AGU PUBLICATIONS

Geophysical Research Letters

RESEARCH LETTER

10.1002/2015GL066173

Key Points:
- An alternate approach to quantifying ensemble ENSO forecast spread is presented
- The 2014 El Nino forecast falls within the expected spread from noise-driven processes

Correspondence to:
S. M. Larson,
slarson@rsmas.miami.edu

Citation:
Larson, S. M., and B. P. Kirtman (2015), An alternate approach to ensemble ENSO forecast spread: Application to the 2014 forecast, *Geophys. Res. Lett.*, *42*, 9411–9415, doi:10.1002/2015GL066173.

Received 11 SEP 2015
Accepted 15 OCT 2015
Accepted article online 21 OCT 2015
Published online 6 NOV 2015

An alternate approach to ensemble ENSO forecast spread: Application to the 2014 forecast

Sarah M. Larson[1] and Ben P. Kirtman[1]

[1]Rosenstiel School of Marine and Atmospheric Science, University of Miami, Coral Gables, Florida, USA

Abstract Evaluating the 2014 El Niño forecast as a "bust" may be tapping into a bigger issue, namely that forecast "overconfidence" from single-model ensembles could affect the retrospective assessment of El Niño–Southern Oscillation (ENSO) predictions. The present study proposes a new approach to quantifying an "expected" spread and uncertainty from noise-driven processes and supplementing these measures with actual ENSO forecasts. Expanding on a previously developed coupled model framework that isolates noise-driven ENSO-like errors, an experimental design is implemented to generate an expected December Niño-3.4 spread from March initial condition sea surface temperature errors that have similar structure to the 2014 and 2015 observed. Results reveal that the 2014 ENSO forecast falls within the expected uncertainty generated by ENSO-independent, forecast-independent, noise-driven errors.

1. Introduction

Many dynamical forecast models predicted a 2014 El Niño event. Yet despite the moderate sea surface temperature (SST) warming that was observed, the 2014 forecast is often described as a "busted" forecast. The question that arises is of practical importance—was 2014 actually a bust or is the usual method of calculating ensemble spread (defined below) underestimating forecast uncertainty and by extension, affecting the retrospective evaluation of El Niño–Southern Oscillation (ENSO) predictions?

A continuing challenge in ENSO prediction is determining how the range of possibilities for a forecast can be better represented without requiring increasingly more ensemble members [*Kumar and Hoerling*, 2000; *Kumar et al.*, 2001]. In particular, single-model ensembles often produce too small of spread or "overconfidence" [e.g., *Buizza and Palmer*, 1998; *Richardson*, 2001]. One method of tackling this issue is the implementation of multimodel ensemble prediction systems, including the North American Multimodel Ensemble (NMME) [*Kirtman et al.*, 2014] and other modeling groups [e.g., *Doblas-Reyes et al.*, 2000; *Palmer et al.*, 2004]. Ensemble predictions combined from multiple models increase forecast diversity and provide a better estimate of uncertainty owing to model formulation [*Kharin and Zwiers*, 2002; *Palmer et al.*, 2004; *Kirtman et al.*, 2014]. Nevertheless, there persists a need to develop alternative approaches to represent uncertainty for single-model ensembles as well, because individual models provide the basis for the multimodel ensemble forecast spread and the number of ensemble members provided is often financially limited.

The present study proposes a new approach to assigning ensemble spread to single-model ensembles in an attempt to better represent an "expected" uncertainty generated by noise-driven errors. Error growth behavior is often model dependent and requires a model-by-model approach to diagnosing not only systematic biases that are studied extensively in the literature [e.g., *Guilyardi et al.*, 2009; *Bellenger et al.*, 2014] but also dynamically driven errors that are difficult to isolate in complex coupled models. The present effort is based on the quantification of dynamically driven error growth for seasonal ensembles using a recently developed National Center for Atmospheric Research Community Climate System Model, version 4 (CCSM4), approach [*Larson and Kirtman*, 2015; hereafter LK15]. The experimental design isolates noise-driven coupled instability error growth that is independent of the ENSO cycle. Expanding on the LK15 methodology, we examine the question—how much forecast spread is expected from noise-driven error growth for the 2014 and 2015 CCSM4 NMME forecasts?

The fundamental assumption is that we are considering error growth independent of the ENSO cycle, which includes both the classical equatorial SST structure and the subsurface heat content precursor [*Cane et al.*, 1986; *Meinen and McPhaden*, 2000; *McPhaden*, 2003]. The introduction of errors to the system is due to intrinsic ENSO-independent perturbations, either present in the initial conditions or occurring stochastically later in the evolution. Note that all measures of spread are defined as the average deviation about the ensemble mean.

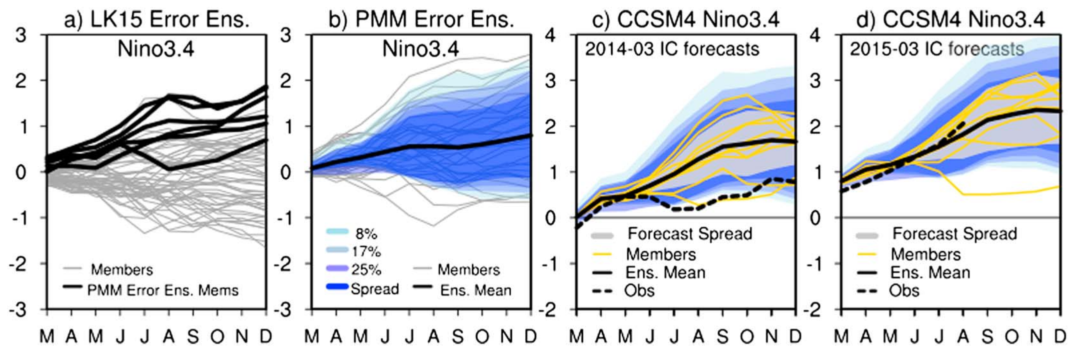

Figure 1. (a) Niño-3.4 SST error swath for the March error growth ensemble in *Larson and Kirtman* [2015]. Black bold curves indicate the six members whose ocean initial conditions are used in the PMM error ensemble. (b) Niño-3.4 error swath, ensemble mean, spread, and uncertainty thresholds for the PMM error ensemble. Uncertainty thresholds are calculated by averaging the most extreme three, six, and nine warm or cold members corresponding to 8%, 17%, and 25%, respectively. (c) 2014 CCSM4 March initialized Niño-3.4 forecasts from the NMME, forecast spread (grey polygon), ensemble mean (black solid), observations (black dashed), and expected spread and uncertainty thresholds from Figure 1b. (d) Same as Figure 1c but for the 2015 forecasts.

2. Observations and Forecasts

Motivated by the LK15 framework, an expected spread is added to the 2014 and 2015 March initialized CCSM4 ENSO forecasts from the NMME. March initializations are chosen for their temporal proximity to the spring predictability barrier [*Webster and Yang*, 1992; *Kirtman et al.*, 2002; *Mu et al.*, 2007] and because the spring background state is particularly well suited for coupled feedbacks to amplify errors and produce large ensemble spread in CCSM4 [LK15]. Additionally, both 2014 and 2015 March initial conditions have robust Pacific Meridional Mode (PMM) [*Chiang and Vimont*, 2004] SST anomaly signatures, an ENSO-independent intrinsic mode of variability also present in the LK15 ensembles and a potential source of ENSO forecast error due to PMM's robust precursor relationship with ENSO. The relationship is highlighted in model, observational, and dynamical forecast studies [*Chiang and Vimont*, 2004; *Chang et al.*, 2007; *Zhang et al.*, 2009; *Larson and Kirtman*, 2013, 2014]. Last, both forecasts predict December El Niño larger than 1.5°C from PMM initial conditions.

Figures 1c and 1d show the CCSM4 March initialized Niño-3.4 anomaly forecasts with 10 ensemble members (gold), ensemble mean (black solid), forecast spread (grey polygon), and the observed Niño-3.4 index from the National Center for Environmental Prediction Climate Prediction Center (CPC; black dashed). To be clear, "forecast spread" is computed from the actual CCSM4 initialized forecasts, whereas "expected spread" is the proposed supplemental spread discussed below. The 2015 observed Niño-3.4 evolution falls within the forecast spread, and thus far, is on track to verify, albeit the final amplitude may vary from that predicted by this particular model (2.33°C in December). On the other hand, observed 2014 Niño-3.4 sits well below the forecast spread and verifies at 0.78°C as compared to the ensemble mean forecast of 1.67°C. So was 2014 a "bust" or does 0.78°C fall within the expected spread generated via noise-driven processes in CCSM4, despite the December forecast spread spanning only 1.16–2.17°C?

The 2014 and 2015 forecasts both predict El Niño from PMM-like initial conditions (Figure 2). Depicted are the 0.5-month lead CCSM4 ensemble mean March SST anomaly forecasts and are considered close proxies for the initial condition. In earlier versions of the NMME models (i.e., phase 1 models), strong positive projections of PMM in the 0.5-month lead March initialized forecasts correctly predict a Niño-3.4 forecast that falls within the upper tercile 49% of the time, which is moderately larger than the 33% expected from randomness [*Larson and Kirtman*, 2014]; therefore, the 2014 and 2015 El Niño forecasts are not particularly surprising. Since PMM predicts El Niño with some skill in the dynamical forecasts, then PMM-like initial errors could have an impact on ENSO forecast error.

3. Model Experiment and Results

As previously mentioned, we expand upon the LK15 model framework, specifically utilizing their March initialized error growth ensemble. The ensemble consists of 60 members each branched from different March

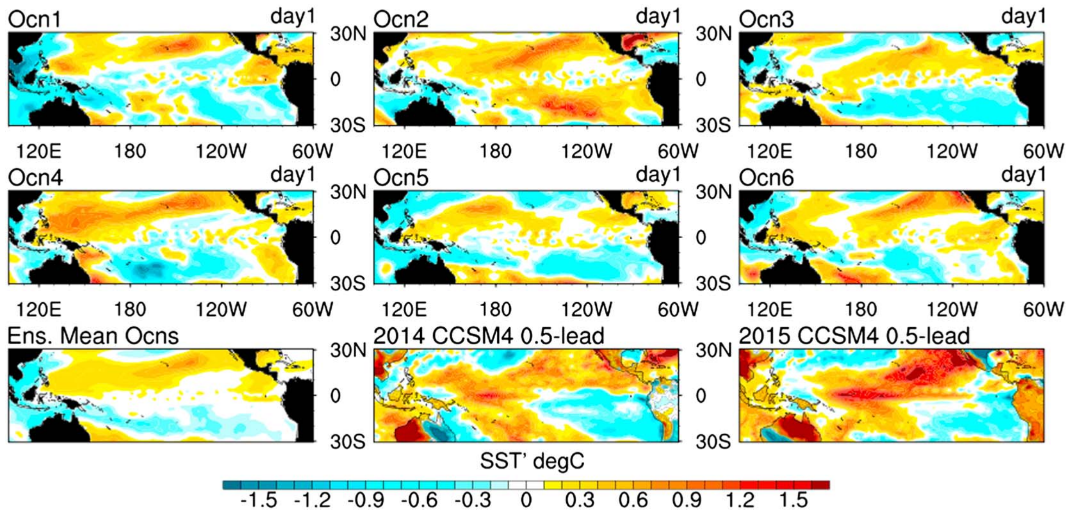

Figure 2. (top and middle rows) Day 1 SST errors in the six ocean initial conditions used for the PMM error ensemble, the (bottom left) ensemble mean, and the (bottom middle) 2014 and (bottom right) 2015 0.5-lead CCSM4 forecasts of March SST anomalies from the NMME.

initial conditions originating from the same base simulation named the Climatological Wind Experiment (CWE; see LK15 for more details). CWE is a mechanically decoupled CCSM4 simulation. Mechanical decoupling is achieved by forcing the ocean component with CCSM4 daily climatological wind stresses while allowing the atmosphere to respond freely without constraint.

Since tropical interannual SST variability is largely wind-driven whereas extra-tropical variability is influenced more by thermal fluxes [*Neelin et al.*, 1994], mechanically decoupling the atmosphere and ocean considerably reduces SST variability in the tropics only. The extratropics are less impacted. Without dynamic wind stress support, the ENSO cycle, large subsurface heat content precursors, and coupled instabilities along the equatorial Pacific are all unsupported in the CWE; therefore, there is no ENSO cycle or subsurface precursors in the initial conditions of the 60 ensemble members. The only perturbations are intrinsic to the system and are dynamically decoupled as per the experimental design. There is practically no interannual SST variability in the ENSO region in the CWE. As a result, overlying atmospheric wind variability is damped by 80% throughout the tropical Pacific compared to a CCSM4 control simulation. Essentially, equatorial Pacific wind perturbations are small and ENSO-independent in CWE. The 60 cases are initialized in March by essentially "turning the coupling back on" (i.e., no longer overriding the atmospheric wind stresses with climatology) to allow for the activation of coupled instabilities via interaction between the perturbations at the air-sea interface, which are treated as errors in the initial condition. Figure 1a shows that several members generate largely warm or cold biased Niño-3.4 SST error growth whereas others remain fairly neutral.

CWE is a free-running simulation, so each member is branched from sequential March initial conditions occurring 1 year apart. In LK15, 10 cases grow noticeably warmer than the other 50, exceeding 1 standard deviation of Nino-3 in December. To discount possible decadal variability that previously went undetected and allow for the maximum amount of model spinup, only the subset of the 10 cases branched from the last 30 year period of the CWE are selected. This results in the selection of the six warm biasing cases as shown in Figure 1a (black bold). We use only the ocean initial conditions from these six cases, referred to as Ocn1-6. The day 1 SST errors from Ocn1-6 and the ensemble mean are shown in Figure 2. All members show some extent of a PMM-like projection; thus, each member is considered having PMM-like errors in the initial condition. This type of structure is desirable to test how errors may grow if error exists in PMM amplitude or structure. The six-member subset is well representative of the 10 warm biasing cases because 9/10 of the day 1 SST errors have PMM-like errors in the initial condition. The fact that the PMM is active in the CWE provides strong confirmation that PMM is ENSO-independent as originally argued in *Chiang and Vimont* [2004].

To test the sensitivity of the ocean perturbations to atmospheric noise, (i.e., errors in the atmospheric initial conditions) the six initial ocean conditions are each paired with six different atmosphere initial conditions from LK15 members that bias cold. These atmospheres are chosen to ensure that the atmospheric noise is entirely decoupled in terms of both buoyancy and momentum fluxes because initial SST error structures can be similar (Figure 2). A total of 36 combinations comprise the PMM error ensemble.

How much scatter can be expected from PMM-like errors in the initial condition SST? Figure 1b shows the Niño-3.4 error swath for the PMM error ensemble. The range of possibilities of December Niño-3.4 error from PMM-like initial errors is large, spanning −0.65–2.57°C with little clustering around a particular final error value. Despite the large spread, the errors are clearly biased warm. This suggests that PMM-like perturbations in the initial condition can shift the distribution warm, implying an increased probability of El Niño. Nevertheless, the noise-driven scatter demonstrates how sensitive March is to noise-driven perturbation growth that affects December ENSO verification (i.e., large expected spread) thus suggesting large forecast uncertainty for these longer lead times.

The PMM error ensemble produces a noticeably larger spread (−0.11–1.72°C) than the forecast spread computed from the CCSM4 forecasts (grey polygons in Figures 1c and 1d). Other thresholds are computed, including averaging the most extreme three, six, and nine ensemble members that bias warm or cold, respectively the 8%, 17%, and 25% levels (Figure 1b). Note that this method may still be underestimating the expected spread as we selected only members that originally bias warm in LK15. A few members have PMM-like initial errors but bias cold instead. The chosen members begin a warm biased trajectory that continues through December, apart from seasonal waxing and waning of the growth rate. Nevertheless, the present method of choosing only six ocean base cases proves to produce a large range of possibilities including members that bias slightly cold. Most importantly, the PMM error ensemble clearly generates a larger expected spread than the actual forecast spread for either 2014 or 2015.

Since the noise-driven scatter from PMM-like errors is large, does this measure of ENSO-independent expected spread change our evaluation of the 2014 forecast? The expected spread and uncertainty thresholds in Figure 1b are added to the 2014 and 2015 ensemble mean forecasts (Figures 1c and 1d). For 2014, observed Niño-3.4 still sits below all uncertainty thresholds during summer, but the final amplitude (0.78°C) falls within the 25% expected threshold (0.51–2.87°C) and is within the lower bounds of the expected spread (0.75–2.58°C), meaning that even though the observed amplitude sits below the forecast spread (1.16–2.17°C), it is an unsurprising outcome when considering the expected noise-driven scatter. This shows that ENSO-independent, noise-driven processes alone can produce large spread and perhaps, the expected spread is a better representation of the uncertainty.

We emphasize that noise-driven scatter is present in the forecast spread, but because of the small ensemble size, the forecast spread does not adequately represent the forecast uncertainty at longer lead times. We also stress that we are not suggesting the removal of the forecast spread polygon, but the addition of an expected spread or uncertainty threshold from noise-driven processes that can serve as a benchmark for expected uncertainty from the stochastic component of coupled system. The benefit being that this approach does not require an increase of forecast ensemble size, but a diagnostic assessment of the noise-driven error behavior away from the forecast setting. Expected spread and uncertainty thresholds can be applied to each forecast, thus providing a measure of spread that is independent of the forecast itself. This way, even if single-model ensemble overconfidence is evident in the forecast spread, a reality check of the potential spread from noise-driven errors is available to supplement the "expert assessment" of the model forecast.

4. Discussion

Ensemble spread is an important estimate of forecast uncertainty and assigning confidence in ENSO predictions requires an expert assessment that may benefit from measures of uncertainty besides forecast spread alone, which can be affected by ensemble size. Quantifying the expected noise-driven spread in fully coupled models that are used in real-time seasonal predictions is essential to providing a forecast-independent measure of the possible uncertainty that can occur. LK15 present a coupled model framework to isolate such error growth for single-model ensembles.

The present study expands on the LK15 experimental design to produce an expected noise-driven spread for the March initialized 2014 and 2015 CCSM4 NMME predictions of December Niño-3.4. The initial conditions of both forecasts contain PMM SST signatures and predict December El Niño. A PMM error growth ensemble is constructed from six ensemble members that originally bias warm in the LK15 March error ensemble, each also having PMM-like errors in the initial conditions. The constructed ensemble is essentially a sensitivity test, pairing each base ocean with six different atmospheric noise initial conditions.

In reality, the 2015 prediction, so far, appears realistic, whereas the 2014 prediction is often discussed as a bust. Based on the expected spread produced by the PMM error ensemble, we argue that the observed 2014 December Niño-3.4 warming falls well within the expected uncertainty for noise-driven error growth originating from the interaction of atmospheric noise with PMM-like errors in the initial condition SST.

We are not suggesting that a new error ensemble must be constructed for each actual forecast but merely demonstrating that the expected spread from initial SST errors with similar structures as the March 2014 and 2015 initial conditions used in the real-time CCSM4 NMME forecasts is large for longer lead times. In fact, the spread in the PMM error ensemble (Figure 1b) is similar to that in the original LK15 March error ensemble (Figure 1a), suggesting that the expected spread is similar for differently constructed noise-driven error ensembles, as long as the initial conditions are ENSO-independent. Thus, the spread of the LK15 ensembles suffice as a benchmark for expected spread from noise-driven processes alone in CCSM4. In this case, more information is gained by using ocean base members with PMM-like initial errors, including that the presence of PMM-like perturbations in March can increase the probability that El Niño will occur (i.e., the warm bias), but does not guarantee it due to the large sensitivity of the coupled system to perturbations in March (i.e., large expected spread).

The present study is only looking at error growth that is independent of the ENSO cycle. The expected spread may change if an ENSO cycle is present in the initial condition and is a topic worth investigating as it has potentially large practical benefits for the assessment of forecast uncertainty. We are actively developing and implementing an experiment design to test this distinction.

Acknowledgments
We thank Bruce Anderson and an anonymous reviewer for helpful comments and suggestions. The NMME project is supported by NOAA, NSF, NASA, and DOE. This research was supported by NSF grants AGS1137911 and AGS1450811. The CPC Niño-3.4 index is available at http://www.cpc.ncep.noaa.gov/data/indices/sstoi.indices. The authors acknowledge computational support from the University of Miami Center for Computational Science. The data for this paper are available upon request.

References

Bellenger, H., E. Guilyardi, J. Leloup, M. Lengaigne, and J. Vialard (2014), ENSO representation in climate models: From CMIP3 to CMIP5, *Clim. Dyn.*, *42*(7–8), 1999–2018.

Buizza, R., and T. N. Palmer (1998), Impact of ensemble size on ensemble prediction, *Mon. Weather Rev.*, *126*, 2503–2518.

Cane, M. A., S. E. Zebiak, and S. C. Dolan (1986), Experimental forecasts of El Niño, *Nature*, *321*, 827–832.

Chang, P., L. Zhang, R. Saravanan, D. J. Vimont, J. C. H. Chiang, L. Ji, H. Seidel, and M. K. Tippett (2007), Pacific meridional mode and El Niño—Southern oscillation, *Geophys. Res. Lett.*, *34*, L16608, doi:10.1029/2007GL030302.

Chiang, J. C. H., and D. J. Vimont (2004), Analogous Pacific and Atlantic meridional modes of tropical atmosphere–ocean variability, *J. Clim.*, *17*, 4143–4158.

Doblas-Reyes, F. J., M. Déqué, and J.-P. Piedelievre (2000), Multi-model spread and probabilistic seasonal forecasts in PROVOST, *Q. J. R. Meteorol. Soc.*, *126*, 2069–2087.

Guilyardi, E., A. Wittenberg, A. Fedorov, M. Collins, C. Wang, A. Capotondi, G. J. van Oldenborgh, and T. Stockdale (2009), Understanding El Niño in ocean–atmosphere general circulation models: Progress and challenges, *Bull. Am. Meteorol. Soc.*, *90*, 325–340.

Kharin, V. V., and F. W. Zwiers (2002), Climate predictions with multimodel ensembles, *J. Clim.*, *15*(7), 793–799.

Kirtman, B. P., J. Shukla, M. Balmaseda, N. Graham, C. Penland, Y. Xue, and S. Zebiak (2002), Current status of ENSO forecast skill, *A report to the Climate Variability and Predictability (CLIVAR) Numerical Experimentation Group (NEG), CLIVAR Working Group on Seasonal to Interannual Prediction*.

Kirtman, B. P., et al. (2014), The North American Multimodel Ensemble (NMME): Phase-1 seasonal to interannual prediction; Phase-2 toward developing intra-seasonal prediction, *Bull. Am. Meteorol. Soc.*, *95*, 585–601.

Kumar, A., and M. P. Hoerling (2000), Analysis of a conceptual model of seasonal climate variability and implications for seasonal prediction, *Bull. Am. Meteorol. Soc.*, *81*, 255–264.

Kumar, A., A. G. Barnston, and M. P. Hoerling (2001), Seasonal predictions, probabilistic verifications, and ensemble size, *J. Clim.*, *14*, 1671–1676.

Larson, S., and B. Kirtman (2013), The Pacific meridional mode as a trigger for ENSO in a high-resolution coupled model, *Geophys. Res. Lett.*, *40*, 3189–3194, doi:10.1002/grl.50571.

Larson, S. M., and B. P. Kirtman (2014), The Pacific meridional mode as an ENSO precursor and predictor in the North American multi-model ensemble, *J. Clim.*, *27*, 7018–7032.

Larson, S. M., and B. P. Kirtman (2015), Revisiting ENSO coupled instability theory and SST error growth in a fully coupled model, *J. Clim.*, *28*, 4724–4742.

McPhaden, M. J. (2003), Tropical Pacific Ocean heat content variations and ENSO persistence barriers, *Geophys. Res. Lett.*, *30*(9), 1480, doi:10.1029/2003GL016872.

Meinen, C. S., and M. J. McPhaden (2000), Observations of warm water volume changes in the equatorial Pacific and their relationship to El Niño and La Niña, *J. Clim.*, *13*, 3551–3559.

Mu, M., H. Xu, and W. Duan (2007), A kind of initial errors related to "spring predictability barrier" for El Niño events in Zebiak-Cane model, *Geophys. Res. Lett.*, *34*, L03709, doi:10.1029/2006GL027412.

Neelin, J. D., M. Latif, and F. F. Jin (1994), Dynamics of coupled ocean–atmosphere models: The tropical problem, *Annu. Rev. Fluid Mech.*, *26*, 617–659.

Palmer, T. N., et al. (2004), Development of a European Multimodel Ensemble System for Seasonal-to-Interannual Prediction (DEMETER), *Bull. Am. Meteorol. Soc.*, *85*, 853–872.

Richardson, D. S. (2001), Measures of skill and value of ensemble prediction systems, their interrelationship and the effect of ensemble size, *Q. J. R. Meteorol. Soc.*, *127*, 2473–2489, doi:10.1002/qj.49712757715.

Webster, P. J., and S. Yang (1992), Monsoon and ENSO: Selectively interactive systems, *Q. J. R. Meteorol. Soc.*, *118*, 877–926.

Zhang, L., P. Chang, and L. Ji (2009), Linking the Pacific meridional mode to ENSO: Coupled model analysis, *J. Clim.*, *22*, 3488–3505.