

On the sensitive areas for targeted observations in ENSO forecasting

Jingjing Zhang^a, Shujuan Hu^{a,*}, Wansuo Duan^{b,c,*}



^aCollege of Atmospheric Sciences, Lanzhou University, Lanzhou, China

^bState Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics (LASG), Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

^cUniversity of Chinese Academy of Sciences, Beijing, China

ARTICLE INFO

Keywords:

Targeted observation

ENSO

Particle filter

Initial error

关键词:

目标观测

ENSO

粒子滤波

初始误差

ABSTRACT

Using the outputs from CMCC-CM in CMIP5 experiments, the authors identified sensitive areas for targeted observations in ENSO forecasting from the perspective of the initial error growth (IEG) method and the particle filter (PF) method. Results showed that the PF targets areas over the central-eastern equatorial Pacific, while the sensitive areas determined by the IEG method are slightly to the east of the former. Although a small part of the areas targeted by the IEG method also lie in the southeast equatorial Pacific, this does not affect the large-scale overlapping of the sensitive areas determined by these two methods in the eastern equatorial Pacific. Therefore, sensitive areas determined by the two methods are mutually supportive. When considering the uncertainty of methods for determining sensitive areas in realistic targeted observation, it is more reasonable to choose the above overlapping areas as sensitive areas for ENSO forecasting. This result provides scientific guidance for how to better determine sensitive areas for ENSO forecasting.

摘要

本研究分别从初始误差增长和粒子滤波同化的角度识别了ENSO预测的目标观测敏感区。结果表明:与粒子滤波同化方法确定的敏感区相比,基于初始误差增长确定的敏感区位置稍微偏东,且在东南太平洋也有分布,但整体而言,这两种方法确定的敏感区存在大范围重合区域,是互为印证的。在实际目标观测中,如果考虑使用确定敏感区方法的不确定性,那么选择上述两种方法确定的敏感区的重合区域作为ENSO预测的目标观测敏感区将更为合理。

1. Introduction

El Niño–Southern Oscillation (ENSO) is the dominant mode of the large-scale coupled ocean–atmosphere in the tropical Pacific. The occurrence of ENSO not only has direct impacts on the atmosphere and ocean over the Pacific regions, but also plays an important role in modulating the global weather and climate anomalies through teleconnection (Cane, 1983; Sakai and Kawamura, 2009). Accordingly, it is important to improve the forecasting skill for ENSO.

However, there are still many uncertainties in realistic ENSO forecasting. In particular, the impact of the “spring predictability barrier” (SPB) is one of the main reasons for the large deviation in ENSO forecasting (Duan and Hu, 2016). The SPB phenomenon refers to the fact that most models when forecasting ENSO often show a significant decline in their forecasting skill during boreal spring and/or the beginning of summer, resulting in large forecast uncertainties (Duan et al., 2009a; Zhang et al., 2014). Zhang et al. (2014) summarized previous studies and found that initial errors with specific spatial distribution characteristics are more likely to cause a notable SPB, and such initial errors are

often located in several key areas. This may offer some useful information on sensitive areas for targeted observations in ENSO forecasting.

The methods used in previous studies to determine sensitive areas for targeted observations can be roughly divided into two categories (Toth and Kalnay, 1997; Bishop and Toth, 1999; Baker and Daley, 2000; Hamill and Snyder, 2002). The methods in the first category—e.g., the linear singular vector (Palmer et al., 1998) and conditional nonlinear optimal perturbation (Mu et al., 2003; Duan and Mu, 2009b)—begin by obtaining the initial errors that have the greatest impact on the forecast, and then the areas with larger and more concentrated initial errors can be determined as the sensitive areas. These methods are based on the initial error growth (IEG), which measures the sensitivity of forecast errors to initial errors. The methods in the other category are aimed at reducing the forecast error variance directly. Those regions where the reduction in the forecast error variance is maximized are deemed as the sensitive areas for targeted observations via data assimilation. The ensemble transform Kalman filter (ETKF; Bishop et al., 2001) and the particle filter (PF; Van Leeuwen, 2009; Kramer and Dijkstra, 2013; Duan et al., 2018) methods, which are based on ensemble dispersion to

* Corresponding authors.

E-mail addresses: hushuju@lzu.edu.cn (S. Hu), duanws@lasg.iap.ac.cn (W. Duan).

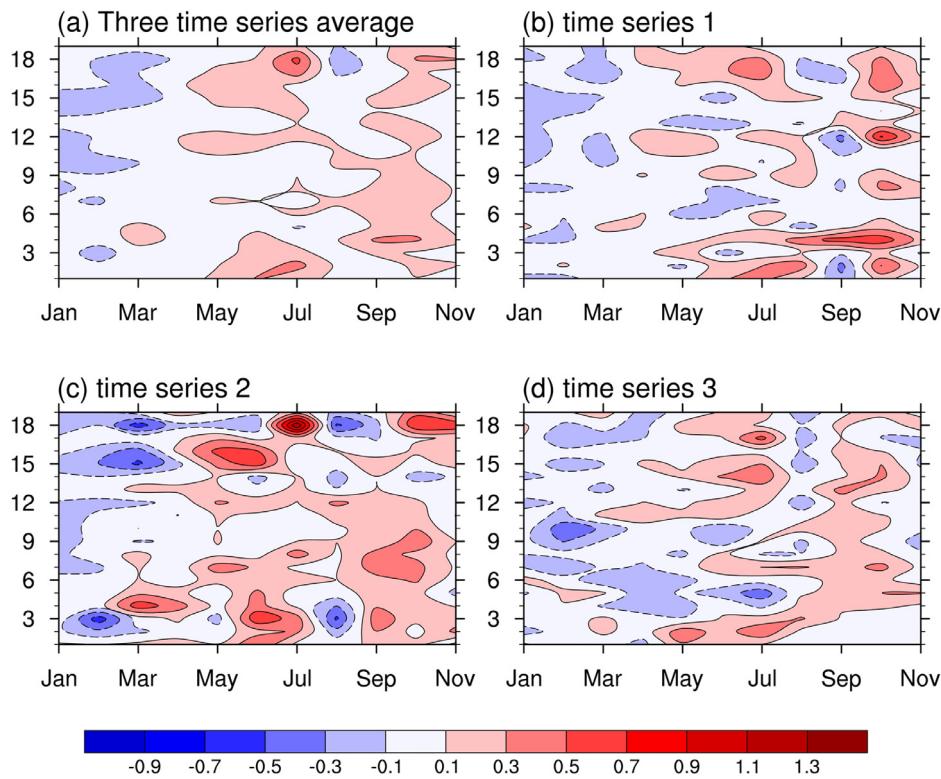


Fig. 1. Monthly mean growth rates of prediction errors for El Niño events. The vertical axes denote the prediction samples. The contour lines represent the monthly growth rates of the increase (or decrease) in prediction errors, where positive values indicate growth of the prediction.

measure the sensitivity of forecast errors to initial errors, are two examples of this category of methods. ETKF is currently a popular data assimilation method, but is based on linear and Gaussian assumptions (Vetra-Carvalho et al., 2018). Meanwhile, due to scientific and technological advances, dynamical models have become increasingly nonlinear, meaning it is more reasonable to require data assimilation methods that can handle non-Gaussian distributions, which cause the limitation of ETKF to be amplified (Vetra-Carvalho et al., 2018). The PF method has developed rapidly under such development needs because, although it and ETKF can be unified through Bayes theorem, the PF method holds the prospect of completely nonlinear data assimilation and is not limited to Gaussian distributions (Vetra-Carvalho et al., 2018; Van Leeuwen et al., 2019).

In the above two ideas for determining sensitive areas, we ask the following questions: Can the sensitive areas determined from the perspective of IEG and those determined by the PF method be mutually supportive? What are the similarities and differences between them? When considering the uncertainty of these methods for determining sensitive areas, how should we then determine more reasonable sensitive areas? By answering these questions, we hope to provide a scientific reference for determining the sensitive areas for targeted observations in ENSO forecasting.

2. Data and methods

2.1. Data

We used the monthly mean sea surface temperature (SST) data of the pre-industrial control (pi-control) runs from the CMIP5 experiments of CMCC-CM (Centro Euro-Mediterraneo sui Cambiamenti Climatici Climate Model). The pi-control runs, which are mainly used to analyze the internal variability of models with time-invariant forcing, are a reference for historical and climate sensitivity experiments. We randomly chose three 20-year time series from the coupled model, and from each

series we identified three typical El Niño events, i.e., a total of nine events. This paper studies typical El Niño events with warming in early boreal spring and peaking at the end of the year (Fig. S1).

2.2. Methods for determining sensitive areas for targeted observations

2.2.1. Method based on IEG

To study the distribution features of initial errors, we need to judge whether the SPB phenomenon for El Niño events exists in CMCC-CM. First of all, the SST of each typical El Niño year was treated as the “observation”, and then the other 19 years of SSTs in each time series were regarded as 19 “predictions” of the “observation”. Each “observation” corresponded to 19 “predictions”. Then, following the definition of Mu et al. (2007) regarding the growth tendency of prediction errors, we were able to estimate the monthly growth tendency of prediction errors. According to the season-dependent evolutions of prediction errors, we could judge whether the SPB existed in CMCC-CM. That is, if the growth tendency of the prediction errors reached its maximum in spring and/or the beginning of summer, we were able to confirm that the SPB phenomenon for El Niño events existed in CMCC-CM.

In this study, because we adopted the pi-control runs, we assumed that prediction errors were only caused by initial errors. For the predictions that yielded an SPB, we studied the corresponding initial errors through empirical orthogonal function (EOF) analysis. Then, we determined the areas with larger and more concentrated initial errors as the sensitive areas for targeted observations.

2.2.2. PF method

The PF method is an assimilation method that uses the Monte Carlo algorithm to achieve Bayes theorem (Duan et al., 2018). The core of the PF method is to adjust the weight of particles by using ‘sequential importance sampling’ (Van Leeuwen, 2009; Kramer et al., 2012; Kramer and Dijkstra, 2013). The specific details of the PF method can be referred to in Kramer et al. (2012) and Kramer and Dijkstra (2013).

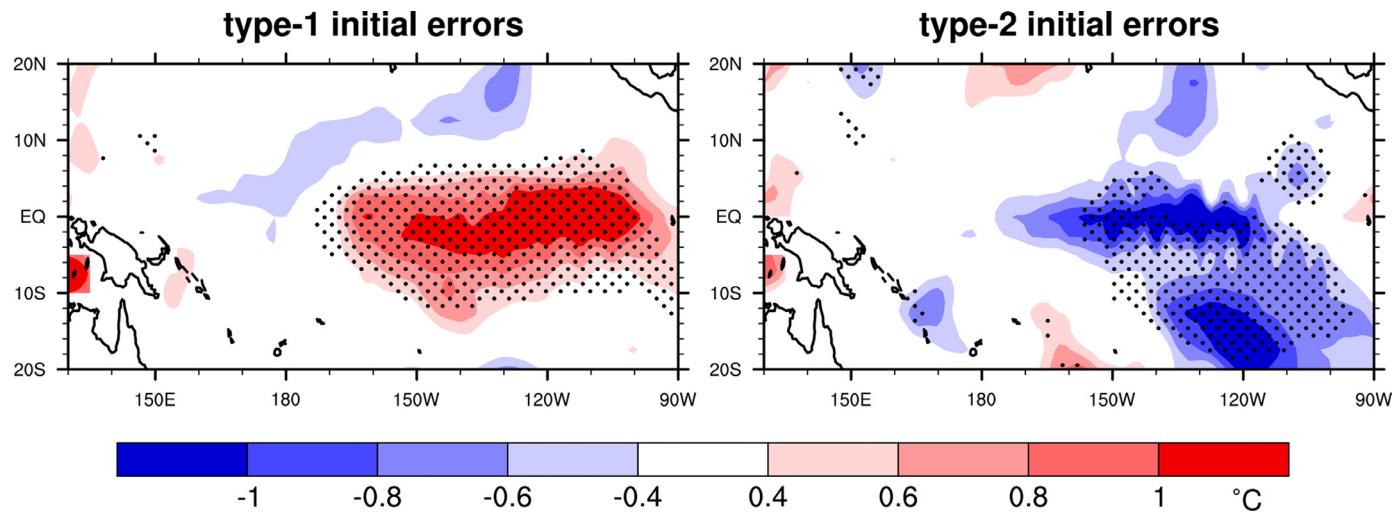


Fig. 2. Composite patterns of SSTAs for two types of SPB initial errors. Dotted areas indicate the composites of SSTAs exceed the 95% confidence level.

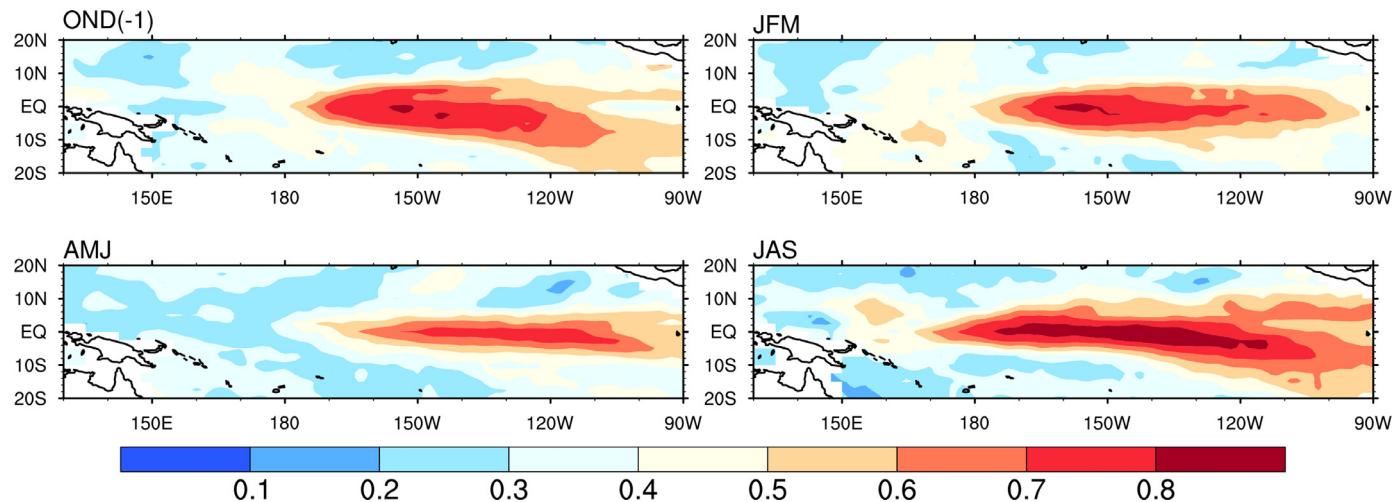


Fig. 3. Ensemble mean of the PP index obtained by assimilating the three-month observations of SST at OND(-1), JFM, AMJ, and JAS, respectively.

Using the PF method to determine sensitive areas involved first splitting the 330-year SST integration from CMCC-CM into one-year segments, which yielded 330 ensemble members (i.e., particles). Since the external forcing of the pi-control runs is constant, the ensemble could be regarded as having an identical distribution. The *prior* probability density function (PDF) of the system state will be the climatological PDF given by these ensemble members. Then, we adjusted the weights of ensemble members by assimilating the observations. The observations here were not real as such, but simulated. It was necessary to take the SST in each typical El Niño year selected in the previous part of this study as the “true value”, and superimpose the random error on the “true value” to produce an idealized observation. Therefore, the new ensemble could be obtained by continuously assimilating the three-month idealized observations, and the *posterior* PDF was given by these new ensemble members.

In this study, we used the predictive power (PP; Schneider and Griffies, 1999)—an entropy-based metric—to measure the degree of reduction in the uncertainty of the *posterior* PDF relative to the uncertainty in the *prior* PDF. The PP is limited to the range $0 \leq \text{PP} \leq 1$. The larger the PP, the greater the degree of reduction in the forecast uncertainty. Therefore, the regions with high PP were determined as the sensitive areas for targeted observations.

3. Sensitive areas determined by the IEG method

According to the method described in Section 2.2.1, we are able to judge whether the SPB phenomenon exists in CMCC-CM. Fig. 1(a) shows the monthly mean growth rates of prediction errors for the nine El Niño events. From Fig. 1(a) we can see that prediction errors usually start to grow in April, and the most significant growth occurs around June. That is, we can demonstrate that the SPB phenomenon exists in CMCC-CM. Furthermore, we can study the monthly mean growth rates of prediction errors in each time series (Fig. 1(b-d)), from which we can see that the time when the prediction errors start to grow rapidly and the growth rates of prediction errors are different in each time series. The results indicate that the SPB may have interdecadal characteristics. From Fig. 1 we can also conclude that some of the above predictions yield the SPB phenomenon, while others cannot yield it. As mentioned above, the prediction errors are only caused by initial errors. Nevertheless, some initial errors are obviously too large to be compared with actual initial SST errors. Thus, we selected predictions with initial errors less than 0.5°C , and we refer to the initial errors yielding the SPB as “SPB initial errors” for convenience.

But what are the spatial distribution characteristics of the SPB initial errors? To address this, we used EOF analysis to extract the dominant

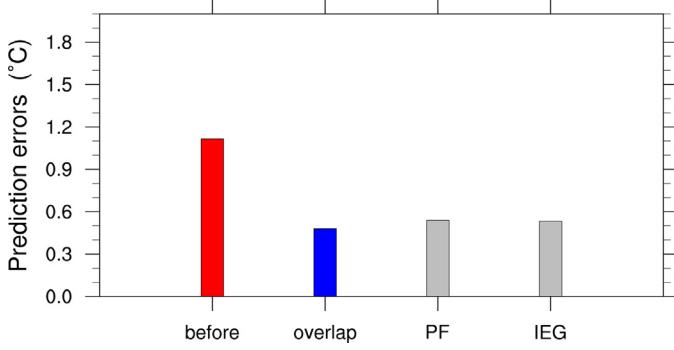


Fig. 4. Mean prediction errors averaged over April–December in the ensembles, obtained after assimilating SST in JFM at the first 40 maximum points from the overlapping areas (blue bar), PF areas and IEG areas (gray bars), respectively. As a reference, the red bar represents the counterpart before assimilation.

mode (i.e., EOF1) of the SPB initial errors, and then selected original initial errors with the same and opposite sign as EOF1 to form two types of initial errors, and finally obtain two main patterns of SPB initial errors through composite analysis (Fig. 2). One of the patterns possesses a positive SST anomaly (SSTA) pattern in the eastern equatorial Pacific, while the other consists of a negative SSTA component that is located in the eastern equatorial Pacific and the southeast equatorial Pacific. It can be seen that the distribution of the main energy of the SPB initial errors has local characteristics; that is, the distribution of errors has large-value areas.

Overall, the main initial errors yielding the SPB are concentrated in the eastern equatorial Pacific (10°N – 10°S , 150° – 100°W) and the south-east equatorial Pacific (10° – 20°S , 140° – 110°W), which means that initial errors in these areas are more likely to cause the SPB and produce large prediction errors. Therefore, these areas may represent sensitive areas for ENSO forecasting. If we deploy additional observations in these areas, and then assimilate observations to the initial field of the model, the forecasting skill for El Niño events could be significantly improved, as compared to doing so in other areas.

4. Sensitive areas determined by the PF method

In this next part of our study we used the PF method to determine the sensitive areas of the nine El Niño events. Fig. 3 shows the PP index (average for the nine El Niño events) obtained by assimilating the three-month observations of SST at OND(-1), JFM, AMJ, and JAS, respectively, where OND(-1) represents the period from October to December of the year before the typical year, JFM represents the period from January to March of the typical year, and so forth. As mentioned above, the regions with high PP can be determined as the sensitive areas for targeted observations. Therefore, from Fig. 3 we can see that the positions of the sensitive areas obtained by assimilating the observations in different seasons are not much different and basically located in the central-eastern equatorial Pacific ($\sim 10^{\circ}\text{N}$ – 10°S , 180° – 120°W). However, the values of the PP index have some differences in different seasons. The PP obtained by assimilating observations in AMJ is the smallest. Moreover, the PP values obtained by assimilating the observations in JAS are bigger than the assimilation results in other seasons. This

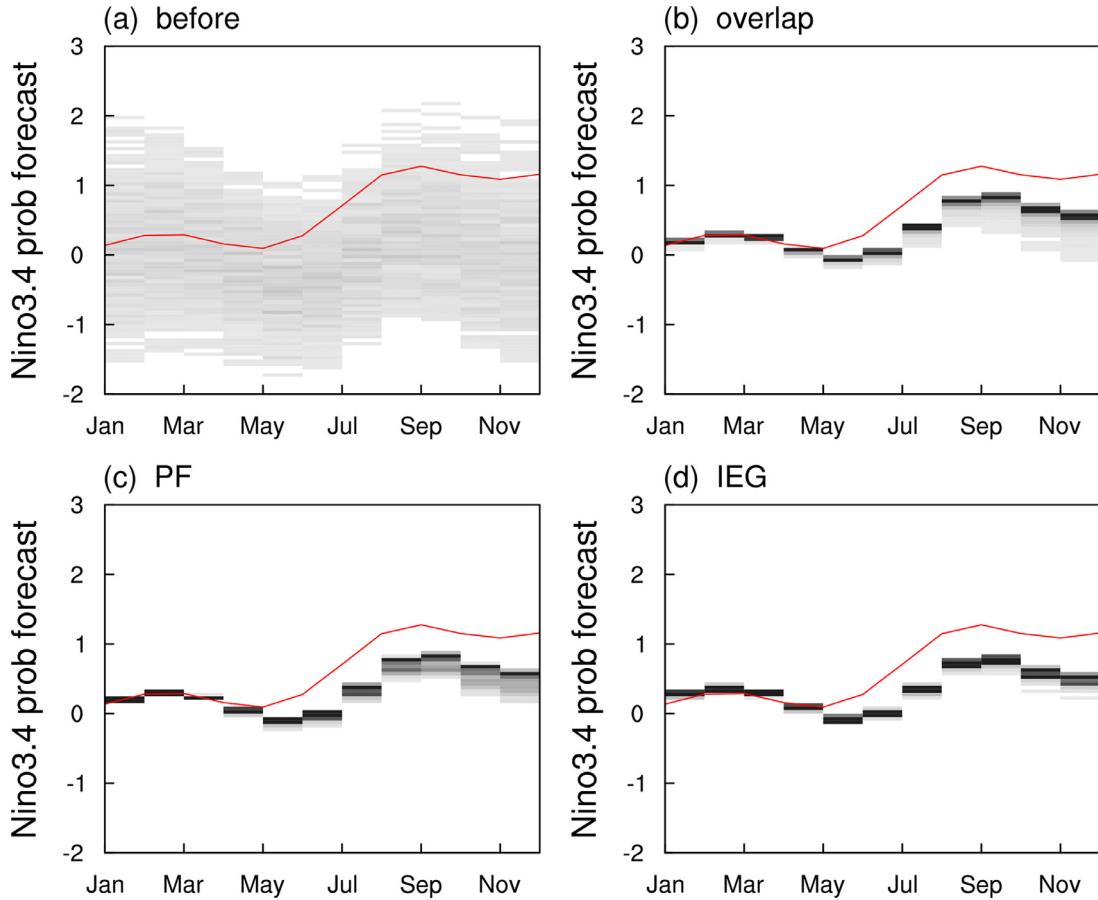


Fig. 5. (a) Climatological probability distribution of Niño-3.4 index before assimilation. (b–d) Probability distribution of Niño-3.4 index, obtained after assimilating SST in JFM in the (b) overlapping areas, (c) PF areas, and (d) IEG areas. The red line represents the development of the “true value”. The gray shading represents the probability value, and the darker the color, the greater the probability.

is because JAS is the season when the development of El Niño events strengthens and the signal-to-noise ratio for SST is relatively large.

5. Effectiveness of the sensitive areas

Through the previous analysis we found that there is a large-scale overlap (10°N – 10°S , 150° – 120°W) of the sensitive areas determined by the above two methods in the equatorial eastern Pacific. Therefore, the sensitive areas determined by the two methods are mutually supportive. When considering the uncertainty of the methods in realistic targeted observations, it is more reasonable to choose the above overlapping areas as the sensitive areas for ENSO forecasting.

To illustrate this inference, we selected the first 40 maximum points from the above overlapping areas, the sensitive areas determined by the PF method (i.e., PF areas), and the sensitive areas determined by the IEG method (i.e., IEG areas), respectively. Next, we used the PF assimilation method to assimilate the SST in JFM for each group of grid points, and then let the three groups of ensemble members obtained after assimilation to develop freely during the following April–December. By comparing the prediction skills of the three ensembles, it was possible to verify whether or not the overlapping areas were valid for improving the El Niño forecast skill.

Fig. 4 shows the mean prediction errors averaged over April–December in the three ensembles, obtained by assimilating the SST in JFM at the first 40 maximum points from the overlapping areas (blue bar), PF areas and IEG areas (gray bars), respectively. From **Fig. 4** we can see that the mean prediction errors obtained by assimilating SST in the overlapping areas is obviously smaller than the mean before assimilation, which is also smaller than the mean in the PF areas and IEG areas. That is, choosing the overlapping areas as the sensitive areas is reasonable and effective, and is more effective than adopting the PF areas or IEG areas.

Finally, we analyzed the reasons why the overlapping areas are more valid, from the perspective of the probability distribution. **Fig. 5** shows the probability distribution of the Niño-3.4 index, obtained by assimilating the SST in JFM in the overlapping areas, PF areas, and IEG areas, respectively. Compared with the climatological probability distribution (**Fig. 5(a)**), the spreads of the three ensembles obtained after assimilating observations in the above three areas decrease to varying degrees. At the assimilation stage, the three groups of ensemble members are all located near the “true value”, and over time the errors gradually increase at the forecast stage; but overall, the development of all members is skewed to the warm events. The number of members skewed to the warm events after assimilating SST in the overlapping areas is more than its counterparts in the other two areas, and the errors with the “true value” after assimilating SST in the overlapping areas are lowest.

6. Conclusions

In this study we identified the sensitive areas for targeted observations for ENSO forecasting from the perspective of the IEG method and PF method. Results showed that the sensitive areas determined by the PF method are targeted in the central-eastern equatorial Pacific ($\sim 10^{\circ}\text{N}$ – 10°S , 180° – 120°W), while their counterparts determined by the IEG method are concentrated in the eastern equatorial Pacific (10°N – 10°S , 150° – 100°W) and the southeast equatorial Pacific (10° – 20°S , 140° – 110°W). We found that the sensitive areas determined by the above two methods overlap widely in the eastern equatorial Pacific (10°N – 10°S , 150° – 120°W). Therefore, the sensitive areas determined by the above two methods are mutually supportive. When considering the uncertainty of the methods in realistic targeted observation, it would be more reasonable to choose the above overlapping areas as the sensitive areas for target observations in ENSO forecasting. By comparing the prediction skills of the three ensembles, obtained after assimilating SST in the

overlapping areas, PF areas, and IEG areas, respectively, it was further verified that the overlapping areas are more valid for improving ENSO forecast skill. The approach in this study to determining sensitive areas provides a reference for further research on targeted observations.

Funding

This study was supported by the National Natural Science Foundation of China [grant numbers 41930971, 41775069, and 41975076].

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.aosl.2021.100054](https://doi.org/10.1016/j.aosl.2021.100054).

References

- Baker, N.L., Daley, R., 2000. Observation and background adjoint sensitivity in the adaptive observation-targeting problem. *Q. J. R. Meteorol. Soc.* 126 (565), 1431–1454. doi:[10.1002/qj.49712656511](https://doi.org/10.1002/qj.49712656511).
- Bishop, C.H., Toth, Z., 1999. Ensemble transformation and adaptive observations. *J. Atmos. Sci.* 56 (11), 1748–1765. doi:[10.1175/1520-0469\(1999\)0562.0.CO;2](https://doi.org/10.1175/1520-0469(1999)0562.0.CO;2).
- Bishop, C.H., Etherton, B.J., Majumdar, S.J., 2001. Adaptive sampling with the ensemble transform Kalman filter. Part I: theoretical aspects. *Mon. Weather Rev.* 129 (3), 420–436. doi:[10.1175/1520-0493\(2001\)129<0420:ASWTET>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<0420:ASWTET>2.0.CO;2).
- Cane, M.A., 1983. Oceanographic events during El Niño. *Science* 222 (4629), 1189–1195. doi:[10.1126/science.222.4629.1189](https://doi.org/10.1126/science.222.4629.1189).
- Duan, W.S., Hu, J.Y., 2016. The initial errors that induce a significant “spring predictability barrier” for El Niño events and their implications for target observation: results from an earth system model. *Clim. Dynam.* 46, 3599–3615. doi:[10.1007/s00382-015-2789-5](https://doi.org/10.1007/s00382-015-2789-5).
- Duan, W.S., Mu, M., 2009b. Conditional nonlinear optimal perturbation: applications to stability, sensitivity, and predictability. *Sci. China Ser. D.* 52 (7), 883–906. doi:[10.1007/s11430-009-0090-3](https://doi.org/10.1007/s11430-009-0090-3).
- Duan, W.S., Feng, F., Hou, M.Y., 2018. Application of particle filter assimilation in the target observation for El Niño-Southern oscillation. *Chin. J. Atmos. Sci.* 042 (003), 677–695 (In Chinese). doi:[10.3878/j.issn.1006-9895.1711.17264](https://doi.org/10.3878/j.issn.1006-9895.1711.17264).
- Duan, W.S., Liu, X.C., Zhu, K.Y., Mu, M., 2009a. Exploring the initial errors that cause a significant “spring predictability barrier” for El Niño events. *J. Geophys. Res-Oceans* 114 (C04022). doi:[10.1029/2008JC004925](https://doi.org/10.1029/2008JC004925).
- Hamill, T.M., Snyder, C., 2002. Using improved background-error covariances from an ensemble Kalman filter for adaptive observations. *Mon. Weather Rev.* 130 (6), 1552–1572. doi:[10.1175/1520-0493\(2002\)1302.0.CO;2](https://doi.org/10.1175/1520-0493(2002)1302.0.CO;2).
- Kramer, W., Dijkstra, H.A., 2013. Optimal localized observations for advancing beyond the ENSO predictability barrier. *Nonlinear Proc. Geoph.* 20 (2), 221–230. doi:[10.5194/npg-20-221-2013](https://doi.org/10.5194/npg-20-221-2013).
- Kramer, W., Dijkstra, H.A., Pierini, S., Van Leeuwen, P.J., 2012. Measuring the impact of observations on the predictability of the Kuroshio Extension in a shallow-water model. *J. Phys. Oceanogr.* 42 (1), 3–17. doi:[10.1175/JPO-D-11-014.1](https://doi.org/10.1175/JPO-D-11-014.1).
- Mu, M., Duan, W.S., Wang, B., 2003. Conditional nonlinear optimal perturbation and its applications. *Nonlinear Proc. Geoph.* 10 (6), 493–501. doi:[10.5194/npg-10-493-2003](https://doi.org/10.5194/npg-10-493-2003).
- Mu, M., Duan, W.S., Wang, B., 2007. Season-dependent dynamics of nonlinear optimal error growth and El Niño-Southern Oscillation predictability in a theoretical model. *J. Geophys. Res.* 112 (D10113). doi:[10.1029/2005JD006981](https://doi.org/10.1029/2005JD006981).
- Palmer, T.N., Gelaro, R., Barkmeijer, J., Buizza, R., 1998. Singular vectors, metrics, and adaptive observations. *J. Atmos. Sci.* 55 (4), 633–653. doi:[10.1175/1520-0469\(1998\)055<0633:SVMAAO>2.0.CO;2](https://doi.org/10.1175/1520-0469(1998)055<0633:SVMAAO>2.0.CO;2).
- Sakai, K., Kawamura, R., 2009. Remote response of the East Asian winter monsoon to tropical forcing related to El Niño-Southern Oscillation. *J. Geophys. Res.* 114 (D06105). doi:[10.1029/2008JD010824](https://doi.org/10.1029/2008JD010824).
- Schneider, T., Griffies, S.M., 1999. A conceptual framework for predictability studies. *J. Clim.* 12 (10), 3133–3155. doi:[10.1175/1520-0442\(1999\)0122.0.CO;2](https://doi.org/10.1175/1520-0442(1999)0122.0.CO;2).
- Toth, Z., Kalnay, E., 1997. Ensemble forecasting at NCEP and the breeding method. *Mon. Weather Rev.* 125 (12), 3297–3319. doi:[10.1175/1520-0493\(1997\)125<3297:EFANAT>2.0.CO;2](https://doi.org/10.1175/1520-0493(1997)125<3297:EFANAT>2.0.CO;2).
- Van Leeuwen, P.J., Kunsch, H.R., Nerger, L., Potthast, R., Reich, S., 2019. Particle filters for high-dimensional geoscience applications: a review. *Q. J. R. Meteorol. Soc.* 145 (723), 2335–2365. doi:[10.1002/qj.3551](https://doi.org/10.1002/qj.3551).
- Vetra-Carvalho, S., Van Leeuwen, P.J., Nerger, L., Barth, A., Altaf, M.U., Brasseur, P., Kirchgessner, P., Beckers, J.M., 2018. State-of-the-art stochastic data assimilation methods for high-dimensional non-Gaussian problems. *Tellus A* 70 (1), 1–43. doi:[10.1080/16000870.2018.1445364](https://doi.org/10.1080/16000870.2018.1445364).
- Zhang, J., Duan, W.S., Zhi, X.F., 2014. Using CMIP5 model outputs to investigate the initial errors that cause the “spring predictability barrier” for El Niño events. *Sci. China Earth Sci.* 55 (8), 685–696. doi:[10.1007/s11430-014-4994-1](https://doi.org/10.1007/s11430-014-4994-1).