

Interactive comment on “The construction of a Central Netherlands temperature” by G. van der Schrier et al.

G. van der Schrier et al.

schrier@knmi.nl

Received and published: 28 January 2011

Reply to reviewer 1

We would like to thank the reviewer for the time spent on reviewing this paper and for the useful suggestions provided.

The referee finds the construction of a Central Netherlands Temperature an interesting topic and relevant for this journal and the study clear, interesting and comments on the carefulness with which the metadata is handled.

The main concerns of the reviewer concern the application of the homogeneity checks and the construction of the reference series. Concerning the homogeneity checks;

C1407

urged by the comments of the referee we evaluated the weight we gave to detected breaks in the records which had only weak supporting evidence of the metadata. Our feeling now is that we may have been too conservative. We may have failed to correct for breaks which did warrant an adjustment.

detailed reply:

1. Introduction

As stated in the introduction, this reconstructed series is very close to the one previously computed in Van Ulden et al. in 2009. So an obvious question is : what does this series improve?

The series presented in this study is indeed very similar to the one introduced by Van Ulden et al. (2009). The fact that it is so similar, despite completely different methods for the homogenization and for the construction of the reference series adds to our claim that the presented CNT series is very robust. In the Introduction and Conclusion sections we will make our claim concerning the robustness of the results more clear. A small, but not unimportant point, is that the current study is based on additional research in the KNMI archives, which makes that some observed breaks can now be dated more precisely than was hitherto the case.

2. Construction of long records *The Van Der Hoeven (1992) method to reconstruct T24 is described (not very well: subscripts and superscripts have disappeared in 2), but... it is unclear whether it has been used or not! Note that this is a very empirical method. One may wonder why a direct regression model has not been used, although there might be colinearity problems here.*

The Van der Hoeven (1992) technique has indeed been used. We will make additional comments to make this more clear. A little more background why Van der Hoeven chose this particular method will be provided.

3. Method

“Both references and target series have their seasonal cycles removed”. What

C1408

happens when changes affect seasonality? Especially when the main criterion for assessing changes is a Durbin-Watson test. In Menne & Williams, detection is performed on serial monthly series. From what I understood in your examples (5) and (7) is empirical ("visual confirmation"). But those are nested models, so choosing which one is best can be performed straightforwardly by means of a Fisher test.

The referee has a point here. Removal of the seasonal cycle is indeed superfluous. The methodology and text will be adapted.

Regarding the choice of models: the choice is indeed empirical and based on visual confirmation. It is possible to formalize the choice between models with e.g. a Fisher test as the referee suggests, but we refrain from doing this. In an earlier attempt we formalized the choice between models using a statistical test. We noticed that model (7) (i.e. a combination of a discontinuous trend and step) was chosen in more cases than what could be confirmed by the available metadata. This observation made us change the procedure and allowed for an adjustment of both trend and step when the metadata provided any clues that these adjustments were required. The text will be modified to make this reasoning more clear. Note that distinguishing between models (3) (a step) and (5) (a step in the presence of a continuous trend) is not really necessary since we do not adjust for the continuous trend in model (5), as explained in the manuscript.

4. Reference

- PCA is not a "new" technique for building reference series, references should be searched. - (10) (11) and (12) do not take into account the fact that the "ci" were established using the target series as well. Repeating PCA just excluding the target series should not be so difficult and computationally expensive!

We were not aware that using PCA in building a reference is not new - many thanks for alerting us to this.

The referee is right in noting that the coefficients in eqs. (10) and (11) are

C1409

based on a calculation that included the target series as well. The adjustment explained in eqs. (9)-(11) indeed ignores this effect, although we think that this is of secondary importance since all PC's explain about the same amount of variance.

The argument of computational costs in not calculating reference series which exclude the target series is indeed rather weak. We will remove this argument from the ms. In order to quantify the difference between an adjusted reference series, as done in the ms. and explained in eqs. (9-11), and a reference series excluding the target series, we will make a calculation using De Bilt as target series. Based on a comparison of the reference series we will judge if changing the procedure is apt or not.

Note that the high variance explained by the first mode is not related to seasonality - this has been removed from the series prior to the PC calculation. This mode reflects the warming trend (as stated in the ms.) and due to the relative smallness of the Netherlands, this trend is very homogeneous. Additionally, the smallness of the country makes that changes in (monthly averaged) circulation are very homogeneous over the country as well.

5. Quality checks

Quality checks Standard deviation is interesting, but exhibiting the difference series candidate minus ref is more revealing.

Showing the difference series is indeed interesting, although a running standard deviation reveals much of the interesting details. In the revised ms. we will add a figure showing the (low-pass filtered) difference series.

6. Detection breaks and trends

This paragraph is a long list of "we detected a significant artifact, but did not

C1410

correct it since we did not find supporting metadata". If you assume that your metadata are perfectly complete, no need to use statis you use tests at a 5% level, that is very conservative tests, the latter should occur quite often too. Caricatural examples:

The referee characterizes the approach of the study by "we detected a significant break but did not correct for it since we did not find supporting metadata". The observation of the referee is only partly correct; many corrections are made while no metadata was available to support this action. The incompleteness of the metadata is something of which we are very aware. One of the problems we ran into when collecting metadata is that there are many details like routine changes of thermometers, thermometers which needed to be re-calibrated etc. These routine changes are made about once a year, which makes it unlikely that a step change or a discontinuous trend is related to a faulty thermometer. However, it shows that the detail of reporting for these stations is high, when even relatively minor maintenance is added to the metadata records. Consistent and detailed reporting for a station is difficult to combine with vast changes in the surroundings (leading to the break or discontinuous trend) which remained unnoticed. This reasoning misses from the description in §6.3 and makes this passage particularly unclear and the reasoning behind the ignorance of the detected break difficult to understand. In §6, we will comments along the lines of the above discussion.

6.4 Oudenbosch

Here the referee has a point. The Van der Hoeven method uses max. and min. temperatures as part of the input. For this particular break (around 1946-47) we will again consider the evidence.

The 1971-72 break will be reconsidered as well.

6.9 Hoorn

C1411

The 1970-73 break will be reconsidered.

6.11 Deelen

Deelen is a military airport which are notorious for their scanty metadata. A break "large enough to warrant adjustment" is a detected break which convincingly exceeds the critical significance levels. This will be made more clear in the new ms.

7. CNT

Why not use anomalies, which would allow compute the series directly, even with varying number of series? The process leading to corrections of the varying number of series is not really described. If your purpose is to make something representative of the "Central Netherlands", then you should use an interpolation method. I do not understand why you extensively describe the homogeneization of some series that you do not include in the end in the CNT.

Using anomalies is a possibility, but we opted for a record which shows actual values. One of the uses of the CNT record is to put monthly averaged temperatures in a historical perspective.

Using an interpolation method is indeed a possibility and corrects for artifacts which result from an inhomogeneous distribution of stations. However, the distribution of the series which are used to construct the CNT is very homogeneous, they are evenly spread over a region in the central-south eastern part of the Netherlands. Given the even distribution of station, interpolation would not add information which is not already in the set of stations.

Here the referee has a point. The paper is somewhat bifocal; it aims to give a set of homogeneous monthly averaged records for daily averaged temperature for the Netherlands and it aims to present the CNT. We will reconsider the scope/aim of the paper.

C1412

8. *Conclusions*

As a conclusion to this review, for your purpose, it is more important to remove significant changes, even if it sometimes leads to unnecessary adjustments, that will be small anyway, than letting artifacts uncorrected. So please redo the job, correcting the artifacts you put into evidence.

Urged by the comments of the referee we evaluated the weight we gave to detected breaks in the records which had only weak supporting evidence of the metadata. Our feeling now is that we may have been too conservative. We may have failed to correct for breaks which did warrant an adjustment. In the new submission, this will be taken into account.

Interactive comment on Clim. Past Discuss., 6, 2517, 2010.