Review comments on the MS Spatial patterns of North Atlantic Oscillation influence on mass balance variability of European Glaciers.

This work is a good contribution into connecting mass balance studies of glaciers into a larger picture in continental scale. Work like this present manuscript is needed for the modeling community to better understand the effects on mass balance on large regional climatic signals, and paves the way for better predictions of the future of glacial mass balances. The authors uses three different techniques to mimic / model mass balance of a series of European mass balance records, where two complete statistical methods is used and one model that is based on climatological data. The authors achieves very good correlations between their models and the measured data, which in itself arguments for the credibility of the present work.

Comments on the details of the MS

1. The results the authors get when comparing their modeled data with observation are very good, which gives their efforts good credibility. Despite this, I have a few methodological questions on the model part:

a). Was the termini elevation regarded as fixed over the time domain? Some glaciers have almost 50 years of time series, and have retreated, so termini elevation has changed. How was this treated in the modeling? Was the termini position a static point?
b). The vertical data in the CRU grids were used to calculate lapse rates, was this used to get the surface air temperature (SAT) at the glacier termini? What is the vertical resolution in the CRU data, and is that vertical resolution enough to get the SAT needed?
c). Further, is the SAT given exact where ice meets bare ground, or where exact is the point denoting termini elevation. The ice fronts are sometimes like a cliff, and the glacier close to front, and the ground ca differ 50 m or so….
d). How was the precipitation calculated? Normally massbalance models use elevation bands to integrate stepwise change in T and P (e.g Radic, V. and R.Hock. 2011. Regionally differentiated contribution of mountain glaciers and ice caps to future sea-level rise. Nat.Geosc., 4(Feb), 91-94). But, as I understand this MS one point on the glacier was used to derive Pi (Solid precipitation). Is this correct? In that case, which elevation was used to calculate Pi for the glacier? Perhaps elevations bandswas used, but it is not well described in the text. If you used a single point, like the termini for Pi you must argument why you used it, since Pi probably differs quite a bit over the elevation span.
e). If Pi is solid precipitation only, did you not regard water retention from rain fall, and refreezing melt water in the firn pack in the income of the mass balance. Probably not, but in case, argument why you did not use it. In the simplest form water retention can be used in a simple parameterization scheme, eg the Radic and Hock reference above.
f). The best way to get precipitation is probably to do a dynamically downscale gridded data. My experience is that gridded data often underestimate precipitation close to large topographical object, i.e. underestimate the orographic component of precipitation. This modeling is though out of the scoop of this present work, but may be a remark worth to bring in mind.
g) on page 10, li 10 and thereabouts it is a discussion of how to increase the skill score, and correlations between model and data. One thing we must remember is that the measured data still is an estimate of reality. To measure the exact mass balance is a difficult task on most glaciers, and almost an impossible task in a few settings. By this I think we must always understand that there is a deviation of measured mass balance and real mass balance, giving that modeled mass balance could even be a better proxy for real mass balance than measured mass balance in a few (lucky) cases.
h) A final comment I have is that the models likely would give even better performance if daily data is used instead of the very crude average over a whole month; but data handling and availability in this specific data base probably puts the frame of this issue. Perhaps a few words can be used to discuss possible improvement in future work using daily data from example ERA40, or ERA :Interim.

2. The results and comparisons of model data with NAO clearly show how NAO has a regional influence on the mass balance of the European glaciers, with some regions better related, and some less to NAO. My comment here is why NAO is used over the summer months? NAO is not well defined over JJA, due to the flatter pressure field in the summer months (See Hurrell, 1995). This is further shown by the authors by finding DJF has better scores than the full year in their tests. In an older study it was

found that regarding Scandinavian glaciers both winter and summer balances were better correlated to a pressure index they called Norwegian Sea Index, being regionally derived between the pressure centers over about West Siberia/Barents Sea and British Isles (Pohjola, V.A. and J.C.Rogers. 1997. Atmospheric circulation and variations in Scandinavian mass balance. *Quaternary Research,* 47(1), 29-36). Could a development of the present study be expanded into refining the study into more regional pressure centers, instead of the more "global" NAO?

Detailed comments:

P2. Abstract li2. Was it really 7735 individual time series, or 7735 modeled mass balance years? Perhaps best to take out this number, it is not necessary, and not described (as far as I could see) in the further text.

P5. The mean model is not described very well. Perhaps add a line or two about it.

P11. I did not understand li10-15, how did you get a r = 03?

Table 1 and 2. You need to explain better in head what the parameters in the tables represent. I guess the tables is the comparison model to measurements? Is MB in years? In Table 1 is rows 1-2 individually trained, and rows 3-7 the mean model?

In general
*Use r2 instead of r, it is a better measure of skill.