential areas	N/A
Snow	SnowCover	• Days with > 1 cm and > 10 cm snow	• Recreation (C)
Cropping	metaROIMPEL	• Crop yields (unlimited by nutrients and water availability; unlimited by nutrients availability; and limited by nutrients and water availability	N/A
Forestry	metaGOTILWA+	• Wood yield in managed forests	• Timber production (P) • C sequestration (R) • C balance (R) • Water storage in soils (R) • Naturalness, tranquillity, isolation (C)
Rural land use	metaSFARMOD	• Total crop production • Biomass energy • Food energy • Irrigation water demand • Intensively farmed, Forested and Abandoned land	• Food production (P) • Animal production (P) • Bioenergy production (P) • Fibre production (P) • Irrigation use (P) • Attractiveness of agricultural landscapes (C) • Naturalness (C)
Water	WGMM	• Naturalised high & average monthly river flow • Water availability • Water availability per capita • Real low, average and high flows • Water stress • Total water use	• Drinking water (P) • Cooling water (P) • Water storage (R) •
Flooding	CFFlood	• Area at risk of flooding • Damages caused by flooding • People affected by flooding • People in flood risk zones • Areas of coastal grazing marsh, salt marsh, intertidal flats and inland marshes	• Flood protection (R)
Pests	Pestmm	• Number of generations per season (6 species) • Ecoclimatic index (quality of the ecoclimatic niche for 6 species)	N/A
Biodiversity	SPECIES	• Species Presence/Absence • Species Vulnerability Indices	• Wild food plants (P) • Pollination (R) • Charismatic or iconic wildlife (C) • Species for hunting (C)
	metaLPJ-GUESS	• Net Primary Production (by Plant Functional Type, species and grid square) • Biomass (by Plant Functional Type, species and grid square)	• Biomass production (P) • Timber production (Provisioning) • Vegetation influence on local climate (Regulating) • Attenuation of runoff (Regulating) • Attractiveness of forest landscapes (C) • Charismatic or iconic wildlife (C)