Geosci. Model Dev. Discuss.,
doi:10.5194/gmd-2015-273-AC2, 2016

Geoscientific Model Development Discussions

Interactive comment on "ClimateLearn: A machine-learning approach for climate prediction using network measures" *by* Q. Y. Feng et al.

Q. Y. Feng et al.

q.feng@uu.nl

Received and published: 17 September 2016

1. *It is not clear then what this toolbox offers to other researchers interested in applying ML to climate data, how this toolbox compares to existing libraries of ML algorithms, and what set of applications it might be applicable to.*

   The reviewer is right that relatively few details on the software were provided. In the new version of the paper, the capabilities of the `ClimateLearn` package are now described in section 3.3.

2. *There is, of course, value in replication of existing results, but the description in this paper falls far short of a useful replication, as it does not compare in any detail either the methods or the results with previous work. Most of the choices*

made in applying the techniques are presented with no rationale, and no analysis of the impact on the results. For example, why a 3×3 ANN in the first study and a 2×1 in the second? What difference do these choices make? Why a prediction lead time of 12 months (On this point, the paper says similar to Ludescher et al. What does similar mean? Is it an exact replication, and if not, what changes were introduced and why?). How was the data filtering done, and how does it impact the results?

The main purpose of replicating Ludescher et al. (2014) is to show that our toolbox can provide better predictions without subjective decisions like the choice of thresholds, which is the foundation of the predicting method provided by Ludescher et al. (2014). Therefore, we keep using the same dataset over the same period, the same prediction lead time $\tau = 12$ months, and the same chronological way of presenting results as in Ludescher et al. (2014). Please note, we only replicated Ludescher et al. (2014) in section 4.1 for forecasting the occurrence of El Niño events, which has 9 attributes for the ANN, while in section 4.2 we focus on forecasting the development of NINO3.4 index based on Feng and Dijkstra (2016), which only has 3 attributes for the ANN. This is the reason that we have different configurations of ANNs for these two studies. The result in Fig. 3a is without filtering and by eliminating the isolated and transient events, and batching the adjacent events together, we obtained the filtered result in Fig. 3b. As will be better explained in the revised paper, the filtering is an important part of the method, crucial for the classification task.

3. *The paper places way too much emphasis on the algorithms and the training procedure, as if these alone were responsible for useful results, and largely neglects the role of prior knowledge in pre-processing the data, selecting candidate variables to include in the models, and expert interpretation of the results. The role of domain expertise is hinted at throughout the description of*

the method, as the authors draw on existing work to select candidate variables (e.g. skewness and wind stress residuals, as analyzed by Feng and Dijkstra). Surely if there is value in this work it is in the investigation of why these variables are good candidates for model finding, rather than the generation of yet another predictive model from them via a machine learning algorithm?

We agree and in the revised text, we have emphasised the importance of prior knowledge in pre-processing the data using network approaches. `ClimateLearn` therefore provides an important tool way to investigate why network measures are useful variables for climate prediction.

Interactive comment on Geosci. Model Dev. Discuss., doi:10.5194/gmd-2015-273, 2016.