Atmos. Chem. Phys. Discuss., 12, C5247–C5251, 2012
www.atmos-chem-phys-discuss.net/12/C5247/2012/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "A new El Niño-Southern Oscillation forecasting tool based on Southern Oscillation Index" *by* C. A. Varotsos and C. Tzanis

Anonymous Referee #1

Received and published: 1 August 2012

General comments:

This discussion paper is a revised version of an article on a statistical method for prediction of the Southern Oscillation Index (SOI). As the instrumental observed SOI record is relatively long (from 1876) for monthly climate data, it is a popular choice for testing statistical prediction methods. Here the method is based on an analysis of the observed monthly-average SOI record, making use of a 'natural time' re-ordering and calculation of a predictor based on entropy contained in data windows of various lengths preceding the predictand. A 'receiver operating characteristics' (ROC) measure is used to assess the skill of prediction of the SOI in the next month and also at 24 months lead. Assessment is for the skill of prediction of SOI values more negative than two particular negative threshold levels. (Negative SOI is associated with the El Nino phase of the El Nino Southern Oscillation (ENSO) cycle.) While new methods for SOI prediction are of high practical interest, as described in the introduction, the merit of this method is difficult to judge without rigorous comparison with the leading existing methods. In this case no such evidence is provided, and without such evidence it is not possible to recommend publication of this article. Physical explanation of the mechanism for skill obtained is also highly desirable, but again no such explanation is provided. In the abstract the claim is made that 'This finding improves the accuracy of the short-term prediction models of the ENSO extreme events': there is no evidence provided to support this claim.

Specific comments:

From the example (Fig. 1) added in revision to explain 'natural time', it is clear that for the present application, using an SOI record with regularly-spaced monthly values (and no missing data), the 'natural time' is effectively a re-labelling of the calendar time, which should have no effect on the predictive skill. In this case 'natural time' simply serves to provide various weights for the data within the window used for calculating a predictor. (Fig. 1 is helpful, but somewhat misleading in that the conventional time data are shown as irregularly spaced, whereas for SOI they are regularly spaced.)

The calculation of a windowed statistic is described in section 3. The entropy-based statistic S is effectively a non-linear weighted average of data values in the window, intended as a measure of 'disorder' in the windowed data. For $S_{\_}$ the weighting is reversed, and the predictor used is the difference $\Delta S = S - S_{\_}$.

Aside: page 4 line 29: I do not understand how the 'normalized intensity' of an event is related to the 'probability to observe it'.

In Fig. 2 ΔS is shown for a window length of 36 months, along with the SOI timeseries. It is not clear what point is being made here, as the text just describes the figure. (It is also not clear if here the window precedes each SOI value by one month, or if the window is centred on the corresponding SOI values shown: this should be clarified.)

ACPD
12, C5247–C5251, 2012

Interactive Comment

The relation between ΔS and SOI would be clearer if a smaller subset of the timeseries was shown.

In section 4 the method is applied to SOI prediction, using windows of various lengths preceding the data points to be predicted. (The authors need to state more clearly that the windows precede the data points.) Note that the ΔS is not a directly predicted SOI value, but rather is some statistic whose relation to SOI values is to be determined.

That relation is explored as follows. Given some (negative) threshold SOI value T, choose a predictor value ($\Delta S0$ say). Consider months when the the SOI is less than or equal to T. If for some month the predictor value ΔS is larger than $\Delta S0$ then the prediction is true (a hit), otherwise it is a miss. Thus hit rate and false alarm rates can be calculated for each $\Delta S0$, and used to construct ROC graphs. (The explanation is not clear in the text: an example would be helpful, e.g. provide numbers of hits, misses and rates for an example point on one of the curves in Fig. 3.) (It would be helpful if the number of cases with SOI less than or equal to T could be provided for each T considered.)

Fig. 3 shows ROC graphs for various window lengths and threshold values. (Note: the target thresholds are two negative values of SOI, which roughly correspond to weak and strong El Nino events: this aspect should be emphasised here and in the abstract.) The results in Fig. 3 are for a lead time of zero months: i.e. the predictand is SOI in the month immediately following each predictor window. The authors note that skill is best for windows of 20 and 24 month lengths. It would be helpful if the values of $\Delta S0$ used to construct the points on the curves could be indicated on the graphs, at least for a subset. (Note that, as can be seen in Figs. 5 and 6, the SOI timeseries contains a substantial persistent component, so at this lead time values well above the 'no skill' diagonal in Fig. 3 could be expected for a simple persistence prediction scheme.)

At this point the onus is on the authors to demonstrate that the levels of skill obtained by this method are an advance on other methods. However no comparison is made,

ACPD
12, C5247–C5251, 2012

Interactive Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

without which the merit of the new scheme is unknown.

Comparison is made difficult by the choice of ROC and particular thresholds for the skill measure. Although ROC is sometimes used in this field, I am not aware of the same ROC and threshold combination being used elsewhere in the extensive ENSO and SOI prediction literature. The choice of a zero month lead monthly SOI value as the predictand is also unusual. For ENSO prediction, practical value largely lies in predictive skill at longer ranges. (At zero month lead, it is usually evident that an ENSO event is underway.) Thus most SOI/ENSO prediction schemes provide skill information for a range of lead times, typically several months to several seasons ahead, and are assessed on their performance over such ranges. To reduce the effects of month-to-month variability, it is also common to average the predictand over, say, 3 months when assessing long-range performance. Possibilities that the authors could pursue readily to aid comparison are: (a) consider other measures. Although the emphasis here is on specific thresholds, comparison would be aided by the addition of other simple statistics that are widely used for assessment. (b) consider other methods. Comparisons could be made by applying their measure to predictions with other statistical methods (e.g. optimal normals that make use of simple averages over various window lengths). (c) consider other SOI predictands – e.g. SOI at various ranges up to seasons ahead.

(Note: seasonal dependence of skill can be substantial and is also of interest to the ENSO prediction community, and this aspect could also easily be investigated.)

Case studies of the major 82/83 and 97/98 El Nino episodes are discussed (Figs. 5,6). The need for a 'time increased probability' switch is unclear: it just seems to be a threshold indicator. The choice of $\Delta S0$ corresponding to a 50% false alarm rate is unclear and needs explanation. It is unjustified to claim prediction 'well in advance' based on results for zero month lead times.

Information about skill at a lead time of 24 months has been added in revision. Fig. 7 shows ROC statistics for predictors with 40 and 48 month windows, and for a simple predictor which is the SOI monthly value 24 months before. Modest skill is obtained, providing some support to the claim of ‘prediction well in advance’ for the T=-15 threshold. It is not clear though that the new method outperforms the simple predictor for all the window/threshold combinations presented, and the significance of the results is also untested. As with the zero month lead results, further comparison with other methods is essential to assess the merit of the new method.

I did not understand the explanation in the last paragraph of section 4, involving the difference in disorder looking in the immediate future versus the immediate past and its behaviour approaching ’critical points’.

An aside: I would be curious to know to what extent ΔS is better than S or S_- as a predictor: there is no mention of this aspect.

Interactive comment on Atmos. Chem. Phys. Discuss., 12, 17443, 2012.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper