Evaluation of a climate model consensus forecast for Victorian farmers, seven years on

Dale Grey[1], and Graeme Anderson[2]

[1] Department of Economic Development, Jobs, Transport and Resources, PO Box 3100, Bendigo DC, VIC, 3551, dale.grey@ecodev.vic.gov.au
[2] Department of Economic Development, Jobs, Transport and Resources, PO Box 103, Geelong, VIC, 3220, graeme.anderson@ecodev.vic.gov.au

Abstract
For seven years DEDJTR have been producing a newsletter called the Fast Break that summarises climate forecasts for Victoria, effectively bringing two hours of web searching onto an A4 page. The newsletter has a distribution list of 2500 people, 62% of whom are farmers with 37% agribusiness and government advisors. Forecasts from seven Coupled Global Circulation Models (CGCM), two ensembles of CGCM's, and two statistical models were collated in July of each year for predictions of the spring period August-October since June 2008. The eleven model predictions were analysed for agreement and the most common prediction was compared against the actual values. This consensus methods ability to predict NINO 3.4 was excellent, but predictions for Indian Ocean, rainfall and temperature were mixed, but had useful skill in the direction in which to "jump", if not "how far". Consensus of models was almost certainly better than picking one model with most skill and following it. Individual model accuracy was erratic and in some years could lead to perverse outcomes if only one was followed. Best overall model performance for rainfall prediction was for the 2010 La Nina event.

Key words
Newsletter, coupled, global, circulation, ensemble, survey

Introduction
Increases in modern computing power have meant that statistical seasonal forecasts have been augmented or replaced with Coupled Global Circulation Models (CGCM). Such models have the power to predict climate phenomena based on the laws of physics, without using historical patterns like the statistical systems do. Rather than facing the confusion of which model to choose in which year and at what time, it is possible that there is greater strength in looking at an ensemble of models. Such an ensemble could be better at removing the noise from individual models and seek a consensus prediction. Ensemble predictions have been used for Indian monsoon rainfall out to five days (Kumar *et.al.* 2012), climate predictions for the Asian monsoon (Krishnamurti *et.al. 2006) and the International Research Institute seasonal climate forecasts for the world (Barnston et. al.* 2010). These authors and others have found greater statistical skill in the Multi Model Ensemble (MME) approach, compared to individual models. A cheaper approach (called a poor man's ensemble PME) is used by some meteorological and climatological agencies, where the outputs from other agency models are combined rather than running many expensive models yourself. The Australian BoM uses such an approach in its Water and the Land eight day rainfall prediction using the approach of Ebert (2001). In 2007 the Australian Bureau of Meteorology created a summary of CGCM predictions of temperature in the NINO 3.4 section of the Pacific Ocean. We thought that by looking at a simple MME the eastern Indian Ocean surface temperature, rainfall and land temperature predictions could also be assessed and be of use to Victoria's Agriculture sector. Following seven years of use, we assessed how this consensus forecast performed for spring. Commonly verification studies involve correlation gridded data sets between forecast and observed values. This study has been low-tech, and forecast and observed values have been visually compared.

Methods
In the last week of every month the outputs of the selected models were obtained off the web. Models were chosen according to their ease of web based access and /or English translation. Predictions were obtained from seven CGCM's (System 4 ECMWF 2015 , POAMA2 BoM1 2015, SINTEX-F JAMSTEC 2015, DFSv2 NCEP 2015, Glosea5 UKMO 2015, GEOS-5 NASA 2015 and CGMCM1.0 BCC 2015), two ensembles of CGCM's (IRI 2015, APCC 2015) and two statistical models (Qld SOI phase system DSITIA 2015, WA ESS AEGIC 2015). The outputs for the sea surface temperature (SST) in the NINO3.4 region and Eastern Indian Ocean, rainfall and land temperature were collated into a table and published in the Fast

Break e-newsletter. Thresholds used for El Niño and La Niña and Indian Ocean Dipole positive and negative were > +/- 1.0 oC. If temperatures in both oceanic regions were > +/ - 1.0 oC they were classified as "warm"/"cool", when between +/- 0.5-1.0 oC they were classified as "slightly cooler"/"slightly warmer". When between +/- 0.5 oC they were classified as "neutral". For rainfall, forecast values between +/- 0.1 oC were classified as "average", between +/- 0.1-0.6 mm a day anomaly were classified as "slightly wetter"/"slightly drier", > +/- 0.6 mm/day were classified as "wetter"/"drier". For temperatures values between +/- 0.3 oC were classified as "average", anomalies between +/- 0.3-1.0 oC were classified "slightly warmer"/"slightly cooler", values > +/- 1.0 oC were classified as "warmer"/"cooler". The majority consensus of all eleven model predictions was distilled into one prediction. No weighting was used. Where models were split between two outcomes, we gave both predictions split by "/". Where there was no model consensus we used the prediction of "mixed".

For this study, we looked at the July predictions for spring (August-October) and compared these to the archived actual values for SST from the OSPO NOAA site (NOAA 2015), and the Bureau of Meteorology historic rainfall and temperature maps (BoM2 2015). A qualitative verdict was given for the correctness of the forecasts, taking into consideration the extent of predicted versus actual across the ocean or state of Victoria. When the actual results for the state were varied predictions were split with a "/". When rating forecasts for SST anomalies, a rating of "excellent" was chosen where the model consensus was the same as the actual outcome, "Good" was chosen if the direction of the temperature signal was correct, "Poor" was used if the actual outcome was not in the direction of that predicted. For rainfall and temperature a rating of "Excellent" was given if greater than 66% of the state's area anomalies were similar to that predicted. "Good" was chosen if 33%-66% of the state had an outcome predicted. "Poor" was chosen if less than 33% of the state was represented by the prediction. Not all models present outputs for all parameters we tested, Table 1 shows the various models and the parameters viewable.

Table 1. Models and the parameters available for use in this study.

	Model	Organisation	Country	NINO 3.4	Eastern Indian Ocean	Rainfall	Temp.
CGCM's	System 4	ECMWF	UK	yes	yes	yes	yes
	POAMA2	BoM	Australia	yes	yes	yes	yes
	SINTEX	JAMSTEC	Japan	yes	yes	yes	yes
	CFSv2	NCEP	USA	yes	yes	yes	yes
	GEOS-5	NASA	USA	yes	yes	yes	yes
	CGMCM1.0	BCC	China			yes	yes
	UKMO	GloSea5	UK	yes	yes	yes	yes
Ensembles	IRI	USA		yes	yes	yes	yes
	APCC	Korea		yes	yes	yes	yes
Statistical	SOI phase	DSITIA Qld	Australia			yes	
	ESS	AEGIC WA	Australia	yes		yes	

Results

Data are presented for the seven years of August to October predictions and their actual outcomes. The consensus of 11 models was excellent at predicting NINO3.4 sea surface temperatures, with the exception of 2012 where a weak El Niño was predicted to form but failed to materialise in spring (Table 2).

Table 2. Predicted and actual outcomes for August-October NINO3.4 temperatures, and a qualitative assessment of correctness.

	Predicted	Actual	Assessment
2008	neutral	neutral	excellent
2009	El Niño	El Niño	excellent
2010	La Niña	La Niña	excellent
2011	slightly cool	slightly cool	excellent
2012	Weak El Niño	neutral	poor
2013	neutral	neutral	excellent
2014	slightly warm	slightly warm	excellent

Consensus of models had fair performance at predicting eastern Indian Ocean sea surface temperatures correctly predicting three out of seven years (Table 3). Furthermore, predictions were partly correct during 2010 and 2011. In 2012 an IOD+ (Indian Ocean Dipole) occurred with little notice that was only predicted by the ECMWF model. In 2013 all models, bar the IRI ensemble predicted an IOD- which failed to occur.

Table 3. Predicted and actual outcomes for August-October eastern Indian Ocean temperatures, and a qualitative assessment of correctness.

	Predicted	Actual	Assessment
2008	IOD+	IOD+	excellent
2009	neutral	neutral	excellent
2010	slightly warm	weak IOD-	good
2011	neutral/weak IOD+	weak IOD+	good
2012	Mixed	IOD+	poor
2013	IOD-	neutral	poor
2014	slightly warm	slightly warm	excellent

Consensus of models generally had good success at predicting spring rainfall, although only half of the models generally made the correct prediction (Table 4). The exception was the wet La Niña year of 2010 where at least the trend of the rainfall response was in the right direction.

Table 4. Predicted and actual outcomes for August-October rainfall, and a qualitative assessment of correctness.

	Predicted	Actual	Assessment
2008	average/slightly drier	drier	good
2009	slightly drier/average	average/slightly drier	good
2010	slightly wetter	wetter	excellent
2011	average	slightly drier/average	good
2012	average/slightly drier	drier	good
2013	average	mixed	good
2014	Average/slightly drier	drier	good

Consensus of models had mixed performance at predicting spring temperatures (Table 5). Actual temperatures were dominated by slightly warmer springs, with the exception of the cooler 2010 La Niña which tripped many models up, with the exception of POAMA. In 2013 only IRI and NASA successfully predicted the warmer spring.

Table 5. Predicted and actual outcomes for August-October temperatures, and a qualitative assessment of correctness.

	Predicted	Actual	Assessment
2008	slightly warmer/average	slightly warmer/average	excellent
2009	slightly warmer	average/slightly warmer	good
2010	Slightly warmer	average/slightly cooler	poor
2011	average/slightly warmer	Slightly warmer	good
2012	average/slightly warmer	Average/slightly warmer	excellent
2013	mixed	slightly warmer/warmer	poor
2014	slightly warmer/average	warmer	good

Discussion

Due to the year to year erratic nature of individual model performance (data not presented), the use of multi-model ensemble forecasts provides some greater clarity, but is still less than the perfection that farmers desire, but may never obtain. Greatest skill was in the prediction of NINO3.4 temperatures. Predictions for the eastern Indian Ocean, Victorian rainfall and temperature were not as good but exhibited useful skill at getting the signal direction right on many occasions. The rainfall predictions are the most agricultural important as nitrogen fertiliser decisions can be made through August and September in many medium to high rainfall districts. Interestingly, there were no poor predictions made by the consensus system where the actual outcome was opposite to that predicted. Such perverse outcomes would cause decision makers to lose the most money from an unpredicted event. Using the results of just one model would occasionally lead to perverse outcomes where the prediction was radically different to the consensus. Rainfall predictions often

contained a signal to the actual outcome but were rarely perfect. At some times a mixed signal could be considered useless, but often implies that individual model variability is so great, that decisions should be made on more known parameters, such as rainfall to date and stored soil moisture. The August to October period is the time where models should have the greatest skill, so it could be implied that predictions made outside this period might have poorer skill than was seen here.

Conclusion

CGCM consensus forecasting had good skill for the prediction of NINO3.4 temperature in spring but mixed skill in other parameters. Improved climate literacy of the how and when of using seasonal forecasting is critical if farmers and advisors are to apply these tools in their on farm decisions. While climate modelling and computer power is improving forecast skill, it will be important to deliver development and extension programs such as The Break to ensure knowledge, trust and utilisation of forecasts by farmers improves.

References

Barnston AG, Li S, Mason S.J, DeWitt D.G, Goddard L, and Gong X (2010). Verification of the First 11 Years of IRI Seasonal Climate Forecasts. J. Appl. Meteor. Climatol. 49, 493-520.

Ebert EE (2001). Ability of a poor man's ensemble to predict the probability and distribution of precipitation. Mon. Wea. Rev. 129, 2461-2480

Krishnamurti TN, Mitra AK, Kumar TSVV, Yun WT and Dewar WK (2006). Seasonal climate forecasts of the South Asian monsoon using multiple coupled models. Tellus A. 58(4), 487-507

Kumar A, Mitra AK, Bohra AK, Iyengar GR and Durai VR (2012). Multi-model ensemble (MME) prediction of rainfall using neural networks during monsoon season in India. Meteorol. Appl. 19, 161-169