Title
Comment on “Nonparametric forecasting of low-dimensional dynamical systems”

Journal
Physical Review E, 93(3)

ISSN
2470-0045

Authors
Kondrashov, Dmitri
Chekroun, Mickaël D
Ghil, Michael

Publication Date
2016-03-01

DOI
10.1103/physreve.93.036201

Peer reviewed

Comment on “Nonparametric forecasting of low-dimensional dynamical systems”

Dmitri Kondrashov, Mickaël D. Chekroun, and Michael Ghil

Phys. Rev. E **93**, 036201 — Published 16 March 2016

DOI: 10.1103/PhysRevE.93.036201

Comment on "Nonparametric forecasting of low-dimensional dynamical systems"

Dmitri Kondrashov,* Mickaël D. Chekroun,† and Michael Ghil‡

Department of Atmospheric and Oceanic Sciences, 405 Hilgard Ave, Box 951565, 7127 Math Sciences Bldg. University of California, Los Angeles, CA 90095-1565, U.S.A.

(Dated: December 10, 2015)

The comparison performed in [1] between the skill in predicting the El Niño-Southern Oscillation (ENSO) climate phenomenon by the prediction method of [1] and the "past-noise" forecasting (PNF) method of [2] is flawed. Three specific misunderstandings in [1] are pointed out and set straight.

I.

To test their proposed forecasting method on climate data, the authors of [1] considered the so-called Niño-3.4 index that represents anomalies of sea surface temperatures (SSTs) in the central equatorial Pacific; in the climate sciences, an anomaly is simply the difference between the instantaneous or monthly mean value and the climatological average. The Niño-3.4 scalar index is widely accepted as a good indicator of the ENSO phase and amplitude [3, 4]. The forecasting model of [1] was trained on monthly data for the January 1950 – December 1999 interval and its forecasting skill was verified over the January 2000 – September 2013 interval.

The difference between forecast and verification time series is measured in [1] and [2] by root-mean-squared error (RMSE) and correlation coefficient (CORR). The results of [1] are plotted in their Fig. 3 and the authors claim that their 14-month lead forecast skill of RMSE = 0.60 and CORR = 0.64 is significantly better than that of the PNF method proposed in [2] and reported in Fig. 3 of that paper, with RMSE = 1.4 and CORR = 0.4, respectively. We point out below three crucial misinterpretations of our own results from [2] by the authors of [1], along with associated methodological issues in comparing the results of the two methods.

(i) Comparison with an "incorrect PNF" and not with the genuine PNF method of [2]. The comparison of the results of Fig. 3 in [2] with Fig. 3 in [1] used the wrong curves in [2, Fig. 3].

We hope that the clarifications presented here will help the authors of [1] and the general readership to better understand Fig. 3 in [2]. As explained in [2], the reshuffled PNF test (green curves in panels B and C of [2, Fig. 3]) was used to emphasize the importance of performing the appropriate selection of the noise snippets from the past, in order to drive the forecasts into the future.

The green curves in these panels illustrate the results when the snippet selection is made according to a random reshuffling vs. the blue curves that provide the correct PNF results, with the selection performed according to the procedure at the heart of the PNF method, in which the noise snippets are selected according to the proper phase of the low-frequency variability (LFV) present in the Niño-3.4 time series; see *Step S1* on p. 11 769 of [2], right column.

The comparison of the blue curves in [2, Figs. 3B, C] (correct snippet selection) with the green ones (random selection) shows the former to be a great improvement upon the latter: for the blue curves, the correct-PNF forecast skill scores at a 14-month lead are RMSE = 0.99 and CORR = 0.4. Having clarified this key point, we give the reasons for the still apparently better RMSE score of the method proposed by [1] in points (ii) and (iii) below.

(ii) **Normalization of the RMSE values.** The RMSE skill scores in [2] are properly normalized, while those in [1] are not.

The correct way to normalize the prediction error in RMSE is by the climatological value of the fluctuations about the mean; this normalization corresponds to forecasting simply the climatological mean. Hence the zero-anomaly forecast (dashed blue horizontal line in the upper panel of [1, Fig. 3]), labeled "Climatological Error," should correspond to the value 1.0 and not $\simeq 0.78$, as reported there. With the proper normalization, the RMSE at 14-month lead in [1] increases to $0.60/0.78 \simeq 0.77$. On the other hand, the reported RMSE forecast skills in Figs. 3 and 4 of [2] have been properly normalized, as explicitly mentioned in their respective captions.

(iii) Comparison of two forecasting methods on different benchmarks: the Niño-3.4 index in [1] with the Niño-3 index in [2].

The comparison in [1] uses the prediction results in Fig. 3 there and those in [2, Fig. 3]. The latter, though, shows forecasting skill for a different time series, namely for the Niño-3 index. Both indices have been used in operational ENSO prediction [5, 6]. But the Niño-3.4 region (5S–5N, 170–120W) is located to the west of the Niño-3 region (5S–5N, 150W–90W), and these two regions exhibit different ENSO dynamics. Niño-3, on which [2] focus, is usually associated with classical ENSO events that occur mainly in the Eastern Tropical Pacific, while Niño-3.4 is more representative of Central Pacific ENSO events [7]. Hence the prediction skill for the two is not expected to be the same.

In fact, in [2], predictions were carried out for the entire equatorial SST field, and not just for one of these two scalar indices, as in [1]; doing so confirmed, in particular, that predictive skill varies spatially and that the Niño-

* dkondras@atmos.ucla.edu
† mchekroun@atmos.ucla.edu
‡ ghil@atmos.ucla.edu

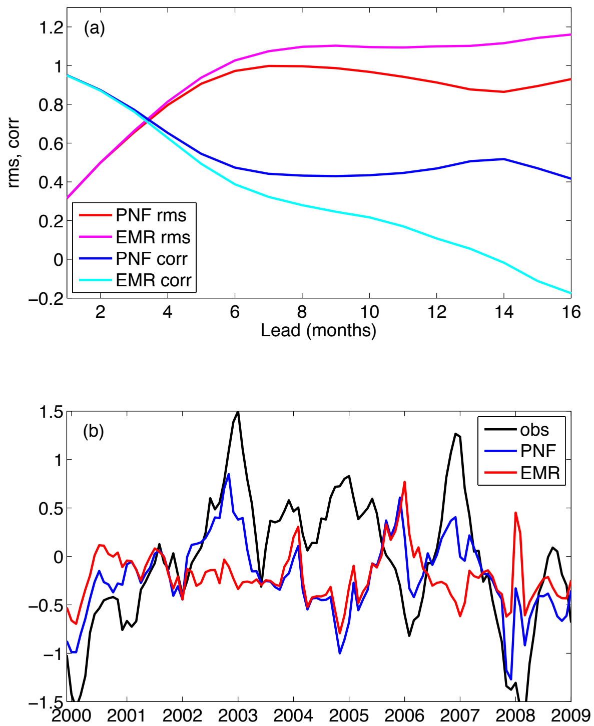

FIG. 1. (Color) Niño-3.4 prediction skill for the 2000–2009 interval, based on the results of [2]. (a): normalized RMSE and CORR ensemble mean forecast skill obtained by the empirical model reduction (EMR) method of [8] and by the PNF method of [2]; (b): PNF forecasts (blue curve) and EMR forecasts (red) validated at 14-mo lead, compared to the actual data (black).

3.4 region has, in fact, greater predictability, i.e., lower RMSE and higher CORR than the Niño-3 region; see panels B and D of Fig. 4 in [2], respectively.

For the sake of clarity, we plot here in Fig. 1 the corresponding PNF results for the Niño-3.4 index in the same way as for Niño-3 in Fig. 3 of [2]. These new plots show PNF skill at 14-month lead of RMSE = 0.86 and CORR = 0.52, respectively.

To summarize, for the Niño-3.4 index used by [1], their now properly normalized skill scores at 14-mo lead of RMSE = 0.77 and CORR = 0.64 are, at best, only modestly better than the RMSE = 0.86 and CORR = 0.52 values of PNF results reported here. Recall also that — given the shortness of the validation interval allowed by the availability of accurate instrumental data — the skill scores reported for the two methods are subject to a sampling error that might exceed the reported differences.

Furthermore, comparing the lower panel of Fig. 1 here with the corresponding panel of Fig. 3 in [1] shows that the extreme episodes — i.e. the strong El Niños in 2003 and 2007, as well as the strong La Niña in 2008 — are better predicted by the PNF method, while the modestly better overall skill of method [1] results from the much more extensive quiet episodes. Moreover, as [2] explains, the PNF improvement for these extreme episodes is due to the constructive interference between energetic phases of the quasi-quadriennial (QQ) and quasi-biennial (QB) modes of LFV [5]. The PNF improvement is not guaranteed during episodes with weak LFV, such as the borderline El Niño of 2005, cf. the cyan line for QQ+QB in Fig. 3a of [2]). It is the prediction of the strongest ENSO episodes, though, that is climatologically and socio-economically most interesting [4, 6].

PNF has also been successfully applied to forecasting the Madden-Julian Oscillation (MJO), where the shorter MJO time scale, along with the relatively longer dataset, allows for robust identification of strong LFV episodes. Once more, the latter episodes coincide with pronounced PNF improvement, cf. the cyan and blue lines in Fig. 1d of [9].

Finally, it is worth mentioning that the PNF method of [2] aims to improve long-term ENSO prediction beyond one year, while at shorter leads — i.e., shorter than 6 months — its skill is about the same as that of the currently operational empirical model reduction (EMR) method introduced in [8]. This roughly equal skill is apparent at up to 6 months lead time when comparing the red and magenta lines for RMSE, and blue and cyan lines for CORR in the upper panel of Fig. 1 here.

EMR has been identified by an independent study [6] as one of the best among half-a-dozen statistical and a dozen dynamical models for Niño-3.4 prediction being continuously monitored at the International Research Institute for climate and society (IRI). The most recent EMR forecast contributed to IRI's ENSO "forecast plume" by the UCLA team, and based on Tropical Pacific SST conditions through July 2015, calls for an exceptionally strong El Niño during the 2015-2016 winter; see the UCLA-TCD statistical model at `http://iri.columbia.edu/climate/ENSO/currentinfo/`.

The research herein was supported by grants N00014-12-1-0911 and N00014-16-1-2073 from the Multidisciplinary University Research Initiative (MURI) of the Office of Naval Research and by the US National Science Foundation grant OCE-1243175.

---

[1] T. Berry, D. Giannakis, and J. Harlim, Phys. Rev. E, **91**, 032915 (2015).
[2] M. D. Chekroun, D. Kondrashov, and M. Ghil, Proc. Natl. Acad. Sci. USA, **108**, 11766 (2011).
[3] M. A. Cane, S. E. Zebiak, and S. C. Dolan, Nature, 827.
[4] S. G. H. Philander, *El Niño, La Niña, and the Southern*

Oscillation (Academic Press, San Diego, 1990).
[5] M. Ghil and N. Jiang, Geophys. Res. Lett., **25**, 171 (1998).
[6] A. G. Barnston, M. K. Tippett, M. L. Heureux, S. Li, and D. G. DeWitt, Bull. Amer. Meteorol. Soc., **93**, 631 (2012).
[7] K. Ashok and T. Yamagata, Nature, **461**, 481 (2009).
[8] D. Kondrashov, S. Kravtsov, A. W. Robertson, and M. Ghil, J. Climate, **18**, 4425 (2005).
[9] D. Kondrashov, M. D. Chekroun, A. W. Robertson, and M. Ghil, Geophys. Res. Lett., **40**, 5305 (2013).