Ocean Sci. Discuss., 3, S679–S680, 2006
www.ocean-sci-discuss.net/3/S679/2006/

Ocean Science Discussions

Interactive comment on "Towards measuring the meridional overturning circulation from space" *by* D. Cromwell et al.

D. Cromwell et al.

Received and published: 10 November 2006

The reviewer has 'a few fundamental problems [...] with the statistical model and the interpretation of the prediction'. We will show below that these problems are unfounded.

The first problem the reviewer has is overfitting. Overfitting occurs when the model we use not only fits the signal we are trying to extract but also the noise. This often occurs, as the reviewer points out, when too large a number of parameters is fitted. The classic example is fitting a polynomial of high order. We are very conscious of the problems of overfitting. For this reason we did not fit the model to the entire dataset and rely upon the regression diagnostics to tell us how good a fit we have. Instead, we split the dataset into two, fitting to the first half and using the second to test the predictive power of the model. The only real test of such a model is the mean square error predicting on an independent dataset. This is the method we use. Overfitting would result in very

poor predictions in this circumstance. The quality of our predictions, we believe, shows that we have not 'overfitted' the data.

The second problem the reviewer has is with the quality of the prediction. He/she claims that a one-parameter fit - the mean MOC value - would give a better fit. This is a rather artificial example as we have no means of estimating the mean MOC in practice, but let's stick with it. If we knew the mean MOC in the second half of the series we would have a mean square error of 0.48 Sv^2. This is lower than our best result of 0.63 Sv^2. But this is a totally unrealistic situation. Any predictor must be able to predict the magnitude of the MOC as well as its variability from year to year. If we use only the training set, as in our linear model, then the mean square error from using the mean MOC is 3 Sv^2. So our models easily outperform the past mean. The mean square errors from the two means are very different. This is because the magnitude of the signal has changed between the training set and the test set. Our models capture this fairly well. As a metric, mean square error is better than error variance or correlation because it includes the magnitude of the signal, which the others don't as they assume zero mean.

We will incorporate the above remarks in the revision. To sum up: we believe that the quality of our predictions on an independent dataset shows that we have not been guilty of overfitting; and a single value such as the mean of the training set performs very badly as a predictor, thus showing the predictive power of our model.

Interactive comment on Ocean Sci. Discuss., 3, 1623, 2006.