International Environmental Modelling and Software Society (iEMSs)
2012 International Congress on Environmental Modelling and Software
Managing Resources of a Limited Planet, Sixth Biennial Meeting, Leipzig, Germany
R. Seppelt, A.A. Voinov, S. Lange, D. Bankamp (Eds.)
http://www.iemss.org/society/index.php/iemss-2012-proceedings

A EU27 Database of Daily Weather Data Derived from Climate Change Scenarios for Use with Crop Simulation Models

Donatelli, M.[1,2], Fumagalli D.[1], Zucchini A.[1], Duveiller G.[1], Nelson R.L.[3], Baruth B.[1]

[1]*Joint Research Centre, European Commission, Ispra (VA), Italy.*
[2]*CRA- Research Centre for Industrial Crops, Bologna- Italy*
[3]*Dep. Biological Systems Engineering, WSU, Pullman, WA, USA*
(marcello.donatelli@jrc.ec.europa.eu)

Abstract: Global Circulation Models (GCM) estimate future climate under scenarios of greenhouse gases emissions. Such estimates include several meteorological parameters but the two direct outputs are air temperature at earth surface and precipitation. The estimates are spatially downscaled using different methodologies, but it is accepted that such data require further processing for use with simulation models. Daily values of solar radiation, wind, air humidity, and at times rainfall may have absolute values which are not realistic, and/or the daily record of data may prove not to be consistent across meteorological variables. The final problem is related to the fact that GCM estimate the dynamics of climate, providing one instance of data per date in time series. Typically, crop models are deterministic and run in a stochastic fashion, hence requiring multiple years of weather data representing each time horizon of interest. Furthermore, if the time horizons of interest are very close (e.g. 2020 and 2030), sampling without overlap GCM outputs creates instability in means which may even show, in specific cases, apparent inversions of trends, creating artifacts also in the simulation via impact models. This paper presents a data base of daily weather data, with EU27 coverage at a 25 km grid, derived from the ENSEMBLES downscaling of the global circulation models HadCM3 and ECHAM5 realizations of the IPCC A1B emission scenario, in which solar radiation, wind and relative air humidity where estimated or collected from historical series, and derived variables reference evapotranspiration and vapour pressure deficit were estimated from other variables, ensuring consistency within daily records. Synthetic time series data were also generated using the weather generator ClimGen. All data are made available via web services in a data portal that also contain link to a reference domain ontology, and meta-information, according to the specifications for the semantic web publishing of the EU framework project TaToo.

Keywords: climate change; weather data; crop growth modeling, Europe,

1. Introduction

The basis for assessing potential impacts of climate change is future climate predictions. To obtain such predictions, it is necessary to have a reliable model of the climatic system and to use it to estimate possible future outcomes. A clear distinction has to be made between these two concepts: models, which are based on physical laws, and scenarios of green house gases emissions, which are a coherent, internally consistent and plausible description of a possible future state of the world. Climate change projections realized by running GCMs (Global Circulation/Climate Models) or RCMs (Regional Climate Models) under different emission scenarios are intrinsically subject to a significant amount of uncertainty.
Translating climate forecasts to estimate of impact on agriculture remains a challenge, due to the significant differences in spatial and temporal scales between GCMs and crop growth models (Hansen et al., 2006). Despite an increasing ability

of GCMs to successfully model present-day climate and provide realistic quantitative predictions of climate change at continental scale (IPCC, 2007b), they still have serious difficulties in reproducing accurate daily estimates at local scale. Even though GCMs operate at sub-daily scale, the spatial averaging at the coarse grid-scale distorts the temporal variability of daily weather sequences (Osborn and Hulme, 1997). This is especially true for precipitation. For instance, while a GCM may estimate monthly precipitation correctly, the daily precipitation may be spread throughout the month in a very unrealistic way (e.g. raining a little every day for example). Such distortions of daily weather variability can seriously bias crop model simulations (Semenov and Porter, 1995; Mearns et al., 1996; Hansen and Jones, 2000; Baron et al., 2005).
The objective of this paper is to present the realization of a dataset of weather data covering Europe at a grid 25 x 25 km, suitable for use with crop growth and more in general biophysical models, and made available via web services to non-profit users.

2. Materials and methods

2.1 Surface air temperature and precipitation

The need for bias correcting GCM-RCM projections for use by impact models is well known e.g. (Christensen et al., 2008), and the influence of such biases on hydrological and crop modelling has been extensively investigated by e.g.(Teutschbein and Seibert, 2010), who claimed that unless climate model outputs are corrected, their application to impact models may be unrealistic.
The source of climate data described in this paper is the bias-corrected ENSEMBLES dataset of Dosio and Paruolo (2011). Two realizations were selected, giving priority to the future projections of the A1B emission scenario given by HadCM3 GCM nested with the HadRM3 RCM (the realization is denoted as METO-HC-HadRM3Q0-HadCM3Q0). This represents a “warm” realization of the A1B emission scenario. The simulations based on another GCM, namely ECHAM5, coupled to the HIRHAM5 RCM for the downscaling, were also extracted to provide a milder, with respect to temperature, scenario. This can be considered as a “cold” realization (denoted DMI-HIRHAM5-ECHAM5). These two realizations of a single scenario are the extremes, in terms of surface temperature, within the ones analyzed in the ENSEMBLES project, allowing testing the largest uncertainty available in weather inputs to impact models. The target time horizons to build data were 2020, 2030, and 2050, which can be compared to a baseline centered on 2000. Therefore, given the two realizations (the “warm” based on HadCM3 and the “cold” based on ECHAM5), a total of 8 climate dataset could be made available to be used for the crop simulations.
The baseline period (1993-2007) data was built using the estimates available from the same scenarios of the same years. A comparison was made against data widely used to represent the baseline period chosen, as frequencies in classes representing the range of variability for air temperature and rainfall, not being possible a 1:1 comparison between scenario realizations and data based on observations. The reference weather data used were the Crop Growth Monitoring System (CGMS) weather database of the MARS unit of JRC, and the ECMWF re-run. Both A1B realizations matched acceptably (data not presented here) the reference data series with respect to temperature and precipitation.

2.2 Other weather variables

Crop and other biophysical simulation models require daily inputs of weather. A process-based crop simulation model can be very sensitive to weather inputs, not only as values, but also to consistency of the daily record. If data generation via downscaling is the result of an independent generation of weather variables, such consistency is not achieved. Also, if monthly mean values taken from GCM output are the basis for generating daily weather data, the resulting values do not

necessarily represent observed or known patterns for the variable of interest: As anticipated, the monthly mean of rainfall can be spread over the whole month, or it can be concentrated in few rainfall events. Another example is solar radiation: from monthly averages that can be considered correct, daily values might be derived which do not show the expected range of variation. Crop models are very sensitive to such differences, because of the time step used for simulation, and because the processes they simulate are non-linear.
Although the A1B realizations were initially assumed to be ready for use, a closer analysis of the dataset by Dosio and Paruolo (2011) revealed that part of it was inadequate to properly run process-based crop growth models to assess climate change impacts on yield. There were two different problems to solve. The first problem related to the lack of consistency of weather parameters, which results from the fact that the bias-correction is done on a subset of the necessary variables only, namely air temperature and rainfall. Other required variables, such as global solar radiation and wind speed, have unrealistic distributions when compared to observed data from the MARS-CGMS database or to simulated data from ECMWF over a past period of time. The second is related to sample size as articulated in the dedicated paragraph.

2.3 Global solar radiation

There is no evidence from GCMs that global solar radiation values at earth surface will change in the future. Global solar radiation was hence estimated using the auto-calibration procedure (Bojanowski and Donatelli, submitted) of the method Bristow-Campbell, which does not require reference data (i.e. recorded data of global solar radiation). The methods for estimating global solar radiation using daily air temperature range are based on the assumption that the site is not significantly affected by advection, which of course is not always the case. In case of an attempt to estimate the solar radiation pattern of a specific site, this assumption can be a strong limitation, but when working with abstractions such as interpolated time series associated to a spatial grid, the assumption can be considered non-limiting. This is because the range based method is physically based: clear days show a greater range of temperature because during the days solar irradiance is not filtered by clouds, and, during the night, the long wave emission from soil surface is more rapidly lost in the atmosphere. Also, seasonality is accounted for in the specific model. As described in the relevant paper, the auto-calibration method provides robust estimates of solar radiation, with the advantage of estimating a value which is consistent with temperature data. Given that scenarios of climate change as from GCM do estimate changes in temperature, the Bristow-Cambell b parameter is consequently estimated for each scenario, and solar radiation is estimated accordingly. Of course there are uncertainties on the temperature estimates of GCM and RCM, but it is out of scope of this application to articulate about the variables which are exogenous; however, data integration is done creating data records which are consistent at daily level, as required by crop models. Clear sky transmissivity was estimated for each grid cell from remote sensing data (Bojanowski and Donatelli, submitted), prior to the estimate of the b parameter, being the c parameter kept constant as c=2. Annual cumulated values of estimated solar radiation matched the ones of the reference CGMS and ECMWF dataset.

2.4 Wind and relative air humidity

Global circulation models do not produce estimates of either wind or air relative humidity. A conservative approach is to use historical series of such data, which only empirically can be associated, in general with a weak relationship at each site, to patterns of temperature and rainfall. However, the data to investigate such relationships are certainly not available for future climate scenarios; hence the conservative choice of using unchanged historical measurements was made. The measurements used were extracted by the CGMS weather database: the data of 1996-2005 were used both for the baseline and future scenarios. Wind and relative humidity are in case of direct interest for models for plant diseases.

2.5 Evapotranspiration and vapor pressure deficit

Reference evapotranspiration and vapour pressure deficit were estimated from the variables above using the Penman-Monteith as described in the FAO56 method, and implemented in the CLIMA libraries (Donatelli et al., 2006, 2009). A simpler method could have been chosen given the uncertainty on inputs, but given that no reference data is available, a more physically based model as Penman-Monteith was preferred to empirical models which would have generated data not necessarily consistent with other variables. Furthermore, an empirical model was not an option given no reference data to estimate its parameters.

2.6 Sample size

The time series produced via GCM (or RCM) runs represent the trends expected in climate variables such as temperature; however, there is a random component of variability around such a trend. For a given time horizon, climate studies will typically look at a sample of 30 years around that horizon to characterize a given variable or to derive other data (such as crop yields) from it. Such sample size is deemed large enough so that the short-term random fluctuations – such as daily weather variations – do not influence the outputs derived from the GCM simulations. Having a large sample size is also a reason why climate studies typically look at time horizons that are well separated in time, e.g. 2020, 2050 and 2100, so that the trend effect dominates over the random noise of the yearly weather (which can take values which are much different from the trend). If time horizons of interest are 2020, 2030, that is, close in time, taking windows of 30 years around these close time horizons would result in an overlap that renders the separation into two horizons meaningless. Conversely, when considering only 10 years (thereby avoiding overlap) the sample size becomes too small in order to assume that short-term weather fluctuations do not dominate over the trend. Indeed, 3 or 4 years which are much warmer than the trend during a period of 10 years will have stronger consequences on the average indicators of the crop simulations than if these 3-4 years occur within a period of 30 years. A stochastic weather generator, ClimGen (Stöckle et al., 2001), was hence used to increase the sample size corresponding to each time horizon. A set of 15 years from the GCM-RCM runs was used around each reference year (e.g. 2020 +/-7 years, so from 2013 to 2027), increasing the robustness of the estimate to characterize a time period. The weather generator uses these data to derive monthly parameters resuming the distribution of each weather variable for each grid cell. These parameters are then used to generate a set of 30 synthetic years for every grid cell, which have the characteristics of the 15-year period. Although the 15-year periods, used as source to generate parameters, overlap by 4 years across the time spans centered on the dates of interest, this is not a problem since the new synthetic years are different (although referred to the same weather) from the GCM-RCM ones. All the variables other than surface air temperature and precipitation were estimated / duplicated as done with the original series.
It must be noted that the weather generator is applied independently on every grid cell based on parameters defined for every grid cell individually. As a result, there is an apparent loss of spatial consistency of weather variables if a single synthetic year is considered for all cells. This is not true because each of the 30 years generated is a sample of weather data for that period and cannot be paired to individual years of adjacent cells. The spatial consistency is ensured when averaging the 30 years together and the mean weather parameters are observed. This also applies to variables of indicators derived from this synthetic dataset. The final crop simulations results are therefore based on 30 different runs for each time

horizon, and must be considered as possible outcomes for the considered period. The workflow to generate the dataset is shown in Fig. 1.

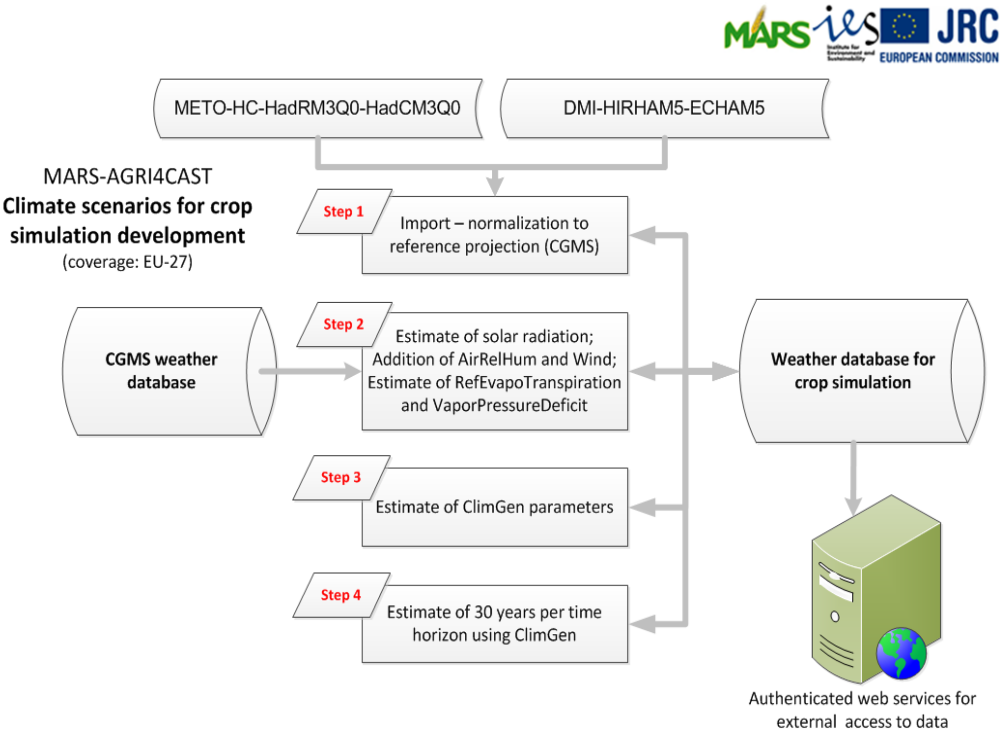

Figure 1. Outline of the processing of climate data starting from ENSEMBLES time series bias corrected by Dosio and Paruolo (2011).

3. Results and Discussion

The following maps show examples of the spatial distribution and of variability of A1B realizations.

3.1 Horizon 2030 under A1B scenario with HadCM3

The increase in temperature is marked over the entire continent and for both seasons (practically all of Europe is now at least 0.5°C above the corresponding temperature in 2000). The rise is higher for maximum temperature in summer and for minimum temperature in winter. With the exception of Scandinavia and the British Isles, European summers are considerably drier in 2030 than in 2020 and 2000. The notable increase in cumulated rainfall around the Italian peninsula disappears in summer but remains in winter. The region comprising Northern Spain and South-Western France which is drier in winter in 2020 becomes even drier in 2030 and extends geographically. Changes in cumulated potential evapotranspiration remain marginal in winter, but increase slightly throughout Europe (except in Scandinavia, British Isles and Northern Atlantic coast). It must be acknowledged that there is apparently a calculation artifact in the cumulated potential evapotranspiration difference map for summer: a horizontal strip going from West to East through Switzerland, Austria, Hungary and Romania. The source of it has to be further investigated given that no error in the data handling and generation was found.

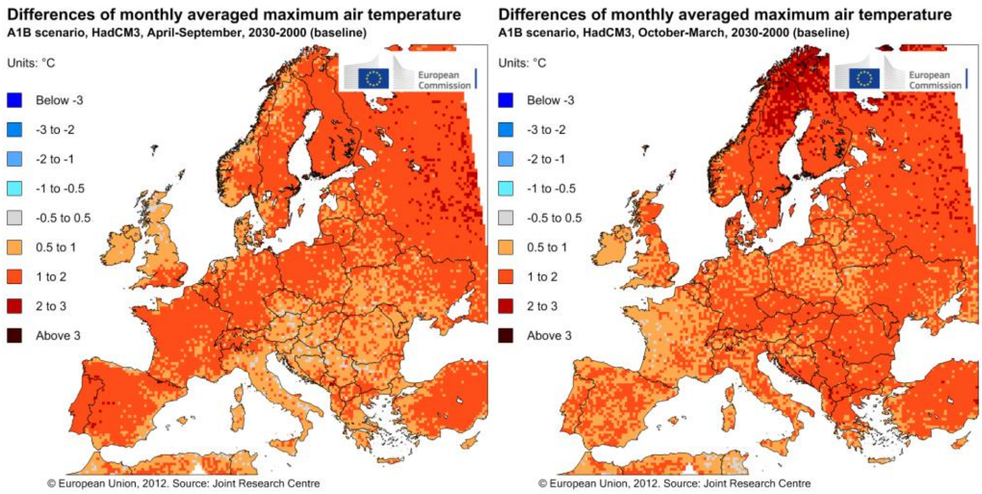

Figure 2. Difference of monthly averaged maximum temperature (HadCM3, A1B, 2030-2000) for April-September (left) and October-March (right)

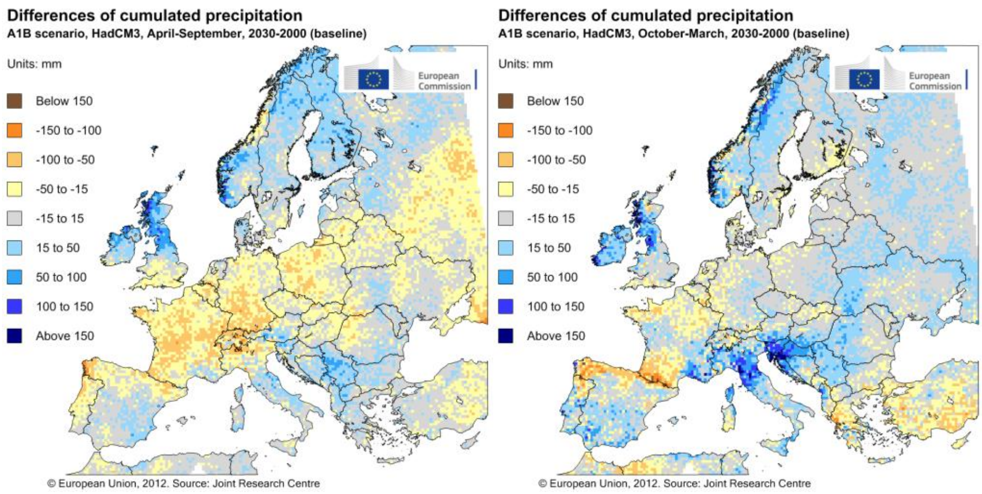

Figure 3. Difference of cumulated precipitation (HadCM3, A1B, 2030-2000) for April-September (left) and October-March (right)

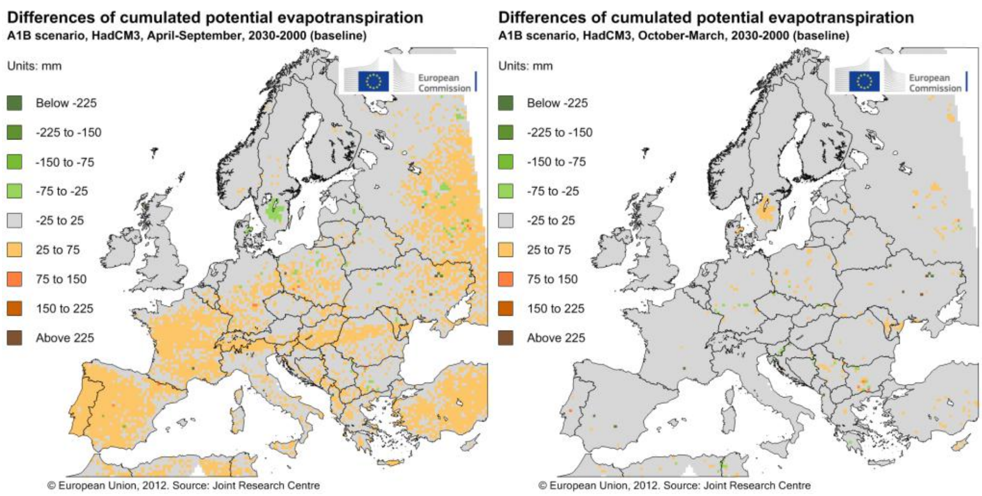

Figure 4. Difference of cumulated potential evapotranspiration (HadCM3,A1B, 2030-2000) for April-September (left) and October-March (right)

3.2 Horizon 2030 under A1B scenario with ECHAM5

The maps of precipitation are shown to allow comparing to the other realization of the A1B scenario. Precipitation increases strongly in France and to a lesser extent in Northern Europe during the cold period, while the warm period is drier than baseline (with the exception of Northern Italy and Scandinavia).

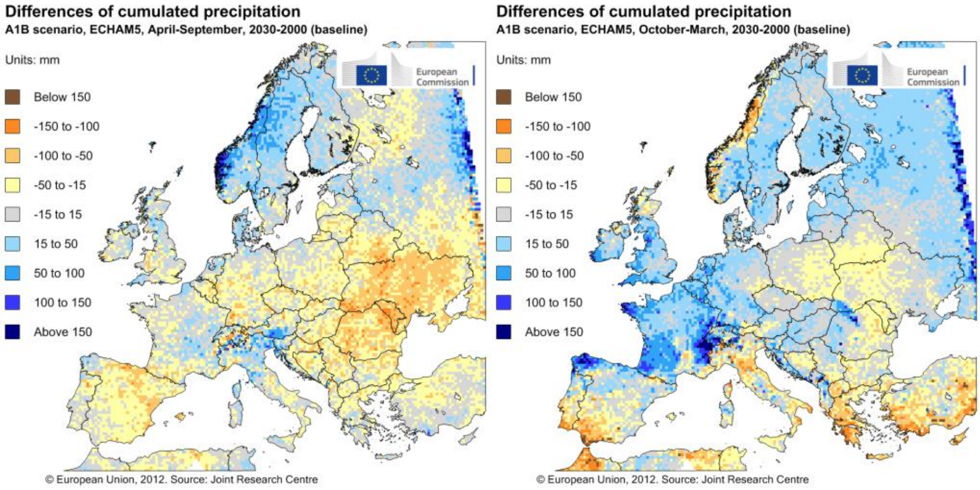

Figure 5. Difference of cumulated precipitation (ECHAM5, A1B, 2030-2000) for April-September (left) and October-March (right)

3.3 Data availability

Data are made available via web services using the SOAP protocol. Data access is given upon request to institutions providing the reference IP from which data will be accessed, and receiving an ID and password (data access is not granted to individuals). Sample applications (Microsoft .NET) are provided to exemplify data access. Data made available are time series as directly derived from the ENSEMBLES scenarios, and as synthetic time series generated using the ClimGen weather generator. The development of a portal including metadata, and the relevant domain ontology and RDF data descriptions is on-going.

4. Conclusions

Weather data have a large impact, as inputs, on the models used to make an impact assessment of climate change on agriculture. Different data processing to create such dataset may lead to different outputs, which would impact on simulation model results. Sharing a database covering EU27 potentially removes a source of variability in climate change and agriculture analyses. The process to build a dataset as the one developed has requested considerable resources, domain-specific knowledge and technological expertise. The data made available, which will be extended in the near future to other emission scenarios, provide a cost-free resource to public institutions.

Acknowledgements

Work carried out under the project AVEMAC of the European Commission, and partially granted by the project AgroScenari of the Italian Ministry for Agriculture, Food and Forestry Policies.

References

Baron, C., Sultan, B., Balme, M., Sarr, B., Traore, S., Lebel, T., Janicot, S., Dingkuhn, M., 2005. From GCM grid cell to agricultural plot: scale issues affecting modelling of climate impact. Philosophical Transactions of the Royal Society B: Biological Sciences 360, 2095.

Bojanowski, J., Donatelli, M., (submitted). Auto-calibration method of Bristow and Campbell solar radiation model. Environmental Modelling and Software.

Christensen, J.H., Boberg, F., Christensen, O.B., Lucas-Picher, P., 2008. On the need for bias correction of regional climate change projections of temperature and precipitation. Geophys. Res. Lett. 35, L20709.

Donatelli, M., Bellocchi, G., Carlini, L., 2006. Sharing knowledge via software components: Models on reference evapotranspiration. European Journal of Agronomy 24, 186-192.

Donatelli, M., Bellocchi, G., Habyarimana, E., Bregaglio, S., Confalonieri, R., Baruth, B., 2009. CLIMA: a weather generator framework, in: 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation, Cairns, Australia.

Donatelli, M., Tubiello, F., Stockle, C.O., Rosenzweig, C., Gristina, L., 1998. Simulating cropping system as affected by climate change and elevated CO2 in Italy, in: Proc. of the 7th ICCTA Conference. Presented at the ICCTA Conference, Firenze, Italy.

Dosio, A., Paruolo, P., 2011. Bias correction of the ENSEMBLES high-resolution climate change projections for use by impact models: Evaluation on the present climate. J. Geophys. Res. 116, D16106.

Hansen, J.W., Challinor, A., Ines, A., Wheeler, T., Moron, V., 2006. Translating climate forecasts into agricultural terms: advances and challenges. Climate Research 33, 27.

Hansen, J.W., Jones, J.W., 2000. Scaling-up crop models for climate variability applications. Agricultural Systems 65, 43–72.

IPCC, 2007a. Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.

Mearns, L.O., Rosenzweig, C., Goldberg, R., 1996. The effect of changes in daily and interannual climatic variability on CERES-Wheat: A sensitivity study. Climatic Change 32, 257–292.

Osborne, T.M., Lawrence, D.M., Challinor, A.J., Slingo, J.M., Wheeler, T.R., 2007. Development and assessment of a coupled crop–climate model. Global Change Biology 13, 169–183.

Semenov, M.A., Porter, J.R., 1995. Climatic variability and the modelling of crop yields. Biospheric Aspects of the Hydrological Cycle 73, 265–283.

Stöckle, C., Nelson, R., Donatelli, M., Castellvi, F., 2001. ClimGen: a flexible weather generation program, in: 2nd International Symposium Modelling Cropping Systems. Florence, Italy. pp. 16-18.

Teutschbein, C., Seibert, J., 2010. Regional Climate Models for Hydrological Impact Studies at the Catchment Scale: A Review of Recent Modeling Strategies. Geography Compass 4, 834-860.