Reply to the Interactive comment on “Regional seesaw between North Atlantic and Nordic Seas during the last glacial abrupt climate events” by Mélanie Wary et al., by Anonymous Referee #3, Received and published: 26 March 2017

RC3: Wary et al present reconstruction of surface water conditions in the North Atlantic region during stage three, mainly derived from dinocyst assemblages using transfer functions. The results are compared with climate model hosing experiments and show evidence of an inverse relationship between temperatures in the Nordic Seas and the North Atlantic Ocean/Greenland. The manuscript is well written but I there are some issues that need to be addressed before it should be accepted.

Reply: We thank Anonymous Referee #3 for his/her careful review of our paper. We will take into account all his/her precious comments for the revised version of the manuscript. Below are our replies.

When I read a manuscript that uses transfer functions, I like to start with the raw assemblage data. I was disappointed that the manuscript does not include these, but I found data for two of the four cores examined in Eynaud et al (2002) and partial assemblage data for a further core in Wary et al (2016). This is true that most of the data are already presented in published papers: MD95-2009 and MD95-2010 data in Eynaud et al. (2002), MD99-2281 data in Zumaque et al. (2012), and MD99-2285 partial data (35-41 ka BP) in Wary et al. (2016). Complete data for MD99-2285 core should published soon (Wary et al., The southern Norwegian Sea during the last 45 ka: hydrographical reorganizations under changing ice-sheet dynamics, Journal of Quaternary Science, in press). All these papers mention where these data are available and can be downloaded. We did not present the raw data here (except for the relative percentage of Islandinium minutum) because of that, but we can for sure add them in Supporting Information and/or indicate the repository where they are available.

Although both of these papers are cited for information about foraminifera assemblages and chronologies, neither is apparently cited for the dinocyst assemblage data. Both these publication also include transfer function derived estimates of sea surface conditions and make similar findings to the present manuscript. The lack of citations to this earlier, overlapping work makes the present manuscript appear more novel than is justified: this must be rectified by citing the authors’ previous work appropriately and explicitly stating which parts of the proxy data in the present manuscript are new. The dinocyst stratigraphies that are not already published should be included in the supplementary material.

We agree and will transfer that information from the Supporting Information to the main text.

From the dinocyst assemblages, the manuscript reconstructs summer and winter sea surface temperature and salinities, and sea ice duration using transfer functions. Seasonality is inferred from the difference between summer and winter temperatures. The reported transfer function performances are all impressive, however, these are leave-one-out estimates which, as has been shown repeatedly (Telford 2006; Telford & Birks 2005, 2009, 2011), severely underestimate the true uncertainty in the reconstruction. This is because the environmental variables in the dinocyst calibration set are spatially autocorrelated, violating one of the basic
assumptions of transfer functions (Birks et al. 2010). There are cross-validation schemes that are more robust to spatial autocorrelation (Trachsel and Telford 2016): performance statistics from these should be used instead.

It is likely that with a robust cross-validation scheme, the transfer function performance statistics will appear worse and some variables will have little or no skill. I suspect that salinity models are the weakest and that it will be difficult to make independent reconstructions of sea ice duration or winter temperature as both have strong non-linear relationships with summer temperature. Without knowing how large the uncertainty is, the reader cannot evaluate how meaningful the stadial-interstadial difference temperature is.

Repeated parallel studies, also using cross-validation schemes, have also shown that “Although strong spatial autocorrelation characterizes the original climate parameter distribution, the results show that the spatial structure of data has relatively low effect on the calculation of the error of prediction” (Guiot and de Vernal, 2011a, citation from their abstract), concluding that “until a higher performance transfer function approach is developed […] we can only encourage the paleoclimate community to continue using MAT, with all statistical precautions required, as it has been successful for the documenting recent Earth’s climate dynamics” (Guiot and de Vernal, 2011b, citation from their conclusion). These studies concerned sea-surface temperature and salinity reconstructions derived from the application of MAT transfer function to dinocyst assemblages. Later, other studies also based on validation exercises showed again the prediction power of the method for sea-ice reconstructions (see Quaternary Science Reviews special issue 79 (2013), especially de Vernal et al., 2013a,b).

As the application of MAT transfer functions to dinocyst assemblages has already been the subject of discussion in CPD, our reply here only provides the main conclusions of works testifying of the reliability and robustness of this method; we refer to the reader to this previous discussion (Milzer et al., 2014: http://www.clim-past.net/10/305/2014/cp-10-305-2014-discussion.html) for further details.

Neither the manuscript nor the precursor papers include any reconstruction diagnostics, such as distance to nearest analogue, which would help the reader evaluate whether the reconstructions can be relied upon.

We can for sure add these data in Supporting Information.

The manuscript needs to make the inclusion criteria for the hosing models explicit. Swingedouw et al (2013) includes six models, but only five are used now. The omitted model is BCM2, which has the opposite temperature response in the Nordic Seas to the other models.

We will move that discussion from the Supporting Information section to the main text.

The combination of the warm dinocyst-inferred surface temperatures and cold planktic foraminifera inferred sub-surface temperatures in the stadials raise some questions. Firstly, why do sub-polar planktic foraminifera not inhabit the surface layer. RC2 raised a similar same point. We suggested that the subpolar planktonic foraminifera did not inhabit the surface layer because SSS were apparently too low (summer $SSS_{dino}$ between 30.9 and 31.6 on average) according to these species tolerances (e.g., Tolderlund and Bé 1971). We will add some discussion about that in the revised version of our manuscript even if it is already extensively discussed in our next JQS contribution.

Secondly, do the models suggest such a thin surface layer.
According to our data and interpretations, the maximum thickness of this surface layer should be less than ~ 250 m water depth, i.e. the maximum depth of *N. pachyderma* s. habitat reported for this area (e.g. Simstich et al., 2003). In the model simulations, the thickness of the surface layer is quite variable between the different individual simulations (see Figure 9 in Swingedouw et al., 2013, enclosed below), varying from ~ 150 m water depth (HadCM3, IPSLCM5) to ~ 750 m (MPI-ESM, EC-Earth), which allows our hypotheses.

![Figure 9 in Swingedouw et al., 2013](image)

# Minor points
Tables 2 and S4 claim to present anomalies, but appear to be the actual reconstructions. If they are anomalies, the baseline needs to be specified.

The term anomaly (GS conditions minus GI conditions) is here used to echo the simulation anomalies (hosing experiments minus control experiments), as the aim is to directly confront the reaction of the system between hosing experiment and mean GS conditions, relatively to control experiment and mean GI conditions taken as respective baselines. We can alternatively use the term “differences” for both.
References: