Nonlin. Processes Geophys. Discuss., 1, C102–C111, 2014
www.nonlin-processes-geophys-discuss.net/1/C102/2014/

Nonlinear Processes in Geophysics
Discussions

Interactive comment on "Implications of model error for numerical climate prediction" *by* O. Martínez-Alvarado

O. Martinez-Alvarado

O.MartinezAlvarado@reading.ac.uk

Received and published: 6 May 2014

To Anonymous Referee #2

Thank you very much for reviewing this contribution. In the following I give a response to your comments indicating whether changes to a final version of the paper would be suitable or not and why the article should be published (Your comments are in *italics*).

The paper investigates if solutions of numerical climate models lie on the projection of the actual climate attractor in the model's phase space.

This is incorrect. The article investigates what the consequences of having a model whose attractor does not lie on the system's attractor by construction are.

The study compares model errors between perfect and imperfect models when using perfect or imperfect initial conditions for Lorenz '63 and snapshots of atmosphere models and observations.

This is incorrect. There is only model error in the imperfect model, not in the perfect model. Therefore model errors between these two kinds of models could not be compared in this study.

The results are correct. However, I feel the conclusions are already well known and I do not see the scientific result that goes beyond existing work of studies of model biases in data assimilation.

One important point to clarify is that this is not a study into data assimilation per se. In fact the study's focus is long-term simulations where data assimilation is not relevant any more because the information in the initial conditions has been lost.

The use of a toy model (Lorenz '63) is a perfectly valid approach, but the possibility to perform a decent statistical analysis with long model simulations is not exploited.

Hopefully the Clarification Note added to the interactive discussion on 29 April 2014 has cleared this point out: the methodology description in the article is incorrect: the long-term simulations were 5000 t.u. long instead of 80 t.u. as originally stated. I can add that enhanced statistics is not the only advantage of using simple models. In this study, for example, the fact that it is possible to know perfectly a system and a model that tries to replicate the system's behaviour is what is mostly exploited.

To my understanding there is scientific consensus that:

- *Weather and climate models show systematic biases. This implies a different attractor for the model and the real world.*
- *If data assimilation is used, the model is initialized close to the state of the climate system but it will fall back into it's model state after a short period of time. While*

a perfect model would not show biases in climate runs (at all forecast times), state-of-the-art models do.

- *The correlation of the model state with initial conditions will reduce with time and become tiny for long term simulations (if the mean state is subtracted).*

The results of this paper state (see abstract):

- *"It is shown that the imperfect model's orbit and the system's orbit are essentially different, purely due to model error and not to sensitivity to initial conditions."*
- *"Furthermore, if a model is a perfect model, then the attractor, reconstructed by sampling a collection of initialized model orbits (forecast orbits), will be invariant to forecast lead time."*

I do not see much of a difference.

The understanding of Anonymous Referee #2 might not be correct when he/she states that a perfect model would not show biases in climate runs. I am genuinely interested in knowing how he/she knows that; would his/her statement be applicable to the actual climate system and to our current means of observing it? Even if Anonymous Referee #2's understanding were completely correct, this would not mean that we know everything about the relationship between errors in models and their manifestation as model biases, especially in long-term simulations. There is certainly a vast amount of research into the relationship between having model error and the quality of short term forecasts, but there is no published research about the relationship between errors in models and their long-term predictions. This article starts exploring this field with one of the simplest systems that show chaotic behaviour and sets the framework to carry out further research in this direction with more complex systems/models. As stated before more than enhanced statistics, the value of using a simple model in this article

comes from the fact that system and model are known perfectly. The use of a simple model enables the researcher to make direct comparisons between the behaviour of a perfect model (initialised with imperfect initial conditions) and an imperfect model (initialised with perfect initial conditions). Furthermore, it allows investigating the more realistic scenario in which the system (or the perfect model) is not known.

The one result which I found to be really interesting/surprising is given on page 138: The perfect model with imperfect initial conditions undergoes a short period in which the distance to the full system is decreasing. However, I can not recommend the publication of the paper in it's current form.

I also found that result very interesting and surprising as well as the rest of the results. All these results come from the perfect knowledge of the system and model, which is the feature the article exploits. These results show the kind of relationships that we need to understand and for which the use of simple but perfectly determined systems is useful. This is a reason for publishing the paper in its current form although I agree that it can be improved as much as anything else.

Major comments:

- *In the introduction it is not stated what is actually investigated within the paper. Just saying: "The objective of this contribution is to show the implications of the second assumption for long-term integrations of a "simple" dynamical system in three-dimensional phase space" is not enough to give the reader any idea about the content of the paper.*

Thank you for pointing this out. This can be made more explicit, for example, by including parts of the previous discussion and part of the following.

- *The use of a toy model such as Lorenz '63 should allow the use of sufficient statistics.*

Hopefully this point has been cleared out with the Clarification Note added to the interactive discussion on the 29 April 2014. In that note I gave a correct description of the methodology so that it is clear now that the long-term integrations comprise 5000 t.u. rather than 80 t.u. Therefore the number of forecast cycles is 1000 rather than 16. In addition, the use of a simple model is not only useful for the production of sufficient statistics. In this case it is useful because it allows a perfect knowledge of both system and model.

However, the statistics for Figure 1, 2 and 3 are not sufficient. The quantity that should be considered in Figure 1 is the PDF of a simulation much longer than the presented simulation for the two attractors.

It is not clear what PDF Anonymous Referee #2 would like to see and I don't see any strong reason why my figure should be a PDF. A PDF (even a 2D PDF) would make it difficult to visualize the differences between attractors, a problem that is not present if you simply use the plots of the two attractors in phase space. This is exactly what the figure in its current form shows. Furthermore, part of the point of the article is to point out that looking at partial statistics while evaluating a climate model can lead to wrong conclusions in terms of assessing its behaviour. I would be falling into this mistake from the start if I followed Anonymous Referee #2's suggestions.

I am sure that the results in Fig. 2(b) will change if the attractor is calculated for a longer time since more trajectories run away from the standard loop. The plot is therefore not very meaningful. Fig. 2(b) should be changed into a comparison to the PDF.

First of all, trajectories in the Lorenz '63 system do not run away from the standard loop or at least not as wildly as this comment would imply. That is the whole point about the existence of a strange attractor. Figure 2(b) is showing the distance of a few initialisation points to the imperfect model's attractor. The exact distance will depend on the length of the attractor segment used to compute this distance, but the conclusions will remain the same as long as system's and model's attractors are not the same. I

don't see any reason to change it. The analysis that Anonymous Referee #2 suggest will also depend on the length of the portion of the attractor. Furthermore, it will also depend on the frequency at which the forecasts are initialised. And still the conclusions will be the same. Therefore once again I don't see any strong reason why my figure should be any sort of PDF.

Showing four instances of orbits from initial conditions in Figure 2(a) and (c) is good for visualization, but not sufficient to draw any meaningful conclusions (e.g. consider mean and variance of the distance against time for a large number of starting points).

The figures included in the article have been carefully designed to illustrate the points made in the article. I therefore thank you for your comment that Fig. 2(a) and (c) are good for visualisation because that is exactly their purpose. The analysis that you suggest is basically the analysis that is carried out in the rest of the article. Therefore I agree that it is a very good idea but it has been done in this article.

More statistics is needed!

I don't agree with this.

It should easy to turn the zigzag lines in Figure (3) into smooth lines if longer simulations are considered. And this will be totally necessary if conclusions are drawn from these plots. For example: With the given zigzag plots I do not believe the following sentence: p. 138 l.10: "The two imperfect models show a short period of very fast divergence from the prototype system's orbit followed by a plateau and a second period of fast divergence." I believe the plateau is caused by insufficient statistics.

Figure 3 has been constructed with 1000 points for each lead time. I don't think the plateau is caused by insufficient statistics (please also see the Clarification Note added to the interactive discussion on 29 April 2014).

- *Figure 4: I have seen similar plots in talks about data assimilation in atmosphere and ocean for several quantities and several time-scales (without the horizontal* *lines). It is well known that models fall back into their own attractor if data assimilation is stopped. I do not see why this result and the conclusions (p. 139 ll26-29) are new!*

Unfortunately I have been unable to identify the talks about data assimilation that Anonymous Reviewer #2 has attended (N.B. this study is not about data assimilation). Therefore, I am not at all sure what the plots he/she is referring to are. I am also not sure in what way my plots are similar to those Anonymous Referee #2 has seen. However, the lack of the horizontal lines would constitute a critical difference because they allow a comparison between the PDFs produced by the system and the imperfect model. They also allow understanding the difference between these PDFs and those produced by an imperfect model initialised with perfect initial conditions and a perfect model initialised with imperfect initial conditions. Therefore, Figure 4 including horizontal lines should remain.

- *p. 140 ll.1-8: As outlined in the introduction, we know that short term forecasts can be used to evaluate models for long term predictions. We also know that models will fall back into the model state when data-assimilation is stopped. What is new?*

These two facts are known and yet there is no published research explicitly linking these two aspects. There are studies on very short-term forecasts showing errors that look similar to climate model biases and studies linking not-so-short-term hindcasts with climate model biases, but not much more than that. Furthermore, those studies have not really been able to correct biases in models, indicating that there is much more to say about these problems. Moreover, the ideas described in this article can be more useful than simply trying the impossible task of constructing a perfect model. They can be useful for example to develop a better way of interpreting biases for example for increasingly important regional climate change studies. What is new in this

article is a formal investigation of how model errors give rise to model biases through a comparison with a perfect model that is simply not available for realistic climate studies. What is new is the realisation that this method works for a very simple model and the opening of questions such as its applicability for more complex systems like the atmosphere, the atmosphere-ocean or the Earth. At this point I have to say that I am genuinely interested in knowing why Anonymous Referee #2 stated that it is know that "... a perfect model would not show biases in climate runs (at all forecast times)". Under what conditions would this be true?

- *p. 143 ll. 1-12: "The fact that no member in the ensemble is close to the actual behaviour of the system might be due to the same effect: in this particular event, an ensemble around accurate initial conditions generates an ensemble forecast with every member tending towards the model's attractor and away from the true future state of the system". This might be true. But it might also be true that the spread in the ensemble is simply not tuned correctly and fails to reproduce the model error. I do not believe this conclusion on the basis of only one forecast. This is a chaotic system: One occasion does not tell much.*

I would not go as far as saying that the spread in the ensemble is not tuned correctly given that it is the operational ensemble prediction system of one of the world's leaders in weather prediction. From this point of view, the results presented come from one of the best models available; therefore we can be certain is been tuned as correctly as possible. However, I agree that the ensemble settings could be improved. I can acknowledge this. To better understand Anonymous Referee 2's point I would need clarification on what is meant by 'fails to reproduce the model error'. Surely the model error is there and does not need to be reproduced. I totally agree that one occasion in a chaotic system does not tell much. I'm certainly not saying otherwise. However, we should not indulge ourselves in the belief that this happened only once and the rest of the time the model might be fine. We very well know that that is not true. What we should really do is ask ourselves many other questions: Why did it happen? Is it a common feature? Does it have any consequence for long term statistics such as those required for climate studies? Can we learn something about the problem of climate model biases by looking at weather forecasts?

- *p. 144 l. 9-12: Please state clearly what this framework is. I believe you mean the conclusions in p. 140 ll.1-8???*

Yes, that is correct. I can state it more clearly in a revised version if this is allowed.

Minor comments:

- *Acronyms are used throughout the paper that can be avoided. E.g.: p. 142 l. 4: no explanation for 2-PVU; Fig 2: IC can only be understood in combination with the text; Fig. 4: l.u. is hard to understand when looking at the Figure only; p. 152: PV*

Correction accepted.

- *p. 132 l. 24: I guess it would be sufficient if the model attractor is very close to the system's attractor.*

The whole point of the article is that it is not sufficient that the system's attractor is very close to the model's attractor. However, your statement might become true if a proper definition of state on the attractors is made. However, this is beyond the scope of the article.

- *p. 134 l. 17 and p. 143 l. 14 and p. 144 l.9: I find the introduction of the word "prototype system" rather confusing. Especially in the conclusion: "...in the prototype system/imperfect model combination based..." is not at all helpful.*

The paragraph at P134L17 is intended to explain why the term prototype system is used. This paragraph can be reworded and perhaps the reader can be reminded later what the term means.

p. 136 l. 16: It is Figure 1.

Correction accepted.

p. 137 l. 15: Lorenz '93 -> Lorenz '63

Correction accepted.

- *p. 140 l. 24: (T + 0d) -> T + 0d*

Correction accepted.

Figure 1: Please increase the size of the black dots, choose a different color than the nearly invisible gray and thicker lines, and use the same color in (a) and (c). The explanation of (b) is cryptic.

This correction might be accepted. All the changes can be done but I am not convinced they are actually necessary.

- *Figure 2: I would suggest to remove the lines for the interquartile range. The label of the y-axis is not useful and should be replace*

I assume Anonymous Referee #2 refers to Fig. 4. The inclusion of the interquartile range is useful (I would say necessary) to get an idea of the dependence of the PDFs with lead time and therefore these lines should remain.

Interactive comment on Nonlin. Processes Geophys. Discuss., 1, 131, 2014.