Geoscientific
Model Development
Discussions

# *Interactive comment on* "AnaWEGE: a weather generator based on analogues of atmospheric circulation" *by* P. Yiou

**H. van den dool (Referee)**

Huug.Vandendool@noaa.gov

Review of AnaWEGE by Yiou (2013). Review by Huug van den Dool (NCEP) The paper describes an approach to obtain horizontal coherence in data produced by a weather generator. The approach, using analogs to SLP fields, is in itself totally sound. But I don't know whether it is new. It may certainly be useful.

Let me state upfront that this reviewer knows something about analogues, but not much about weather generators. So I can't judge whether weather generators were hitherto pointwise, thus making the analogue aspect a novelty. In the context of downscaling, the analogue methods have been used with spatial coherence for a long time (for example Van den Dool 1994), and the idea (of spatial coherence) is much older than 1994. But perhaps this IS new in the context of weather generators.

C2228

The paper is a bit of a mix of a research paper on the one hand and a manual to describe a new piece of software on the other.

As a research paper it falls a bit short. Example 1. There is no discussion of the size of the domain for SLP vis-à-vis the area of application. By what criterion does one choose the domain(s)?. What about the impact of a smaller or larger domain on the results? Is the present domain optimal? What is the smallest domain size acceptable? (Compare Hammill and Whitaker 2006). Example 2. Although extremes are investigated (last sentence abstract) I don't see any solid conclusions about extremes and don't trust what I read too much. Analogue methods do have several weaknesses about extremes (none mentioned). First, it is likely, in general, that analogues are closer to the mean (more to chose from in that direction) than the base case (j), and since the wings are always under-sampled, the same cases keep coming up as analogues to extremes at time j. You may want to have a look at the frequency distribution of the 20 cases picked in say January (for all instances when the base case is deemed extreme) – some of the same analogue cases are picked over and over again, others are never used. What does this imply? Example 3: Other methods, CCA or similar can be invoked if spatial coherence is sought. Maybe analogues are better, but where is this discussed.

As a technical description I wonder whether we need to know such details as Equations 1-3 (which moreover are flawed in an attempt to be formal). No need to redo notions like Julian Day. Just saying in words what you do is enough and less confusing. On the other hand too much inside terminology is used as if readers should know. Examples, calculation in R (not sure what this means), and the use of words like copula (you have to be a real insider for that notion).

Some detailed specific and mainly minor comments (or repeating some of the above) are as follows:

p4747 mentions two rules. Maybe rules is not the right word. I don't see any rules in the next several sentences.

p4748 80W-30E, 30N-70N. Why this domain?

Eq.3. How can abs(j-j') be <=30 if y and y' are not equal; they will differ 10**4 or something like that. I know what you mean, but it is not stated correctly. This 'window' is 30 days to either side of Gregorian calendar date j? i.e. a window of 60 days?

p4749 ln 14 "The 20 first analogues are computed". You compute them all, thousands of them, about 60 days times 60 years, then select the best 20.

p4749 lower correlation in summer is caused, in reviewer's opinion by lower signal to noise ratio. Less contrasted spatial patterns sounds obscure. Compare to Van den Dool's book (2007), ch 7a, and ch 6 (very short). What season is Fig.2 for? Or are you mixing all seasons here? Not a good idea.

p4750 ln 10 says that correlations and RMS are not correlated in general. Maybe because you mix seasons. They are strongly related for a given season. See also the Murphy and Epstein decomposition for the exact relationship.

p4751 ln 11 The probability of not changing j sub s. You lost me.

Fig.3 etc are any of these discrepancies "significant", or something to worry about.

p4759 ln 20 probability of a cold winter is small. OK. Too small???

p4761 ln 18 What is R website.

H. M. van den Dool, 1994: Searching for analogues, how long must one wait? Tellus, 46A, 314-324. Hamill, Thomas M., Jeffrey S. Whitaker, 2006: Probabilistic Quantitative Precipitation Forecasts Based on Reforecast Analogs: Theory and Application. Mon. Wea. Rev., 134, 3209–3229. See chapter 7a in Huug van den Dool, 2007: Empirical Methods in Short-Term Climate Prediction. Oxford University Press: 215 pages. ISBN-10: 0-19-920278-8 ISBN-13: 978-0-19-920278-2