Geophysical Research Letters

RESEARCH LETTER
10.1029/2019GL085914

Key Points:
- To enhance ENSO prediction, we combine two forecast methods by nudging a linear inverse model (LIM) with a model analog (MA)
- Overall, the MA-LIM forecasts tropical sea surface temperature anomalies better than the LIM and MA at all lead months
- This success is due to the use of more accurate initial condition than MA and an implementation of seasonal cycle and nonlinearity into LIM

Supporting Information:
- Supporting Information S1

Correspondence to:
S. Park,
sungsup@snu.ac.kr

Citation:
Shin, J., Park, S., Shin, S.-I., Newman, M., & Alexander, M. (2020). Enhancing ENSO prediction skill by combining model analog and linear inverse models (MA-LIM). *Geophysical Research Letters*, *47*, e2019GL085914. https://doi.org/10.1029/2019GL085914

Received 19 OCT 2019
Accepted 26 DEC 2019
Accepted article online 30 DEC 2019

Enhancing ENSO Prediction Skill by Combining Model-Analog and Linear Inverse Models (MA-LIM)

Jihoon Shin[1], Sungsu Park[1], Sang-Ik Shin[2,3], Matthew Newman[2,3], and Michael A. Alexander[3]

[1]School of Earth and Environmental Sciences, Seoul National University, Seoul, South Korea, [2]CIRES, University of Colorado Boulder, Boulder, CO, USA, [3]NOAA/ESRL, Boulder, CO, USA

Abstract To enhance El Niño–Southern Oscillation (ENSO) forecast skill, we devise a model analog (MA)-linear inverse model (LIM) by nudging sea surface temperature and sea surface height anomalies forecasted by the LIM into the MA. The performances of the LIM, MA, and MA-LIM are compared to general circulation model simulations and observations. At short (long) lead month τ, the LIM (MA) predicts the Niño 3.4 SST anomalies better than the MA (LIM). On the other hand, the MA-LIM shows the best performance at all τ. At $\tau = 6$, the MA performs better than the LIM in the eastern equatorial Pacific and Indian Oceans but worse in other regions. The MA-LIM substantially remedies the undesirable aspects of the MA. The success of the MA-LIM appears to come from the use of more accurate initial conditions than the MA and an ad hoc implementation of seasonal cycle and nonlinearities into the LIM through nudging to the MA.

Plain Language Summary The El Niño–Southern Oscillation (ENSO) is the most important tropical atmosphere-ocean phenomena driving changes of weather and climate over the globe. Two empirical forecast methods—the linear inverse model (LIM) and model analog (MA)—are known to have nearly the same ENSO forecast skill as general circulation models. To further enhance the ENSO forecast skill, we develop a model analog-linear inverse model (MA-LIM) by combining these two empirical methods. We found that the MA-LIM performs much better than individual LIM and MA. Thus, the MA-LIM can be effectively used to improve ENSO predictions.

1. Introduction

The El Niño–Southern Oscillation (ENSO) is the dominant tropical atmosphere-ocean coupled mode. The impact of ENSO is not confined within the tropical Pacific but extends into remote tropical areas, subtropics, and midlatitudes via anomalous Walker and Hadley circulations, equatorial Rossby or Kelvin waves, and quasi-stationary Rossby wave throughout the year (Alexander et al., 2002, 2004; Lau & Nath, 1996; Park, 2004c; Park & Leovy, 2004; Rasmusson & Carpenter, 1982; Ropelewski & Halpert, 1987; Rowell, 2001; Trenberth et al., 1998). An accurate forecast of ENSO is necessary for a reliable weather and climate prediction over the globe. Many general circulation models (GCMs), however, have problems in simulating the basic statistical properties of ENSO (e.g., Bellenger et al., 2014) and the ENSO-related tropical SST anomalies simulated by most GCMs extend too far west (e.g., Heureux et al., 2020; Li & Xie, 2014; Zheng et al., 2012).

Many statistical methods have been developed to forecast ENSO. Here, we will focus on two Markovian techniques that predict anomalous sea surface temperatures (SSTs) and sea surface heights (SSHs) throughout the tropical Indo-Pacific, linear inverse models (LIMs; Penland & Sardeshmukh, 1995) and model analogs (MA; Ding et al., 2018, 2019). The LIM is an empirical dynamic model, which assumes that the temporal evolution of the predictand is described by a multivariate linear Markov process plus some noise that represents rapidly evolving (and hence unpredictable) nonlinearities. MA forecast ensembles are extracted from preexisting long GCM simulations, by finding those states that best match each initial observed anomaly and tracking their subsequent evolution. Newman and Sardeshmukh (2017) showed that the hindcast skill of the multimodel GCM ensemble mean (MME, hereafter) is comparable to the LIM in the central western Pacific, but is higher in the eastern Pacific at long lead; generally, however, the LIM had skill that met or exceeded single model ensemble means. Ding et al. (2018) showed that MA not only effectively reproduces the forecast skill of MME throughout the tropics but is also significantly better in the eastern Pacific despite

being based on the same models used for the MME. These studies suggest that potential forecast skill is roughly linear in the central Pacific where the LIM works well, but also has a predictably nonlinear component in the eastern Pacific where the MA works better. Both techniques are anomaly models that, unlike the MME, are identically bias corrected and do not suffer from initialization shock. However, both also have practical limits due to the need for lengthy observational or GCM simulation data sets and the choice of the state vector, which may or may not represent all of the predictive information in the initial climate state. In our study, we investigate combining them in a model analog-linear inverse model (MA-LIM), by nudging the monthly SST and SSH anomalies forecasted by the LIM to those forecasted by the MA.

2. The MA-LIM

In this section, we briefly summarize the LIM and MA and explain how we construct the MA-LIM. As described in Penland and Sardeshmukh (1995), the LIM approximates the temporal evolution of a state vector $\vec{x}$ with the stochastically forced linear dynamical system,

$$\mathrm{d}\vec{x}/\mathrm{d}t = \mathbf{L}\vec{x} + \xi, \tag{1}$$

where $\mathbf{L}$ is the linear system matrix and ξ is a Gaussian white noise forcing vector. From (1), the ensemble mean forecast (and most probable state vector) $\vec{x}(t+\Delta t)$ at forecast lead Δt is $\vec{x}(t+\Delta t) = \exp(\mathbf{L}\Delta t)\vec{x}(t)$. The matrix $\mathbf{L}$ is determined by an error variance minimization procedure as $\mathbf{L} = \ln[\mathbf{C}(\tau_0)\mathbf{C}(0)^{-1}]/\tau_0$, where $\mathbf{C}(\tau_0) = \langle \vec{x}(t+\tau_0)\cdot\vec{x}(t)^T\rangle$ is the covariance matrix at lag τ_0 (= 1 month in this study) and the angle bracket denotes the expected mean. Following Newman and Sardeshmukh (2017), our LIM computes temporal evolution of a state vector in the EOF space, so $\vec{x}$ represents the leading 16 and 9 principal components (PCs) of monthly SST and SSH anomalies, which explain 76% (80) and 61% (64) of the total variance, respectively, for the GCM simulations (observations). The number of EOF modes retained was chosen by trial and error to maximize the cross-validated forecast skill, but the skill is relatively insensitive to this choice.

The MA computes the temporal evolution of SST/SSH anomalies at model grid points by averaging the cases from a long GCM simulation where the SST/SSH fields are similar to the target one (Ding et al., 2018, 2019). More specifically, if the forecast starts from calendar month, $t = 0$, the MA selects the 20 cases with small normalized root-mean-square error in the tropics, $E(i) \equiv 0.5\cdot[rmse(SST'(0), SST'_p(0,i))/\sigma(SST'_p) + rmse(SSH'(0), SSH'_p(0,i))/\sigma(SSH'_p)]$ at the same calendar month in different years. Here, $SST'(0)/SSH'(0)$ are the monthly anomalies of SST/SSH at the current time step, $SST'_p(0,i)/SSH'_p(0,i)$ are the monthly anomalies of SST/SSH at the calendar month 0 in the ith year obtained from a certain population (from now on, the long-term GCM simulation from which the MA cases are drawn will be referred to as a population), *rmse* is the spatial root-mean-square error in the tropics, and $\sigma(SST'_p)$ and $\sigma(SSH'_p)$ are the spatiotemporal standard deviations of monthly SST/SSH anomalies averaged over the tropical oceans in the population. Again, although the ensemble size was chosen by trial and error to optimize the forecast skill, the results are not too sensitive to this choice (see Ding et al., 2018 for further analysis). In this paper, we focus only on combining the single LIM forecast with the ensemble mean MA forecast, although the approach could be employed for individual ensemble members.

The strategy of the MA-LIM is to combine the state vector $\vec{x}$ predicted by the LIM with that predicted by the MA (Figure 1). To combine the LIM in EOF space and MA in model grid space, the monthly SST/SSH anomalies obtained by the MA are first projected onto the 16-SST/9-SSH EOFs, and then combined with the LIM with an appropriate weighting factor, w, defined so that $w = 0$ ($w = 1$) returns the LIM (MA) forecast. We have tested sensitivity to w and results are shown below. It was found that $w \approx 0.15$ produced the best forecast skill for the SST anomalies in the Tropics between 25°S and 25°N. Because of the EOF truncation used, the reconstructed initial monthly anomalies of SST/SSH for the LIM (i.e., $SST'_L(0)$ in Figure 1) are slightly different from the raw monthly anomalies of SST/SSH (i.e., $SST'(0)$ in Figure 1). Moreover, the initial monthly anomalies of SST/SSH obtained from the ensemble mean of the 20 cases for the MA (i.e., $SST'_a(0)$ in Figure 1) are also different from $SST'(0)/SSH'(0)$ due in part to sampling uncertainty. In general, $SST'(0)$ is much closer to $SST'_L(0)$ than $SST'_a(0)$, so that our MA-LIM uses $SST'_L(0)/SSH'_L(0)$ as initial conditions.

We perform two analyses: one is a "perfect" model analysis, where the initial target states are drawn from the control run, and the other is an observation analysis, where the initial target states are obtained

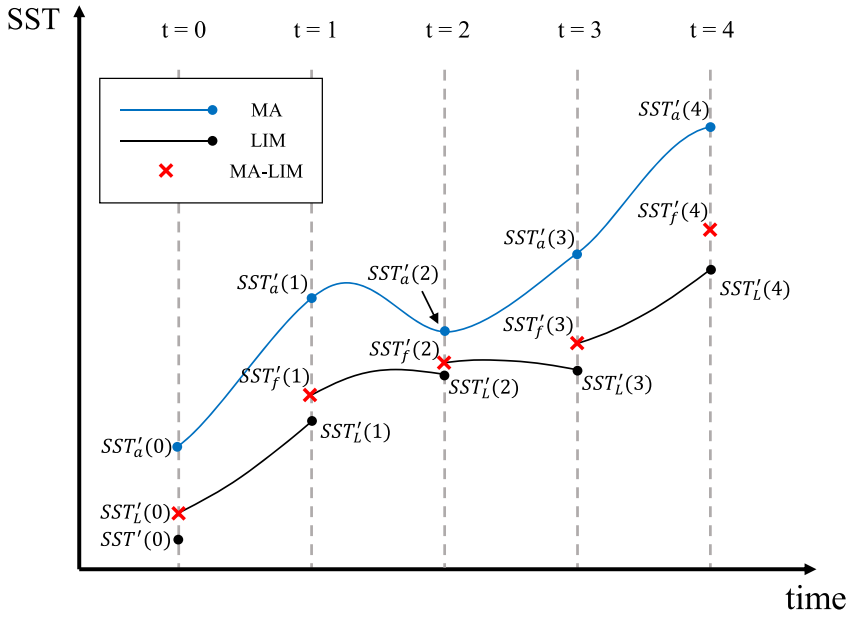

Figure 1. Diagram illustrating the MA-LIM for the tropical oceans. The initial monthly SST anomaly used by LIM, $SST'_L(0)$, is constructed from the 16/9 EOFs and then a 1 month LIM forecast is made to $SST'_L(1)$. Due to the EOF truncation, $SST'_L(0)$ is slightly different from the raw initial monthly SST anomaly, $SST'(0)$. Then, the MA selects 20 cases from the population with anomalous SST/SSH fields similar to $SST'(0)/SSH'(0)$ and computes $SST'_a(0)$, … $SST'_a(t)$, which is the ensemble average of the EOS-reconstructed 20 selected cases at the forecast time, t. Note that $SST'_a(0)$ is different from $SST'(0)$ and $SST'_L(0)$. The initial conditions of the MA-LIM are set to be identical to those of the LIM. The final forecasted monthly SST anomaly at $t = 1$ is obtained by $SST'_f(1) = (1 - w) \cdot SST'_L(1) + w \cdot SST'_a(1)$, where $0 \leq w \leq 1$ is a weighting factor, and these procedures are repeated. In summary, by choosing $w = 0.15$, the MA-LIM can be understood as a LIM slightly nudged by the MA. The same method is used for predicting SSH as well as SST.

from observed conditions during January 1961–December 2010. We performed the analyses using monthly SST/SSH fields at 1° latitude × 1° longitude horizontal resolution in the tropics (25°S to 25°N). For the perfect model analysis, we use 450 years from the preindustrial period simulated by the Seoul National University Atmosphere Model Version 0 with a Unified Convection Scheme (SAM0-UNICON; Park et al., 2019). All model analyses are tenfold cross validation of the 10 subsets of the simulation data: each subset consists of a 45 year evaluation period (10% of the entire data) and the remaining 405 year population. The EOFs of SST and SSH are obtained from the entire 450 year simulation, while the linear system matrix **L** and monthly climatologies of SST/SSH are obtained from the 405 year population. For the hindcast observational analysis, we use 50 years of observed monthly SST (HadISST; Rayner et al., 2003) and SSH (ECMWF ocean reanalysis; Balmaseda et al., 2013). Prior to performing the analysis, we detrended the observed monthly SST/SSH in each grid box. Similar to the perfect model analysis, all observational analyses are from the cross validation of the 10 subsets of the observation data with a 5 year evaluation period in each subset. The EOFs of SST/SSH are obtained from the entire observational record, while the linear system matrix **L** and monthly climatologies of SST/SSH at the grid boxes are obtained from the 45 year data excluding the evaluation period. The LIM approach relies on the assumption that **L** is independent of the lag τ_0. To test the validity of this assumption, we performed the so-called τ test using the method suggested by Winkler et al. (2001) and verified that both the observation and GCM simulation data used in our study passed the τ test, as shown in Figure S1 in the supporting information.

SAM0-UNICON (or simply, SAM0), one of the international coupled GCMs participating in phase 6 of the Coupled Model Intercomparison Project (CMIP6; Eyring et al., 2016), is based on the Community Atmosphere Model version 5 (CAM5; Neale et al., 2010, Park et al., 2014), but CAM5's shallow (Park & Bretherton, 2009) and deep convection schemes (Zhang & McFarlane, 1995) have been replaced by the unified convection scheme described by Park, (2014a, 2014b), with a revised treatment of the convective detrainment processes (Park et al., 2017). Park et al. (2019) showed that the global mean climate and ENSO simulated by SAM0 were roughly similar to those of CAM5/Community Earth System Model version 1 (CESM1;

Hurrell et al., 2013); however, SAM0 substantially improved the simulations of the Madden-Julian Oscillation, diurnal cycle of precipitation, and tropical cyclones.

To quantify forecast skill, we transformed the EOF forecasts back into geographic space and computed the anomaly correlation coefficient, $ACC = Cov(SST'_f, SST')/\{\sigma(SST'_f) \cdot \sigma(SST')\}$ and the root-mean-square-error-based skill score, $RMSSS = 1 - rmse(SST'_f, SST')/\sigma(SST')$ as a function of the lead month. Here, $Cov(SST'_f, SST')$ and $rmse(SST'_f, SST')$ are the covariance and root-mean-square error between the forecasted (SST'_f) and observed (or simulated by SAM0) monthly SST anomalies (SST'), respectively, and σ is the standard deviation.

3. Results

Figure 2 shows the Niño3.4 SST prediction skill of the LIM ($w = 0$), MA ($w = 1$), and MA-LIM ($w = 0.15$), as measured by ACC and RMSSS as a function of the lead month, τ. In the perfect model analysis, the LIM has higher skill than the MA for $\tau < 4$ but the MA is more skillful at longer leads (Figure 2a). The lower skill of the MA than the LIM at short leads could be due to the errors in the initial condition of the MA: if more accurate initial conditions were used (possible with a longer control run), the MA would likely have a better skill at short leads, too. In contrast to the MA, the MA-LIM ($0 < w < 1$) starts from the same initial condition as the LIM. The performance of the MA-LIM with $w > 0.5$ is roughly between those of the LIM and MA at small τ but becomes similar to that of the MA at large τ because the effect of nudging to the MA is accumulated with time. Surprisingly, the MA-LIM with $w = 0.15$ has a better ACC than the LIM and MA at all τ. The analysis with the RMSSS statistics shows similar results, except that the overall relative performance of the MA to that of the LIM is slightly worse than the one measured by ACC and the crossover τ at which the performance of the MA is identical to that of the LIM has shifted from 4 to 5 months (Figure 2b).

In the observational hindcast, however, the crossover τ occurs at ~7 months, indicating that the relative performance of the MA with respect to the LIM is degraded from the perfect model to the observation analysis. This is an anticipated result, since the observationally based LIM also has higher Niño3.4 skill than the MME at short leads (Newman & Sardeshmukh, 2017). Also, apart from effects of initalization shock, which appear small in this region (Ding et al., 2018), the MA should not have higher skill than a corresponding forecast model initialized with the complete climate state. In fact, the SST/SSH anomaly fields in the model simulation are generally different from observations, due to model error, so at $\tau = 0$, the MA in the observational analysis has a lower ACC and RMSSS than does the perfect model analysis. The MA-LIM with a small value of w ($w = 0.15$) performs better than either the LIM and MA for all τ. Since the effect of nugding to the MA is accumulated with time, a small value of $w = 0.15$ does not imply that the nonlinearity captured by the MA is insignificant.

Figure 3 shows the spatial pattern of ACC in monthly SST anomalies obtained from the LIM at $\tau = 6$ and the differences between the MA or MA-LIM and the LIM (see Figure S2 for the RMSSS skill). Results are shown for the perfect model and observation analyses. The forecast skill of tropical SST anomalies has complex spatial variations. In the perfect model analysis, the overall spatial pattern of ACC from the LIM is similar to the well-known SST anomaly pattern associated with the positive phase of ENSO (Figure 3a) with a maximum ACC > 0.8 in the southern portion of the central equatorial Pacific Ocean. This is an anticipated result because ENSO is the dominant mode of tropical SST/SSH anomalies. In this region, both the LIM and MA have similar skill, demonstrating that ENSO evolution in the central Pacific in this GCM is well represented by linear dynamics (Newman & Sardeshmukh, 2017; Penland & Sardeshmukh, 1995). In the eastern equatorial Pacific, western Pacific, and Indian Oceans, the LIM skill is notably worse than the MA skill, suggesting that in these tropical regions, seasonality and/or predictable nonlinearity is important to state evolution. MA performance in other regions, such as the subtropics and Atlantic, is degraded relative to the LIM (Figure 3b). The MA-LIM with $w = 0.15$ retains the positive aspects of the MA and substantially remedies the undesirable aspects of the MA, resulting in domain-averaged ACC = 0.544 that is slightly larger than those of the MA (0.53) and LIM (0.51), with similar improvement seen for RMSSS.

Hindcast skill of the LIM, MA, and MA-LIM in the observational analysis is generally worse than for the perfect model analysis. While this could reflect the fewer samples in the shorter observational record and/or errors in the GCM's simulation of the tropical Pacific, it could also be that tropical Pacific variations in nature

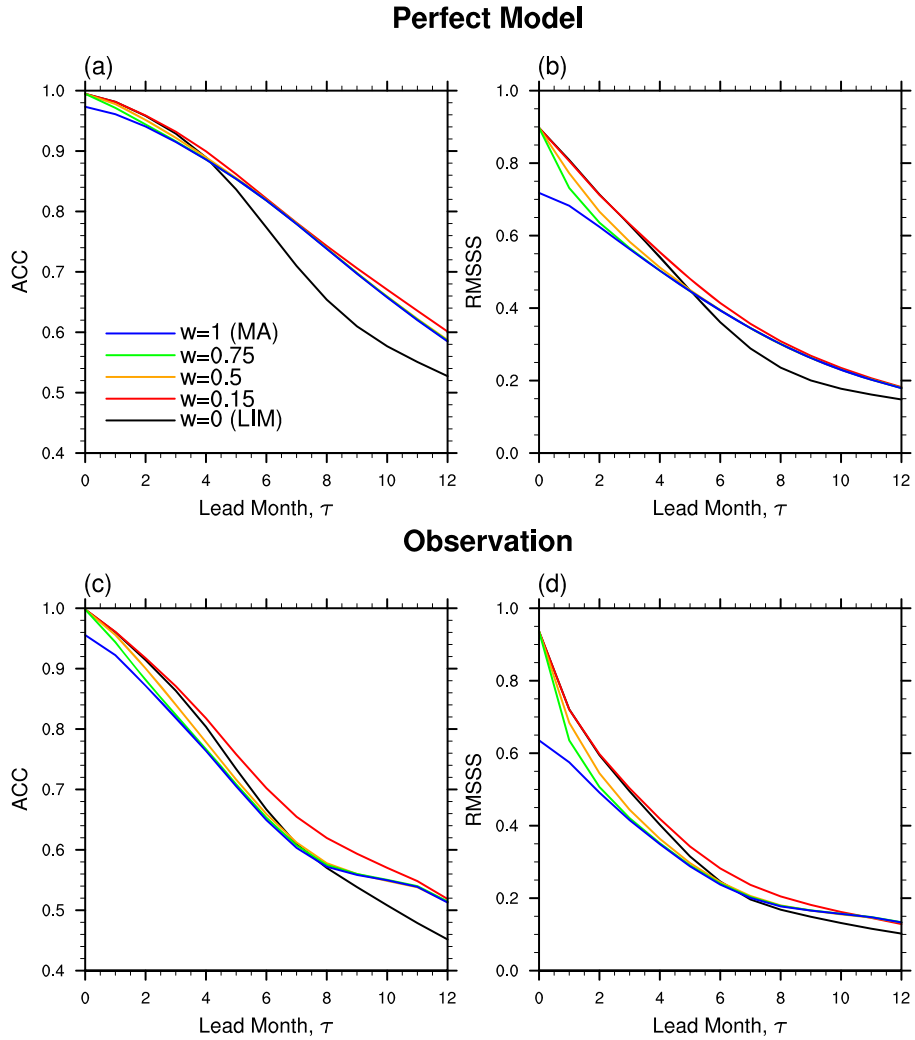

Figure 2. The (a, c) anomaly correlation coefficient (ACC) and (b, d) root-mean-square-error-based skill score (RMSSS) of the Niño3.4 SST anomalies (170°W-120°W, 5°S to 5°N) as a function of the lead month, τ, predicted by the LIM, MA, and MA-LIM with various weighting factor w, obtained from the (a, b) perfect model analysis and (c, d) observational analysis. At $\tau = 0$, both ACC and RMSSS are not exactly 1, since the initial monthly SST anomaly of the MA is computed by averaging 20 cases obtained from the GCM simulation, while the initial SST anomaly for the LIM and MA-LIM are obtained from the EOF reconstruction of 16/9 EOF modes of tropical SST/SSH anomalies, respectively.

are inherently less predictable than within the GCM simulation. Still, the overall spatial patterns of ACC and RMSSS obtained from the LIM in the observational analysis (Figures 3d and S2d) are similar to those in the perfect model analysis (Figures 3a and S2a). Note the similarity of the standard deviation of monthly SST anomalies from the SAM0 simulation and observations (compare solid lines in Figures 3a and 3d). In contrast to the perfect model analysis, except for the eastern Pacific, the MA are generally less skillful than the LIM; the domain-averaged ACC (0.307) and RMSSS (0.058) from the MA are substantially lower than those from the LIM (ACC = 0.379, RMSSS = 0.076). This may be due in part to model error: ENSO-related tropical SST anomalies simulated by most GCMs, including SAM0, typically extend too far westward (Park et al., 2019), and likewise SST skill in the western equatorial Pacific is significantly worse for both the MME and the GCMs that comprise it than for the LIM (Newman & Sardeshmukh, 2017). Because the analog evolution is drawn from the GCM simulation, the MA also performs considerably worse than the LIM in the western equatorial Pacific (Figure 3e), where SAM0 overestimates the observed standard deviation of monthly SST anomalies (solid lines in Figure 3d). Similar to the perfect model analysis, the MA-LIM shows a better performance than the LIM over the eastern equatorial Pacific and Indian Oceans. In addition, the

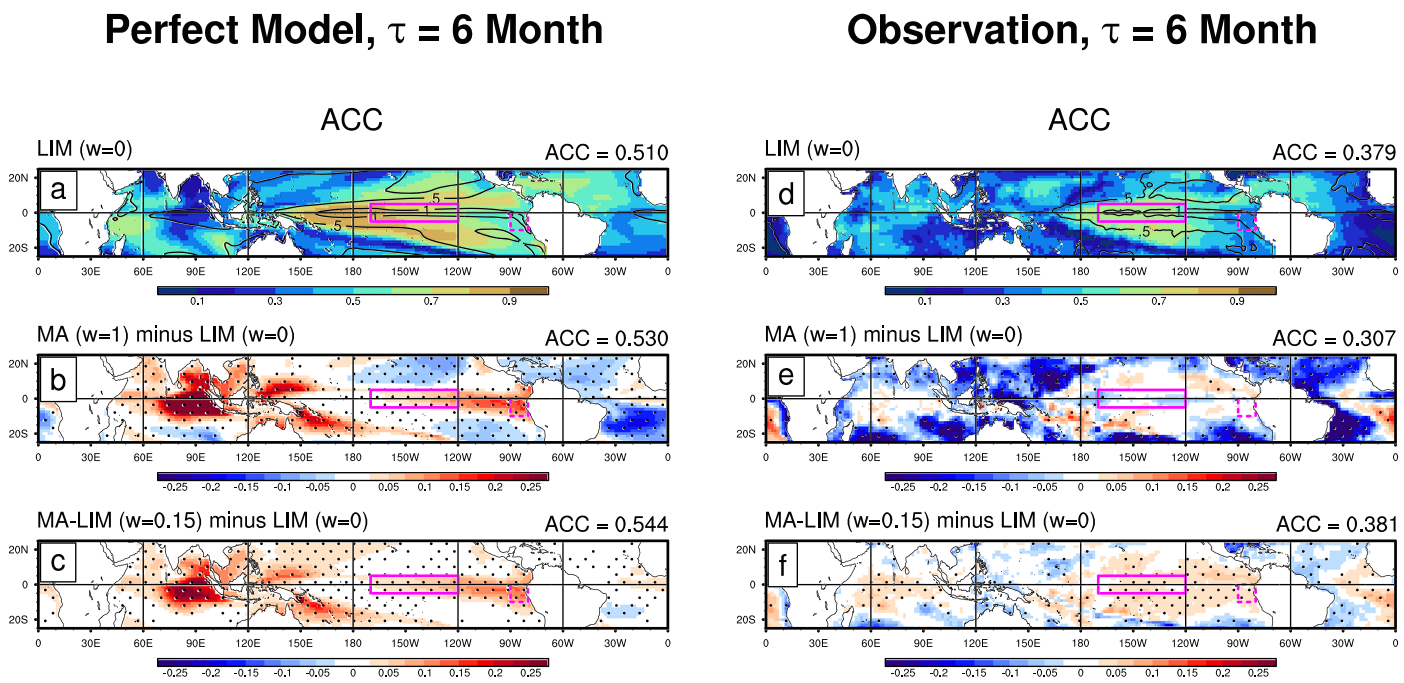

Figure 3. Spatial distributions of (a, d) ACC of the SST anomalies at the 6 month lead obtained from the (a) perfect model analysis and (d) observational analysis from LIM. The other figures show the differences between the MA or MA-LIM with $w = 0.15$ and LIM. The grid boxes with statistically significant ΔACC at the 95% confidence level from the bootstrapping (or Monte Carlo) method are denoted by the dot. A domain-averaged ACC in the tropical region (25°S to 25°N) for the LIM, MA, and MA-LIM is shown at the top right of an individual plot. The solid lines in (a) and (d) are the standard deviations of monthly SST anomaly obtained from the GCM simulation and observation, respectively, with a contour interval of 0.5°C. The solid and dashed pink boxes denote the Niño3.4 and Niño1.2 regions, respectively.

overall MA-LIM hindcast skill for $\tau = 6$ is better than that of the MA and LIM almost everywhere throughout the tropical Indo-Pacific.

Figure 4 shows ACC hindcast skill (see Figure S3 for the RMSSS skill) of the Niño3.4 anomalies as a function of the target month and τ for the three methods, presented by comparing LIM skill with the MA minus LIM and MA-LIM minus LIM difference plots. In all cases, the basic features are similar, with relatively lower prediction skill for target months in late spring and early summer. Similar to Figure 2, perfect model skill of the MA at short τ is again poorer than for the LIM, while the MA performing better than the LIM at longer leads, primarily for seasons where the performance of the LIM is relatively low. The MA-LIM with $w = 0.15$ retains the advantages of both the LIM at short τ and the MA at long τ, resulting in a better performance than the MA and LIM. The difference is more dramatic comparing the MA-LIM to the LIM in the Niño1.2 region (90–80°W, 0–10°S; Figures S4 and S5), primarily because the MA itself is relatively more skillful at longer lead there.

The results from the observational analysis are somewhat noisier than those from the perfect model, but show the same basic picture, with the MA-LIM improving LIM skill for predictions of boreal spring Niño3.4 values, where the LIM had lowest skill. For summer, LIM skill is near its minimum, but the MA had no better skill than the LIM and likewise the MA-LIM yielded minimal improvement. On the other hand, for Niño1.2 (Figures S4 and S5), while year-round skill for the MA-LIM exceeds both the MA and LIM (e.g., Figures 3d–3f), for wintertime verifications the MA-LIM is greatly improved relative to the LIM but is less skillful than the MA itself. The improvements of the SSH forecast skill by the MA-LIM are similar to those of SST both in terms of the spatial distribution and the dependency on τ and target month (Figures S6 and S7). However, in the observational analysis, the prediction skill of the MA-LIM in the Atlantic Ocean was in between those of the LIM and MA, where the MA showed a relatively poor prediction skill.

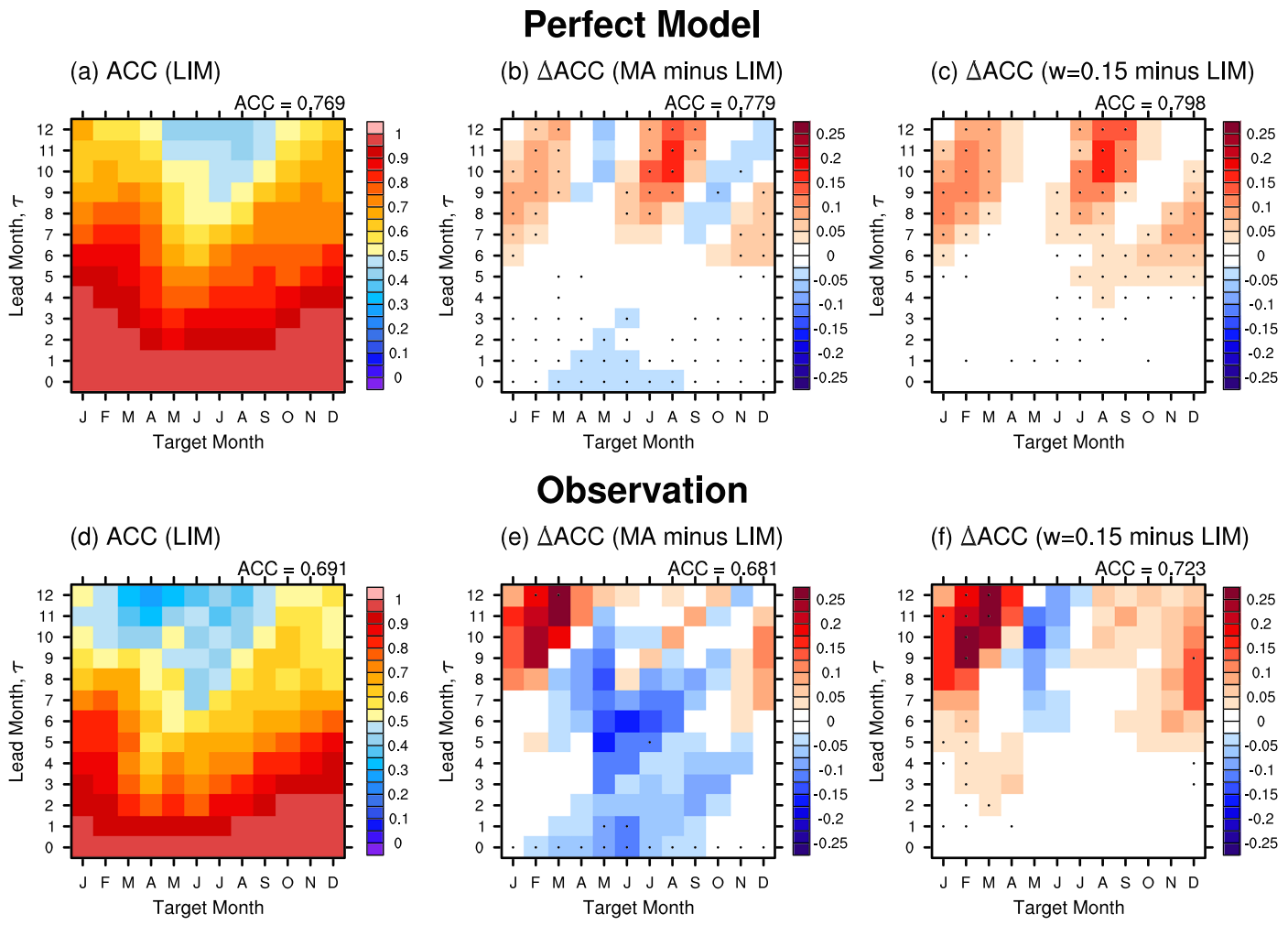

Figure 4. (a, d) ACC of the Niño3.4 SST anomalies as a function of the target month and lead month τ obtained from the (a–c) perfect model analysis and (d–f) observational analysis with the LIM. The differences between (b, e) the MA and LIM and (c, f) the MA-LIM with $w = 0.15$ and LIM are also shown. The domain-averaged ACC from the LIM, MA, and MA-LIM with $w = 0.15$ is shown at the top right of an individual plot. Statistically significant ΔACC at the 95% confidence level from the bootstrapping (or Monte Carlo) method is denoted by the dot.

4. Summary and Conclusions

To enhance the forecast skill of tropical SST anomalies including ENSO, we combined a model analog approach with a linear inverse model (MA-LIM). The MA-LIM nudges monthly SST/SSH anomalies forecasted by the LIM to those forecasted by the MA with an appropriate weighting w at each forecast month. The initial conditions of the MA-LIM are identical to those of the LIM. We compared the performances of the LIM ($w = 0$), MA ($w = 1$), and MA-LIM ($0 < w < 1$) using a long-term coupled GCM simulation (i.e., perfect model analysis) and observations. In predicting the GCM-simulated and observed Niño3.4 SST anomalies, the LIM performs better than the MA at short leads while the opposite is true at long leads. Surprisingly, the MA-LIM with $w = 0.15$ shows the best performance at all leads. The spatial distributions of the forecast skill at 6 month leads in the perfect model showed that the MA performs better than the LIM in the eastern equatorial Pacific, western Pacific, and Indian Oceans but worse in the eastern subtropical Pacific and Atlantic Oceans. On the other hand, the MA-LIM with $w = 0.15$ retains most of the beneficial aspects of the MA and substantially remedies the undesirable aspects of the MA, resulting in the overall best performance in the tropical oceans. Similar results are obtained from the observational analysis but the LIM is better in more areas.

Why does the MA-LIM perform better than either the MA or LIM? A major problem in the MA method is the uncertainty in the initial conditions, which is due in part to insufficient sampling. Because the MA-LIM

is designed to start from the LIM's initial condition that is more accurate than that of the MA, the MA-LIM performs better than the MA. On the other hand, the LIM used in our study is a stationary model, that is, the linear system matrix **L** or the multivariate covariance matrix **C** does not vary with time. The MA-LIM brings the implicit seasonal cycle and nonlinear dynamics in the MA into the LIM in an ad hoc manner, such that the MA-LIM performs better than the LIM. Another factor is that ENSO also has nonlinear dynamics (An & Jin, 2004), which are especially important in the eastern tropical Pacific and can be represented by MA but not LIM.

To further improve the performance of the MA-LIM, it may be necessary to sample more cases from more accurate populations (e.g., more models, improved models, and longer runs); refine the sampling and prediction methods of the MA (e.g., use SST/SSH tendencies instead of SST/SSH themselves); and incorporate seasonal cycle directly into **C** and **L** (e.g., LIM can have variability in the evolution of SST and SSH by constructive interference of the normal modes). We plan to investigate these ways to improve the MA-LIM methodology in the near future.

Acknowledgments
This work was supported by the research project of the National Institute of Meteorological Sciences (NIMS) of South Korea titled "Improving the physics parameterizations and dynamic core of climate prediction models." The data used in this paper are available at this site (ftp://nmlab.snu.ac.kr/datapublic/grl2019a/).

References

Alexander, M. A., Bladé, I., Newman, M., Lanzante, J. R., Lau, N.-C., & Scott, J. D. (2002). The atmospheric bridge: The influence of ENSO teleconnections on air–sea interaction over the global oceans. *Journal of Climate*, *15*(16), 2205–2231.

Alexander, M. A., Lau, N.-C., & Scott, J. D. (2004). Broadening the atmospheric bridge paradigm: ENSO teleconnections to the tropical West Pacific-Indian Oceans over the seasonal cycle and to the North Pacific in summer. *Earth's Climate: The Ocean-Atmosphere Interaction, Geophysical Monograph*, *147*, 85–103.

An, S.-I., & Jin, F.-F. (2004). Nonlinearity and asymmetry of ENSO. *Journal of Climate*, *17*(12), 2399–2412.

Balmaseda, M. A., Mogensen, K., & Weaver, A. T. (2013). Evaluation of the ECMWF ocean reanalysis system ORAS4. *Quarterly Journal of the Royal Meteorological Society*, *139*(674), 1132–1161.

Bellenger, H., Guilyardi, É., Leloup, J., Lengaigne, M., & Vialard, J. (2014). ENSO representation in climate models: From CMIP3 to CMIP5. *Climate Dynamics*, *42*(7-8), 1999–2018.

Ding, H., Newman, M., Alexander, M. A., & Wittenberg, A. T. (2018). Skillful climate forecasts of the tropical Indo-Pacific Ocean using model-analogs. *Journal of Climate*, *31*(14), 5437–5459.

Ding, H., Newman, M., Alexander, M. A., & Wittenberg, A. T. (2019). Diagnosing secular variations in retrospective ENSO seasonal forecast skill using CMIP5 model-analogs. *Geophysical Research Letters*, *46*, 1721–1730. https://doi.org/10.1029/2018GL080598

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development (Online)*, *9*, 10,539–10,583. (LLNL-JRNL-736881).

L'Heureux, M., A. Levine, M. Newman, C. Ganter, J.-J. Luo, M. Tippett, & T. Stockdale (2020). "Chapter 10: ENSO Prediction" *AGU Monograph: ENSO in a Changing Climate*, M. McPhaden, A. Santoso, W. Cai (Eds.), Wiley, in press.

Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., et al. (2013). The community earth system model: A framework for collaborative research. *Bulletin of the American Meteorological Society*, *94*(9), 1339–1360.

Lau, N.-C., & Nath, M. J. (1996). The role of the "atmospheric bridge" in linking tropical pacific ENSO events to extratropical SST anomalies. *Journal of Climate*, *9*(9), 2036–2057.

Li, G., & Xie, S.-P. (2014). Tropical biases in CMIP5 multimodel ensemble: The excessive equatorial Pacific cold tongue and double ITCZ problems. *Journal of Climate*, *27*(4), 1765–1780.

Neale, R. B., Chen, C.-C., Gettelman, A., Lauritzen, P. H., Park, S., Conley, A.J., et al. (2010). Description of the NCAR community atmosphere model (CAM 5.0). *NCAR Tech. Note NCAR/TN-486+ STR*, *1*(1), 1–12.

Newman, M., & Sardeshmukh, P. D. (2017). Are we near the predictability limit of tropical Indo-Pacific sea surface temperatures? *Geophysical Research Letters*, *44*, 8520–8529. https://doi.org/10.1002/2017GL074088

Park, S. (2004c). Remote ENSO influence on mediterranean sky conditions during late summer and autumn: Evidence for a slowly evolving atmospheric bridge. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, *130*(602), 2409–2422.

Park, S. (2014a). A unified convection scheme (UNICON). Part I: Formulation. *Journal of the Atmospheric Sciences*, *71*(11), 3902–3930.

Park, S. (2014b). A unified convection scheme (UNICON). Part II: Simulation. *Journal of the Atmospheric Sciences*, *71*(11), 3931–3973.

Park, S., Baek, E.-H., Kim, B.-M., & Kim, S.-J. (2017). Impact of detrained cumulus on climate simulated by the community atmosphere model version 5 with a unified convection scheme. *Journal of Advances in Modeling Earth Systems*, *9*, 1399–1411. https://doi.org/10.1002/2016MS000877

Park, S., & Bretherton, C. S. (2009). The University of Washington shallow convection and moist turbulence schemes and their impact on climate simulations with the community atmosphere model. *Journal of Climate*, *22*(12), 3449–3469.

Park, S., Bretherton, C. S., & Rasch, P. J. (2014). Integrating cloud processes in the community atmosphere model, version 5. *Journal of Climate*, *27*(18), 6821–6856.

Park, S., & Leovy, C. B. (2004). Marine low-cloud anomalies associated with ENSO. *Journal of Climate*, *17*(17), 3448–3469.

Park, S., Shin, J., Kim, S., Oh, E., & Kim, Y. (2019). Global climate simulated by the Seoul National University atmosphere model version 0 with a unified convection scheme (SAM0-UNICON). *Journal of Climate*, *32*(10), 2917–2949.

Penland, C., & Sardeshmukh, P. D. (1995). The optimal growth of tropical sea surface temperature anomalies. *Journal of climate*, *8*(8), 1999–2024.

Rasmusson, E. M., & Carpenter, T. H. (1982). Variations in tropical sea surface temperature and surface wind fields associated with the Southern Oscillation/El Niño. *Monthly Weather Review*, *110*(5), 354–384.

Rayner, N., Parker, D. E., Horton, E., Folland, C. K., Alexander, L. V., Rowell, D., et al. (2003). Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research*, *108*(D14), 4407. https://doi.org/10.1029/2002JD002670

Ropelewski, C. F., & Halpert, M. S. (1987). Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation. *Monthly weather review*, *115*(8), 1606–1626.

Rowell, D. P. (2001). Teleconnections between the tropical Pacific and the Sahel. *Quarterly Journal of the Royal Meteorological Society*, *127*(575), 1683–1706.

Trenberth, K. E., Branstator, G. W., Karoly, D., Kumar, A., Lau, N.-C., & Ropelewski, C. (1998). Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures. *Journal of Geophysical Research*, *103*(C7), 14,291–14,324.

Winkler, C. R., Newman, M., & Sardeshmukh, P. D. (2001). A linear model of wintertime low-frequency variability. Part I: Formulation and forecast skill. *Journal of Climate*, *14*(24), 4474–4494.

Zhang, G. J., & McFarlane, N. A. (1995). Sensitivity of climate simulations to the parameterization of cumulus convection in the Canadian Climate Centre general circulation model. *Atmosphere-ocean*, *33*(3), 407–446.

Zheng, Y., Lin, J.-L., & Shinoda, T. (2012). The equatorial Pacific cold tongue simulated by IPCC AR4 coupled GCMs: Upper ocean heat budget and feedback analysis. *Journal of Geophysical Research*, *117*, C05024. https://doi.org/10.1029/2011JC007746