11B.3 AUGUST-SEPTEMBER ENSO PREDICTION SKILL 1959-2001: A COMPARISON BETWEEN FOUR STATE-OF-THE-ART SEASONAL MODELS

Adam S. Lea*, Mark A. Saunders and Benjamin Lloyd-Hughes
Benfield Hazard Research Centre, University College London, UK

1. INTRODUCTION

A prime challenge for ENSO seasonal forecast models is to predict boreal summer ENSO conditions at lead. August-September ENSO has a strong influence on Atlantic hurricane activity, Northwest Pacific typhoon activity and tropical precipitation. However, summer ENSO skill is low due to the predictability barrier in boreal spring between March and May (Torrence and Webster, 1998). During 2003/4 four new state-of-the-art ENSO seasonal prediction models have become available providing extended (>40-year) hindcast sets for each ENSO index region. These models are the coupled dynamical models from ECMWF, the Met Office and Meteo France which are contributing to the European DEMETER project (Palmer et al., 2004), and the statistical consolidated ENSO-CLIPER model comprising the ensemble of 18 model variants of the statistical ENSO-CLIPER (CLImatology and PERsistence) prediction model (Lloyd-Hughes et al., 2004).

Here we assess and compare the hindcast skill of these four models in predicting August-September sea surface temperature (SST) for each of the main ENSO index regions (3.4, 3 and 4) for the 43-year period 1959-2001, this being the period of available DEMETER hindcasts. We also examine the improvement in hindcast skill obtained by combining hindcasts from the individual models into a multi-model hindcast. Only leads of 0 and 3 months (corresponding to hindcasts made at the start of August and the beginning of May) are considered, as these are the leads at which the DEMETER integrations provide hindcasts for August-September.

2. DATA AND BIAS REMOVAL

Verifications are made against three different historical monthly SST data sets. These are the NCEP/NCAR reanalysis, the ERA40 reanalysis (www.ecmwf.int/research/era), and the US Climate Prediction Center (CPC) optimal interpolation SST data set.

The output from the dynamical models is post-processed to remove model bias error. The bias correction factor is computed by regressing the model output against observations in a cross-validation framework with 5-year block removal (WMO, 2002). The block is tapered at the time series ends. This procedure is repeated for each lead time and SST data set.

***** *Corresponding author address*: Department of Space and Climate Physics, University College London, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK; e-mail: al@mssl.ucl.ac.uk.

The consolidated ENSO-CLIPER hindcasts are obtained using cross-validation with 5-year block removal (WMO, 2002) and block tapering at the time series ends. Block elimination is employed to minimise potential skill inflation which might arise from the multi-annual persistence of ENSO conditions. NCEP/NCAR SST data for 1948-2001, ERA40 SST data for 1958-2001 and CPC SST data for 1950-2001 are employed in the consolidated CLIPER model building with hindcasts being retained for 1959-2001.

3. SKILL SCORE

August-September ENSO hindcast skill is assessed using the deterministic skill measure recommended by the World Meteorological Organisation for verification of seasonal forecasts (WMO 2002). This measure, called the mean square skill score, MSSS, is the percentage improvement in mean square error over a climatological hindcast. MSSS is a robust skill measure which is immune to the bias problems associated with other measures. Positive (negative) skill indicates that the model is doing better (worse) than climatology. The 43-year (1959-2001) average is used for climatology here.

The MSSS skill from persistence is included for reference. Persistence is calculated over the same length interval as the predictand period (WMO, 2002). For example, ordinary persistence at a lead of 3 months for the August-September target predictand is calculated as the mean anomaly over the prior two-month period March-April. The MSSS skill for the ensemble mean (the mean of the four hindcast models) is also computed and displayed below.

4. HINDCAST SKILL COMPARISON 1959-2001

(a) Niño 3.4 with NCEP/NCAR SSTs

Lead	Mean Square Skill Score (MSSS)					
	Pers	MF	MO	EC	CL	ALL
0	84	90	84	90	89	93
3	-30	55	50	36	36	59

(b) Niño 3.4 with ERA40 SSTs

Lead	Mean Square Skill Score (MSSS)					
	Pers	MF	MO	EC	CL	ALL
0	84	88	83	89	88	92
3	-29	54	51	36	28	57

(c) Niño 3.4 with CPC SSTs

Lead	Mean Square Skill Score (MSSS)					
	Pers	MF	MO	EC	CL	ALL
0	84	89	85	90	89	93
3	-29	56	51	35	39	58

(d) Niño 3 with NCEP/NCAR SSTs

Lead	Mean Square Skill Score (MSSS)					
	Pers	MF	MO	EC	CL	ALL
0	85	92	91	92	89	95
3	-8	51	55	34	29	56

(e) Niño 3 with ERA40 SSTs

Lead	Mean Square Skill Score (MSSS)					
	Pers	MF	MO	EC	CL	ALL
0	79	92	90	92	87	94
3	9	51	56	33	37	57

(f) Niño 3 with CPC SSTs

Lead	Mean Square Skill Score (MSSS)					
	Pers	MF	MO	EC	CL	ALL
0	84	92	91	92	88	94
3	-5	52	55	31	39	56

(g) Niño 4 with NCEP/NCAR SSTs

Lead	Mean Square Skill Score (MSSS)					
	Pers	MF	MO	EC	CL	ALL
0	81	89	81	83	89	90
3	3	50	40	23	56	54

(h) Niño 4 with ERA40 SSTs

Lead	Mean Square Skill Score (MSSS)					
	Pers	MF	MO	EC	CL	ALL
0	84	83	77	80	88	86
3	-5	45	35	23	44	48

(i) Niño 4 with CPC SSTs

Lead	Mean Square Skill Score (MSSS)					
	Pers	MF	MO	EC	CL	ALL
0	79	89	83	85	85	93
3	-5	54	43	26	44	56

Table 1. Mean square skill scores for predicting the August-September Niño 3.4, Niño 3 and Niño 4 SSTs 1959-2001. Verification is against the NCEP/NCAR, ERA40 and CPC data. 'Pers' is the persistence forecast, 'MF' is the Meteo France model, 'MO' is the UK Met Office model, 'EC' is the ECMWF model, 'CL' is the consolidated CLIPER model and 'ALL' represents the ensemble of all four models.

5. SUMMARY

The main conclusions from Table 1 are first that all the models provide at least a 20-25% improvement in mean square error compared to climatology in predicting summer ENSO at a lead 3 months. Second, the 'ALL' hindcast provides an absolute improvement in MSSS over persistence of ~10% and ~50-60% at leads of 0 and 3 months respectively. Third, for Niño 3.4 and Niño 3 the Meteo France and Met Office model's perform best giving MSSS values of 50-55% at lead 3. For Niño 4 the consolidated CLIPER and Meteo France models do best giving MSSS values of 45-55% at lead 3. Combining the four model hindcasts into an ensemble provides, in eight of the nine cases at both leads 0 and 3, an MSSS greater than the highest MSSS from an individual model. Further work will investigate the benefits of ensemble ENSO hindcasts to seasonal predictions of hurricane and typhoon activity.

Acknowledgements. This work is funded by Benfield, Royal & Sun Alliance and Crawford & Company. We thank NCEP/NCAR, ECMWF, the U.S. Climate Prediction Center and the DEMETER project for data and forecasts.

8. REFERENCES

Lloyd-Hughes, B., M. A. Saunders and P. Rockett, 2004: A consolidated CLIPER model for improved August-September ENSO prediction skill, *Wea Forecasting, submitted.*

Palmer, T. N. et al., 2003: Development of a European Multi-Model Ensemble System for Seasonal to Inter-Annual Prediction (DEMETER), *Bull. Amer. Meteor. Soc., submitted.*

Torrence, C., and P. J. Webster, 1998: The annual cycle of persistence in the El Niño-Southern Oscillation. *Quart. J. Roy. Meteor. Soc.,* **124,** 1985-2004.

WMO, 2002: Standardised verification system for long-range forecasts (LRF): new attachment II-9 to the manual on the GDPS, *WMO No. 485, Volume 1*, WMO, Geneva.