Response to "Interactive comment on Reduction of predictive uncertainty in estimating irrigation water requirement through multi-model ensembles and ensemble averaging by S. Multsch et al."

S. Multsch, J.-F. Exbrayat, M. Kirby, N. R. Viney, H.-G. Frede and L. Breuer

Referee #1: The paper describes an attempt to study uncertainties in irrigation water requirements simulated for wheat growing in the Murray-Darling Basin (MDB) in Australia caused by the structure of the model and by the parameters used in it. In general the paper is well written and interesting but I have serious doubts that the setup of the study is useful to address the objectives described in the manuscript. My major points of criticisms are:

Referee #1: 1.) The authors use six empirical methods to calculate evapotranspiration in order to study the structural uncertainty of the model and five crop coefficient sets (needed to convert the ET of a reference crop to the one of a wheat crop) to study the parameter uncertainty in the model. Results of the so created ensemble of model runs (6 ET methods x 5 kc-sets) are weighted then according to their performance in representing measured data to derive a weighted ensemble mean and the ensemble range around the weighted mean. Such a setup is useful to compare results of complex models when the knowledge about the accuracy of the models is limited. However, the methods used to compute ET in this study have been extensively evaluated in previous research.

Authors: We agree that the ET methods we investigate have been widely used and evaluated in ecosystem model applications, but there is no consensus on which ET method is the best. This is particular the case, if (a) relevant input data are missing to drive a more sophisticated method, such as Penman-Monteith or Shuttleworth-Wallace. Moreover, if (b) the scale of the study is regional to global (see p.7529, lines 19-26), the ET method becomes a major source of uncertainty which is independent from the model complexity, because every model in hydrology, climate science and crop modelling relies on an estimate of ET and differences between methods have been reported by others (see page 7540, lines 5-16).

Despite this general knowledge, most studies use a single ET method, disregarding that ET simulations impact model predictions to a large extent. An objective method to evaluate model performance in combination with ensemble predictions could improve this limitation. The "Reliability Ensemble Averaging (REA)" technique is such a method, which has been also applied in hydrology to less complex models. We now use it for the first time to predict irrigation water requirements in this study. Similar to climate science, where REA has been developed, the accuracy of our model results is difficult to assess, as direct validation data are not available. We are able to show that REA leads to a decrease of model uncertainty, which is particular important in data scarce regions where common physically based methods such as the Penman-Monteith approach are limited in their application. We therefore think that the selected approach is valuable and provides interesting insight for modellers from a variety of research disciplines.

Referee #1: Based on comparisons with lysimeter measurements performed in different climatic environments it is well known that the Penman-Monteith (PM) method usually performs best when high quality measurements of all the required weather variables are available. The other ET-methods used in the study are less precise because they ignore some important weather variables or relationships between them determining site specific evapotranspiration. As such, I don't understand the value of applying additional inaccurate methods and of creating some "artificial" uncertainty. In other words: what is the additional value of the weighted ensemble mean as compared to the direct use of the PM method (maybe with site specific adjustment of the resistance terms)?

Authors: We agree that Penman-Monteith performs best. But in many model applications, e.g. simulations for irrigation management, catchment scale modelling or estimation of large scale water consumption (see p.7529, lines 19-26), other methods such as Priestly-Taylor and in particular Hargreaves-Samani are in use. These models give by far different results. We give an example of the "potential" uncertainty which arises if relevant information on climate data is missing which is a crucial information. But we also show that REA is a potential interesting approach to reduce this uncertainty, if more than one ET method is used in simulation studies.

Referee #1: 2.) To "evaluate" the performance of the ET methods the authors used class-A pan data measured at 34 stations (page 7532, lines 16-23). By doing this the authors evaluate evapotranspiration calculated for a reference crop (ET) by measured evaporation from an open water surface (E) which is certainly not the same. The agreement derived from this comparison is then used as a weighting parameter to compute the weighted ensemble mean. Again, I doubt that this weighted ensemble mean will represent an improvement to the direct use of the PM-method. But the authors can test this by comparing the performance of PM56 to the ensemble mean of the other methods.

Authors: For the application of REA, a comparison of model simulations and observations is needed to calculate the model performance criterion *(page 7535, lines 2-7 "…capability of each ensemble member to represent real world data by its bias.").* We could have treated PM56 as being an "observation" in the sense of a benchmark model. However, we think that a more independent test is more appropriate in the sense of REA and therefore decided to use those observations that are at hand: class-A pan observations. To account for the difference of class-A pan evaporation and reference crop ET, we used a commonly applied correction factor (pan-coefficients according to McMahon et al. (2013)) to derive crop ET from class-A pan measurements. Most often, ET estimates are not compared to any measurements at all, leaving modelers with no information on how good their model application is. We therefore think that a comparison to class-A pan is for sure not perfect, but better than no testing at all. This will be acknowledged in a revised version of the paper.

Referee #1: 3.) Parameter uncertainty is evaluated by using 5 sets of crop coefficients. These coefficients relate the ET of a wheat crop to the ET of the reference crop surface. The factors describing the differences between the ET of wheat and the one of the reference grass surface are described in detail in Allen et al. (1998), for example. Methods to reduce uncertainties in crop coefficients would be to (i) adjust standard crop coefficients by considering the local conditions (wheat management, wetting interval, aridity, growing period length) or (ii) using a process based crop model that directly accounts for the underlying processes. APSIM, for example, has been developed in Australia and was frequently applied for the local conditions. I doubt that the set of the crop coefficients used in this study really provides a representative picture on the expected parameter uncertainty.

Authors: We are aware that there are more site specific and regionally adapted Kc values even if crop coefficients are in first instance meant to adjust ETo to a specific crop type, reflecting albedo, crop height, surface resistance, soil evaporation (Allen et al., 1998). But such already mBut in contrast to the widely existing assumption, that better adapted Kc values lead to improved crop ET estimations, we show that the uncertainty related to the choice of Kc is small compared to the uncertainty inherent to the ET model structure (or method) in itself. We highlight the calculation of uncertainty of irrigation water requirement in this manuscript and a method how to reduce it. The importance of the consideration of uncertainty has also been reported elsewhere as stated in the manuscript (page 7543 lines 21-25): "*Despite the growing importance of IRR for today's agriculture and the effect on surface (Hoekstra et al., 2012) and groundwater (Wada et al., 2010) resources, few studies have dealt with the predictive uncertainty of this requirement (e.g. Wada et al., 2013) and how to reduce it*."

Referee #1: 4.) While the authors focus on potential uncertainties caused by the ET calculation method and the crop coefficients, there is little explanation why these two factors were selected and how some other factors may affect the uncertainties in irrigation water requirement calculated in this study.
The model applied here uses some very crude assumptions (e.g. that runoff is fixed to 20% of precipitation, see equation 2). In addition, it does not account for the spatial heterogeneity in soil or crop conditions. From this perspective it's hard to see what readers can learn from the results and what can be generalized for other sites, models and investigated factors.

Authors: The straight forward single crop coefficient concept has been recently applied in various studies (page 7528 lines 27-28, page 7529 lines 1-14). The focus is drawn on crop coefficient parameters and ET methods, as both are crucial features in assessing irrigation water requirements as is already described on page 7528, 15-26.

Authors: Maybe we were not clear enough in describing the scope of the study, and we will certainly address this better in a revised version of the manuscript. We fully agree with the reviewer that there are crude assumptions in some of the ET methods we applied. However, these ET methods are used worldwide in many simulation studies, without any considerations to improve the methods (e.g. the 20% precipitation reduction by runoff in the CROPWAT model). That is why we implemented them as given. In almost all studies, researcher use only one ET method with a single, often spatially independent Kc set. As a result, some scientist ask to at least use better adapted, local Kc sets. However, we show in our work that for any large scale studies, the uncertainty introduced by Kc parameterization is small compared to the uncertainty introduced by the ET method. Of course, this is not the case for any local model application, where crop conditions and spatial heterogeneity needs to be considered. But this is not done in large scale model applications.

Authors: Both factors, i.e. crop coefficient and evapotranspiration, have been reported to be important for the performance of models based on the singe crop coefficient concept as reported by others (see 7540 5-19; 7542 1-15) and we address this part of uncertainty in this study.
The fixed fraction of runoff is adapted from the default setting of the Cropwat model.

Referee #1: Page 7527, lines 9-11: "We find that structural model uncertainty is far more important than model parametric uncertainty to estimate irrigation water requirement." Please notice that only one parameter was tested. Therefore this conclusion is to general.

Authors: Will be rewritten to *"We find that structural model uncertainty among reference ET is far more important than model parametric uncertainty introduced by crop coefficients. These* *crop coefficients are used to estimate irrigation water requirement following the single crop coefficient approach."*

Referee #1: Page 7527, lines 16-18: "We conclude that multi-model ensemble predictions and sophisticated model averaging techniques are helpful in predicting irrigation demand and provide relevant information for decision making." To support this conclusion it is required to show the additional value of the multi-model ensemble predictions, as compared for example to a single application of the Penman-Monteith method. I can still not see it here.

Authors: We disagree in this point. As explained in our rebuttal to comment 2). Using REA, we show that we are able to reduce the predictive uncertainty by considering a number of "uncertain" single models.

Referee #1: Page 7527, lines 21-25: "Globally, the proportion of fresh water consumption by agriculture is large (9087 km3 yr-1) (Hoekstra and Mekonnen, 2012) and is projected to increase in the future in order to support the increasing world population. More precisely, most of the change in freshwater consumption will arise from the increasing irrigation demand by crops (De Fraiture and Wichelns, 2010)." It's required to be more precise. The first figure on fresh water use refers to the sum of irrigation water and natural rainfall while the second statement refers to irrigation only. That future irrigation water requirements will increase is not sure. Models accounting for the reduction in transpiration due to increased atmospheric CO2 concentration show constant or even declining trends. Therefore this section does not reflect the state of knowledge.

Authors: A likely effect of changes of atmospheric CO2 concentration is not part of this study. But even by considering the biophysiological effect of reduced transpiration, the need for additional irrigation water is very likely in the future. This is mainly driven by demographic development and changes in food diets. Accordingly, we will rewrite the passage as follows: *"Globally, the proportion of fresh water consumption by agriculture from rainfall as well as surface and groundwater resources is large (9087 km3 yr-1)* (Hoekstra and Mekonnen, 2012)*. It is projected that water demand is increasing in the future, particular by irrigation agriculture, in order to support the increasing world population with food* (Foley et al., 2011; De Fraiture and Wichelns, 2010; Hanjra and Qureshi, 2010; Wada and Bierkens, 2014)*.*

Referee #1: Pages 7527-7531 (introduction): The authors describe here what they have done in the paper but the objectives remain unclear. Is the objective to quantify uncertainties in irrigation water estimates in models of the same type or is it to develop and present a new method for uncertainty assessment? How does this study compare to all these crop model comparisons published within the last 2-3 years? Wouldn't it be better to replace the crop coefficient approach by a real simulation of crop growth instead of just applying different sets of kc-values with unknown representativeness?

Authors: We might have not been clear enough with the objectives of our study, which we certainly will improve in a revised version of our manuscript. The study is more than a simple model intercomparison, for which a number of studies have been published in the past years and which are cited in our manuscript. We go beyond a simple intercomparison: how can we derive better predictions by using an ensemble of well-known ET methods and which are the likely causes of predictive uncertainty in ET estimations. We are convinced, that ensemble modeling could overcome some of the shortcomings of today's global estimations of water resources, given the large uncertainties in ET estimation.

Authors: Regarding the concern raised in relation to the Kc approach we argue, that the Kc approach is widely applied across many model applications in regions worldwide, in particular for predicting, e.g., irrigation requirements, global water resources, groundwater depletion, water footprint, virtual water trade. The advantage of this approach is that it can be used in regions where less data are available where the application of a comprehensive crop model is not possible. This approach has also been applied for a number of studies in the Murray-Darling Basin (Barton and Meyer, 2005; Harris, 2002; Hughes, 1999; Meyer, 1999). Even though we know that the Kc approach has limitations and that real simulations of crop growth would improve predictions, it remains unlikely that this will be happening on the macroscale.

Referee #1: Page 7531, lines 15-17: "The applicability of six different ETo methods is evaluated by using available measured class-A-pan evaporation measurements of 34 stations in the MDB over a 21 years time period" ET is evaluated with E => does not seem to be very useful

Authors: In order to make class-A pan measurements comparable with reference ETo one has to use pan coefficients (Allen et al., 1998). We converted class-A pan evaporation with pan-coefficients published by McMahon et al. (2013) which are given in a monthly resolution at 68 sites across Australia. Please see page 12 lines 13-16 in the manuscript for further details: *"Pan evaporation differs from evaporation from a cropped surface through a different albedo, heat storage and humidity above the surface. For this reason, the class-A pan data have been adjusted with monthly pan coefficients (McMahon et al., 2013) to better compare them with ET_o simulations of open surface waters. On an annual average, class-A pan evaporation of 1,558 mm yr^{-1} were reduced by 9% to 1,422 mm yr^{-1} across all stations."*

Referee #1: Page 7532, section 2.1 Study site and data: What about uncertainty in input data (e.g. land use, weather) and their interaction with model structure? Uncertainties in humidity and wind speed will likely affect PM but not some other methods like Hargreaves or Priestley-Taylor

Authors: We are glad that the reviewer agrees with us that an accounting of the uncertainty behind ET estimation is complex and includes many sources. A full accounting of the global uncertainty in a spatial context of ET estimation would be for sure interesting, but not achievable at the moment; though on the long term it is highly needed. To our knowledge, our work is one of the few studies that takes a closer look at a part of this uncertainty in the field of macroscale irrigation requirement studies. We focus on two important sources of uncertainty, which have been reported to be relevant for predicting irrigation requirements (Howell et al., 2004; Siebert and Döll, 2010; da Silva et al., 2013). The other sources of uncertainty, i.e. land use, weather and many others, are also important but not in the particular scope of this study.

Referee #1: Page 7533, equation (2): Which data or findings support the very basic assumption that 80% of total precipitation becomes effective?

Authors: The fixed fraction of runoff is adapted from the default setting of the CROPWAT model according to the FAO56 guidelines. It was not our intention to improve any of the ET methods, but rather apply them as given.

Referee #1: Page 7536, lines 17-18: "The median daily ETo for APET is 3.6 mm d-1, PM56 3.9 mm d-1, HS 3.8 mm d-1, PPET 5.2 mm d-1, PT 6.4 mm d-1 and TURC 3.4 mm d-1." Please check the calculation routine and the underlying data for the calculations with Priestley-Taylor. An overestimate in the here reported range is very unlikely and not supported by the previous literature!

Authors: We will again check the amount of the ET predicted by the PT method and include this in a revised version of the paper.

Literature

Allen, R. G., Pereira, L. S., Raes, D. and Smith, M.: Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56, FAO, Rome, 300, 6541, 1998.

Barton, A. B. and Meyer, W. S.: Crop Coefficients - A Benchmark for Use in Australia, Cooperative Research Centre for Irrigation Futures, (CRC IF Technical Report No. xxxx/05), 2005.

Foley, J. A., Ramankutty, N., Brauman, K. A., Cassidy, E. S., Gerber, J. S., Johnston, M., Mueller, N. D., O'Connell, C., Ray, D. K., West, P. C., Balzer, C., Bennett, E. M., Carpenter, S. R., Hill, J., Monfreda, C., Polasky, S., Rockström, J., Sheehan, J., Siebert, S., Tilman, D. and Zaks, D. P. M.: Solutions for a cultivated planet, Nature, 478(7369), 337–342, doi:10.1038/nature10452, 2011.

De Fraiture, C. and Wichelns, D.: Satisfying future water demands for agriculture, Agricultural Water Management, 97(4), 502–511, 2010.

Hanjra, M. A. and Qureshi, M. E.: Global water crisis and future food security in an era of climate change, Food Policy, 35(5), 365–377, doi:10.1016/j.foodpol.2010.05.006, 2010.

Harris, G. A.: Irrigation: water balance scheduling, Queensland Department of Primary Industries and Fisheries, (DPI Note FSO546), 2002.

Hoekstra, A. Y. and Mekonnen, M. M.: The water footprint of humanity, P. Natl. Acad. Sci. USA, 109(9), 3232–3237, 2012.

Hoekstra, A. Y., Mekonnen, M. M., Chapagain, A. K., Mathews, R. E. and Richter, B. D.: Global monthly water scarcity: blue water footprints versus blue water availability, Plos One, 7(2), e32688, 2012.

Howell, T. A., Evett, S. R., Tolk, J. A. and Schneider, A. D.: Evapotranspiration of full-, deficit-irrigated, and dryland cotton on the Northern Texas High Plains, Journal of irrigation and drainage engineering, 130(4), 277–285, 2004.

Hughes, J. D.: Southern Irrigation SOILpak. For irrigated broad area agriculture on the Riverine Plain in the Murray and Murrumbidgee valleys., NSW Agriculture, Orange, 1999.

McMahon, T. A., Peel, M. C., Lowe, L., Srikanthan, R. and McVicar, T. R.: Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: a pragmatic synthesis, Hydrol. Earth Syst. Sci., 17, 1331–1369, 2013.

Meyer, W. S.: Standard reference evaporation calculation for inland, south eastern Australia, CSIRO Land and Water., 1999.

Siebert, S. and Döll, P.: Quantifying blue and green virtual water contents in global crop production as well as potential production losses without irrigation, Journal of Hydrology, 384(3), 198–217, 2010.

Da Silva, V. de P., da Silva, B. B., Albuquerque, W. G., Borges, C. J., de Sousa, I. F. and Neto, J. D.: Crop coefficient, water requirements, yield and water use efficiency of sugarcane growth in Brazil, Agr. Water Manage., 128, 102–109, 2013.

Wada Y., van Beek L. P., van Kempen C. M., Reckman J. W., Vasak S. and Bierkens M. F.: Global depletion of groundwater resources , Geophys. Res. Lett, 37(20), L20402, 2010.
DOI: 10.1029/2010GL044571

Wada, Y., Wisser, D., Eisner, S., Flörke, M., Gerten, D., Haddeland, I., Hanasaki, N., Masaki, Y., Portmann, F. T. and Stacke, T.: Multimodel projections and uncertainties of irrigation water demand under climate change, Geophys. Res. Lett., 40(17), 4626–4632, 2013.

Wada, Y. and Bierkens, M. F. P.: Sustainability of global water use: past reconstruction and future projections, Environ. Res. Lett., 9(10), 104003, doi:10.1088/1748-9326/9/10/104003, 2014