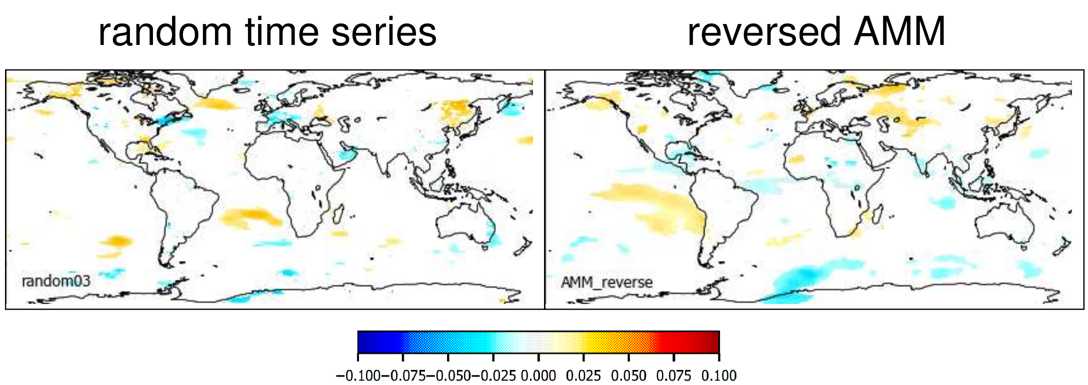
Response to Reviewer 1

We thank the reviewer for the constructive and helpful remarks suggestions. We followed most of these suggestions as described below. Before we give our detailed answers to the individual comments, we first summary the most important changes to the manuscript.
The following major changes were made to the revised version of the manuscript:

A) The motivation and description of the empirical method was improved.
We agree with both reviewers that the description of new method had errors and was complicated. It was also not well motivated.
In the revised version we shifted many technical details of the new method to the appendix (e.g. the description of our normalisation approach as new appendix A1, or the investigation of the effect of time shifts as new appendix A3).
We also added more details about the fit function to section 4, and more details about the calculation of the temporally reversed indices to section 5.1.
For the motivation of our new approach we added the following information to section 5 (see also new Fig. 6):
'... Information about the significance of the fit results can be obtained from the fit function itself. However, in practice, the significance information from the fit has several limitations:
a) The determination of the significance is based on several assumptions about the data sets, e.g. that all data points of the time series have the same uncertainties and follow a normal distribution. However, the errors of the individual data points can be very different, e.g. the effect of clouds on the errors of the satellite TCWV data set can be very different for different seasons and regions. Also, the uncertainties are not only random but contain also systematic contributions. It is difficult (if not impossible) to quantify the uncertainties of the involved time series.
b) The determination of the significance is based on prescribed significance levels. The choice of such a significance level is arbitrary and the obtained significance information depends on this choice.
c) In several tests we fitted artificial time series to the TCWV data set. These tests showed that even for such non-geophysical time series 'significant' fit results can be obtained (see the examples in Fig. 6). On the left side of this figure, fit results for a time series containing only white noise, and on the right side fit results for a temporally reversed teleconnection index are shown (the temporally reversed index is obtained from the original index by mirroring the time axis). The blue and red areas show fit coefficients for both time series, which are classified as significant by the fit.
Based on these findings, we conclude the use of the significance of the detection of an index derived from the fit itself is not straight-forward.
To address these difficulties, we developed and applied an empirical approach to determine threshold values for the delta RMS values to decide whether an index is significantly detected in a global data set. The new procedure is described in the next section. It has the following two main advantages:

-the threshold values are determined empirically. Thus no assumptions on the properties of the time series or the significance levels have to be made.
- the method provides a clear procedure and in particular a metric which can be applied in a consistent way to different data sets and thus allows a quantitative comparison (see section 6).'



New Fig. 6 Global maps of the fit results for an artificial time series containing only white noise (left) and a temporally reversed teleconnection index (AMM, right). The white areas represent fit results, that are classified as non-significant by the fit routine (for a 5% significance level).

B) The Scope and aims were made more clear.

Probably one important misunderstanding was that we gave the impression to the reader that we aim to investigate the influence of teleconnections on the TCWV. This has probably even led to the expectation that we could predict monthly TCWV using teleconnections. This was not our intention. To make this more clear in the revised version of the manuscript, we removed the term 'influence of...', in all parts of the text. And in the introdiction we added the following information: 'Here it should be noted that we do not aim to identify causal relationships or even to predict the TCWV based on teleconnection indices' At the end of the introduction, we added more details and explanations to our research goals. We also restructured the conclusions accordingly and provided respective answers to the research questions formulated in the introduction. The following modified text was added to the conclusions:

'....Based on the obtained results, we could derive the following main conclusions related to the science questions mentioned in the introduction:

a) We developed a new empirical approach to determine whether a teleconnection index is significantly detected in a global data set. This approach avoids problems of existing algorithms for the determination of significance, because no assumptions on the significance level or the measurement uncertainties have to be made. We applied the new method to a global data set of the TCWV derived from satellite observations and found that 40 teleconnection indices could be significantly detected.

b) We applied the same method also to TCWV from the ERA interim data set. Here we used two versions of the model data sets, one including all data, the other only clear sky data. The results for both versions agree in general very well with those for the satellite data set. This confirms both the quality of the satellite and model data sets. It also indicates that the satellite observations can be seen as representative for all day mean values. For some teleconnections, however, also systematic differences, mainly over northern Africa, were obtained. Since these differences are not found for the majority of the teleconnection indices, we conclude that they are very probably not related to systematic errors of the satellite data set, but rather indicate shortcomings of the model over these regions.

c) We also applied our method to a variety of other data sets, which are usually used in teleconnection studies (surface temperature, surface pressure, geopotential heights and meridional winds at different altitudes). For most of these data sets less teleconnection indices were significantly detected than for the TCWV data sets, while for zonal winds, more teleconnection indices (up to >50) were significantly detected. These results indicate that our global TCWV data set is well suited for teleconnection studies. In our view, this is an important aspect, because our data set is exclusively based on measurements. The strongest teleconnection signals were detected for the data sets of tropospheric geopotential heights and surface pressure. This finding is consistent to the fact that most teleconnection studies are based on these quantities. Another interesting finding is that in none of the global data sets, non-teleconnection indices (like the solar variability, the stratospheric AOD or the hurricane frequency) were significantly detected.

d) We investigated the spatial distribution of the teleconnection patterns. In particular we calculated global maps for the cumulative effect of all teleconnection patterns. For that purpose we first orthogonalised the teleconnection indices to avoid the effect of correlation between the indices. Compared to the original set of indices, much less of the orthogonalised indices (20 compared to 42) were significantly detected in the TCWV data set. Our global map of the cumulative effects of all significantly detected orthogonalsed teleconnections showed the strongest teleconnection signals in the global TCWV data set over the Tropics and in polar regions. These spatial patterns point to importance of different driving mechanisms in different regions.'

C) The relationships between different indices and the motivation for the orthogonalisation of the indices was made more clear:

We added new columns in Table 2 (see below). We now show separate columns for indices similar to ENSO, polar atmospheric indices, MJO indices, as well as other oceanic and atmospheric indices.

Indices similar to ENSO (7)	**Other oceanic indices (16)**	**Atmospheric polar indices (8)**	**MJO indices (15)**	**Other atmospheric indices (8)**	**Others indices (7)**

BEST N34 TPI ONI ENSO N4 IND	HAW PDO PMM N1 TNI NTA TNA WHWP IPO CAR AMO DMI AMM STA TSA EA_ersst	SCA AAO EAWR NAO EPNP AO PE WP	MJ1 MJ2 MJN VPM1 VPM2 VPMN RMM1 RMM2 RMMN OOMI1 OOMI2 OOMIN FMO1 FMO2 FMON	PNA SOI NOI EA QBO Q30 Q50 Q70	Solar indices: RI MGII SWO S107 AP HUR (hurricane frequency) SAOD (stratospheric AOD)

In section 3 we added the following explanation:
' Many of these indices (describing the same phenomenon), but also many of the other teleconnection indices are highly correlated. The strength of these correlations is presented in Fig. 3 as a matrix with correlation coefficients between the different indices (after the seasonal cycles were removed). In spite of the correlations amongst the teleconnection indices, we decided as a first step to include them all in our study, because beforehand it is not clear which index might be best suited to represent a teleconnection phenomenon. Using our empirical approach, however, it becomes possible to quantify the significance and strength of the different indices and thus to select the best suited index for a given teleconnection phenomenon. Finally, we apply an orthogonalisation for the most significant indices (see section 7) to minimise the effect of the correlations and to identify the dominant temporal teleconnection patterns in our TCWV data set.'

To better motivate the orthogonalisation, we modified the respective information in section 7 to:
, To account for correlations between the different indices, we thus applied an orthogonalisation approach. For the orthogonalisation (based on the Gram–Schmidt process), all ‚significant' original indices and significant temporal derivatives (see Figure A11) were considered (in total 57 indices). The order of indices used in the iterative orthogonalisation process was from highest to lowest p99 values. The result of the orthogonalisation approach is a set of modified teleconnection indices, which shows zero correlation amongst each other (for the considered time period). Thus this new set of orthogonalised indices can be used to determine the number of independent significant teleconnection patterns in the global water vapor data sets. We applied our new method to the new set of orthogonalised indices to test which of the modified indices have p99 values above the significance threshold.'

D) The logical flow of the paper and the appearance was improved.

As mentioned above several technical parts were shifted to the appendix. The science questions were better motivated in the introduction, and the corresponding answers were added to the conclusions.
Several Figures were shifted/deleted/modified:
-Fig. 3 was shifted to the appendix
-Fig. 8 was shifted to the appendix
-the upper part of Fig. 9 was shifted to the appendix
-Figs A1 and A2 were deleted as suggested
-the quality of Fig. A4 was improved and the number of the sub-figures was largely reduced (by a factor of 3)
-the quality of Fig. A9 was improved

E) We added a new sub-section (6.1) for the comparison of the spatial patterns of the measured and simulated TCWV.
While for most teleconnection indices very good agreement of the spatial patterns is found between the measuremed and simulated TCWV, for some indices also substantial differences are detected. These differences can point to shortcomings in either the satellite or model data sets (or both) and might be helpful for corresponding improvements.
We added a new Fig. 8 (see below) and the following new text:
'For most of the teleconnection indices, very similar spatial patterns are found in the TCWV data sets obtained from satellite or ECMWF data (see Fig. A9). This confirms both the high quality of the satellite measurements and model simulations. However, for some indices, also substantial differences are found (see Fig. 8). The most obvious differences are found over northern Africa. In principle, they could be caused by errors of both the satellite or model data sets. However, since very good agreement over northern Africa is found for most of the indices, we can very probably exclude systematic measurement biases (like e.g. effects from the high surface albedo over the Sahara). Thus we conclude that the observed differences probably indicate deficiencies in the model simulations, possibly related to the sparseness of observational data over northern Africa used in the model. It is interesting to note that the differences are found for both oceanic and atmospheric indices which have rather different frequencies. These comparison results might help to improve the model performance over norther Africa (and to a lesser degree also over other regions).'

New Fig. 8:

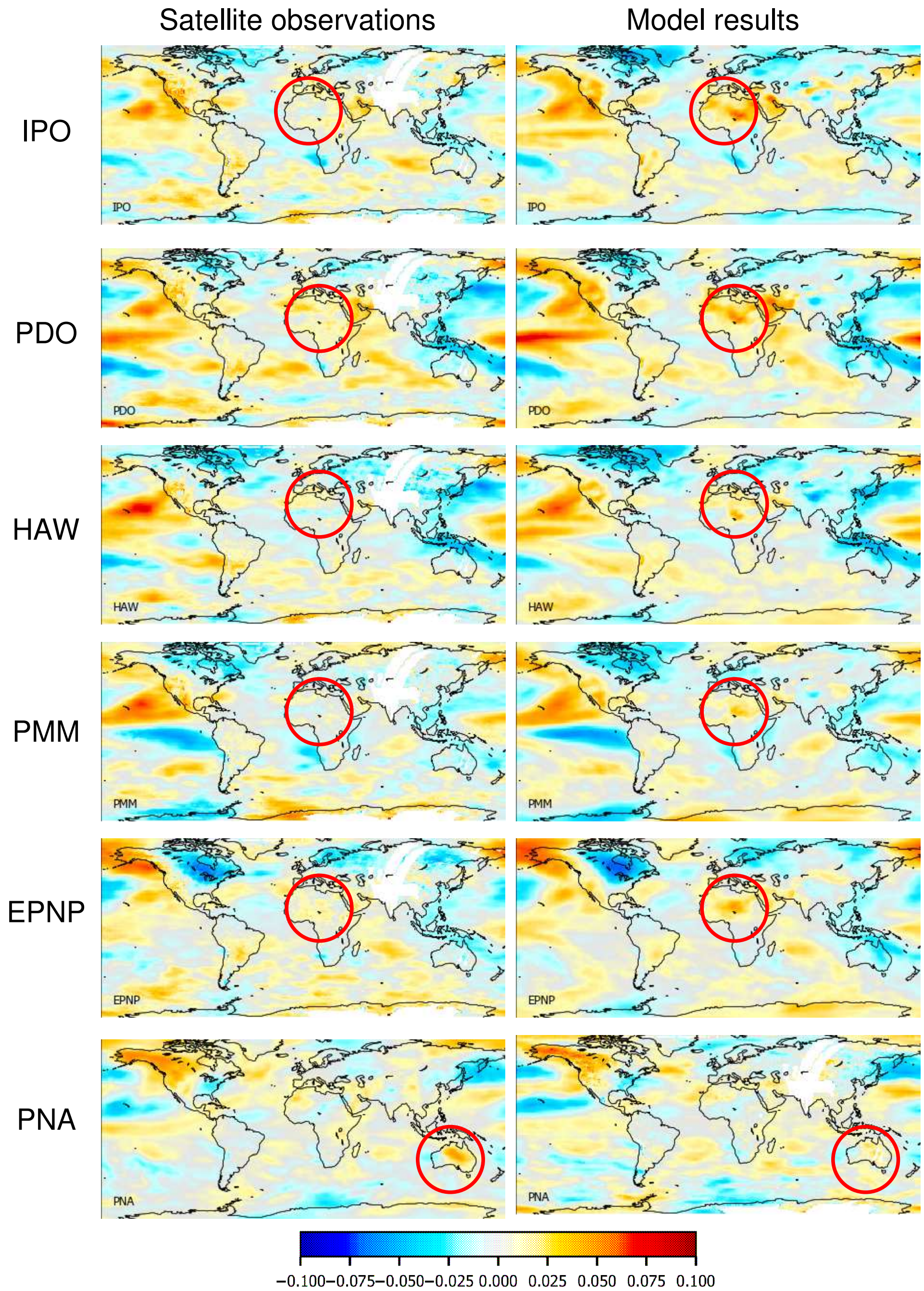


Fig. 8: Fit coefficients for selected teleconnection indices, for which different patterns were found in the TCWV data set from satellite observations (left) and model simulations (right). The red circles indicate regions with substantial differences

More details about the changes are given in the individual replies to the Reviewer comments below:

Reviewer comment:
The manuscript selects extensive existing teleconnection indices and aims to identify teleconnection patterns in a new global dataset. It presents a new method to examine the reproducibility of the teleconnection in the global dataset along with other data sources. Although the manuscript first employs the dataset to the teleconnection research, which shows its novelty, the quality of the presentation needs substantial improvements. In the current version, the manuscript intends to address the research questions mentioned in the introduction, but the presentation of the results is confusing and difficult to follow. For example, the manuscript (e.g., Page 2, line 66-67) frequently mentioned the aim is to investigate the influence of teleconnections on the global distribution of the total column water vapor (TCWV). As far as I understand, the paper does not clearly address this issue. Could the authors clarify and stress the influence of teleconnection on the TCMV in the manuscript?

Author reply:
As mentioned in point B) above, we tried to make our aims more clear. Our aim was not to investigate the influence of teleconnection on the TCMV or to predict monthly TCWV using teleconnections. As described in point B) above, this was made more clear in many parts of the manuscript. For the text mentioned above (Page 2, line 66-67), we modified it to:
'In this study we investigate to which extent the temporal patterns of various teleconnections can be identified in the global distribution of the total column water vapor (TCWV).'

Reviewer comment:
Hence, I suggest some restructure of the body text of the manuscript. For instance, it might be beneficial to clarify the relationship between different groups of indices and their corresponding results.

Author reply:
We agree and made the corresponding changes, see points B), C), and D) above.

Reviewer comment:
Also, I suggest the authors improve the quality of the figures.

Author reply:
We agree and made the corresponding changes, see point D) above.

Specific comments:

Reviewer comment:
1. The manuscript used the water vapor column data from satellite observations in the red spectral range. Is there anything special for the use of red spectral range in the paper? It will be good to give some explanations otherwise I suggest removing “in the red spectral range” from the title.

Author reply:

We changed ‚red' to ‚visible'. The important point here is that the satellite observations observe scattered and reflected sun light. Thus they are sensitive for the total atmospheric column.

We added the following explanation to section 2.1:
' The data analysis is performed in the red spectral range. Since these satellite instruments observe scattered and reflected sun light, the observations are sensitive for the whole atmospheric column including the surface-near layers which usually contain the largest fraction of the total atmospheric TCWV.'

Reviewer comment:
2. Page 3, line 105: Could the authors provide one or two references, which shows that the variations of the TCWV are strongly associated with ENSO events? Or can authors provide the correlations over the tropical band?

Author reply:
The following references were added:
Simpson, J. J., J. S. Berg, C. J. Koblinsky, G. L. Hufford, and B. Beckley, The NVAP global water vapor data set: Independent crosscomparison and multiyear variability, Remote Sens. Environ., 76, 112–129, 2001.
Soden, B. J., The sensitivity of the hydrological cycle to ENSO, J. Clim., 13, 538– 549, 2000.
Wagner, T., S. Beirle, M. Grzegorski, S. Sanghavi, U. Platt, El-Niño induced anomalies in global data sets of water vapour and cloud cover derived from GOME on ERS-2, J. Geophys. Res, 110, D15104, doi:10.1029/2005JD005972, 2005.

Reviewer comment:
3. In the third section (Page 4, line 130-145), the authors did a great piece of work on putting various existing teleconnection indices together. The manuscript divides those indices into groups but indices in the same group can have high correlations, like ENSO indices (Fig. A3). The authors could focus on some selected indices and omit other highly correlated indices unless the differences among those indices affect the conclusion of the manuscript. It would be good to see more discussions in the line 143-144 for Fig.4.

Author reply:
In this comment we see two important aspects:
a) the description of the correlations and the grouping of the indices should be improved.
We followed this suggestion, see point C) above. In particular we added new sub groups of indices to table 2.
Furthermore, we added the information that for indices with high correlation similar spatial patterns are found in the TCWV data set. In section 4.1 we added the following information:
' As expected, for groups of indices with strong temporal correlation also similar spatial patterns are found. This is most obvious for indices similar to the ENSO index (first group of indices in Figures A6 and A8). Similar spatial

patterns are also found for other pairs of indices, e.g. between the Hawaiian Index (HAW) and the Pacific Decadal Oscillation (PDO) as well as between the South Tropical Atlantic index (STA) and the Equatorial Atlantic Index (EA_errst)'

b) It is suggested to ‚focus on some selected indices and omit other highly correlated indices unless the differences among those indices affect the conclusion of the manuscript.'

In principle we agree to this suggestion. However, in our opinion this was already addressed in a systematic way in the original manuscript by applying the orthogonalsiation of the teleconnection indices. Our choice to use an orthogonalsiation has the advantages that it is mathematically straight-forward and avoids any ambiguities and arbitrariness in the selection of the ‚best' index out of a group of similar indices. Overall our procedure should be seen as a two step approach: in the first step all available indices are used, because it is beforehand unclear, which of them are most significantly detected in the TCWV data set. But by applying our method to all indices, we can answer the question which indices are most significantly detected.

In a further step we then apply the orthogonalisation to obtain a new set of indices without any correlation amongst them.

To make our aims and the procedure more clear, we added the following information to the section 3:

' Many of these indices (describing the same phenomenon), but also many of the other teleconnection indices are highly correlated. The strength of these correlations is presented in Fig. 3 as a matrix with correlation coefficients between the different indices (after the seasonal cycles were removed). In spite of the correlations amongst the teleconnection indices, we decided as a first step to include them all in our study, because beforehand it is not clear which index might be best suited to represent a teleconnection phenomenon. Using our empirical approach, however, it becomes possible to quantify the significance and strength of the different indices and thus to select the best suited index for a given teleconnection phenomenon. Finally, we apply an orthogonalisation for the most significant indices (see section 7) to minimise the effect of the correlations and to identify the dominant temporal teleconnection patterns in our TCWV data set.'

In section 7 the explanation was extended to:

' To account for correlations between the different indices, we thus applied an orthogonalisation approach. For the orthogonalisation (based on the Gram–Schmidt process), all ‚significant' original indices and significant temporal derivatives (see Figure A11) were considered (in total 57 indices). The order of indices used in the iterative orthogonalisation process was from highest to lowest p99 values. The result of the orthogonalisation approach is a set of modified teleconnection indices, which shows zero correlation amongst each other (for the considered time period). Thus this new set of orthogonalised indices can be used to determine the number of independent significant teleconnection patterns in the global water vapor data sets.'

and in section 8:

'The cumulative delta RMS map for the orthogonalised indices represents the overall contribution of teleconnections to the variability of the global TCWV distribution.'

Reviewer comment:
4. In the fifth section, the manuscript uses the reversed indices but a clear explanation on the reversed indices and its meaning is needed, e.g. what is the meaning of 'reversed'. The current presentation makes Fig. 7 hard to understand.

Author reply:
The description of the reversed indices was made more clear, see point A) above. In section 5.1 the following clarification was added:
'The basic idea of our new approach is to use non-geophysical indices for the estimation of the significance level. Non-geophysical indices are indices without any temporal correlation with the temporal variations of the investigated geophysical data sets. For that purpose we chose all temporally reversed indices (see Table 2 and Fig. A6), because they cover all relevant frequencies of the true teleconnections. In practice, the time axis is flipped, that means the first entry (July 1995) will be assigned to the last month (October 2015), and so on.'

Reviewer comment:
5. Page 7, line 289-292: the results here are interesting. Could the authors provide more physical or dynamical explanations behind these results?

Author reply:
We added the following information to section 6:
'For the TCWV data sets, surface temperature and pressure, as well as most of the zonal winds, the largest p99 values are found for indices similar to ENSO. For the TCWV data sets and surface temperature, this can be expected, because the ENSO phenomenon is driven by the surface temperature (over the tropical Pacific). Accordingly, also the TCWV data sets will be strongly affected, because the TCWV depends strongly on the temperature in the lowest atmospheric layers. The strong influence of the ENSO phenomenon (BEST index) on the zonal winds at most levels can probably be explained by the fact that large scale phenomena like ENSO can have a strong influence on the quasi-persistent zonal flow patterns in the tropics and sub-tropics. For the geopotential heights and meridional winds, the largest p99 values are found for the polar atmospheric indices (mostly AAO, but also SCA). For the geopotential heights this might be expected because the polar atmospheric indices are defined based on anomalies of the geopotential heights. Why also for the zonal winds, the largest p99 values are found for the polar atmospheric indices is, however, is not clear to us.'

We added also a comparison of the maximum p99 values to section 6 (we also added a new column to table 3). The respective text in section 6 is:

’ Our new method for the determination of the significance level also allows a direct comparison of the strengths at which the different indices are detected in the different data sets. In Table 3 also the maximum p99 values of the delta RMS normalised by the corresponding significance threshold values are shown. The highest normalised p99 values are found for the geopotential heights (except the 50hPa level) and the surface pressure. This finding is consistent with the fact that these quantities are used in most teleconnection studies and many indices are even defined using these quantities. The lowest normalised p99 values are found for zonal winds, for which also the smallest numbers of significant indices are obtained. Intermediate values are found for the water vapor data sets.’

Minor comments:

Reviewer comment:
1. Page 3, line 116-117: Could the authors add more descriptions for Fig. 2?

Author reply:
The following information was added:
’Similar patterns are found in all three data sets indicating the good consistency amongst them. The highest values are found over the tropics, especially over the west Pacific. Lower values are found towards higher latitudes showing the strong dependence of the TCWV on temperature.’

Reviewer comment:
2. Page 3, line 125: I suggest removing Figs. A1 and A2.

Author reply:
Both figures were removed as suggested.

Reviewer comment:
3. Page 4, line 138: The manuscript used the word ‘fit’ here, but the introduction of the fit function is shown in the next section. Moving the description of the fit function here might improve the clarity. Section 3 and 4 should be organized in a logical flow.

Author reply:
We reorganised sections 3 and 4 accordingly.

Reviewer comment:
4. Page 5, line 174: Could the authors clarify the meaning of a larger or lower value of delta RMS (eq. 3)?

Author reply:
After equation 3 the following information is added:
’The delta RMS is a measure for the magnitude of the variance of a considered data set, which can be explained by the chosen teleconnection

pattern. If there is high similarity of the temporal variation of an index with the temporal variation of the considered data set, the delta RMS values for both fits is large. If there is no similarity, the corresponding delta RMS value is zero.'

Also the following information is added:
'It should be noted that instead of the delta RMS values, also the correlation coefficients between the considered data set and the fit function (eq. 1) might have been used since the spatial patterns of both quantities are very similar (see Fig. A8).'

New Fig. A8:

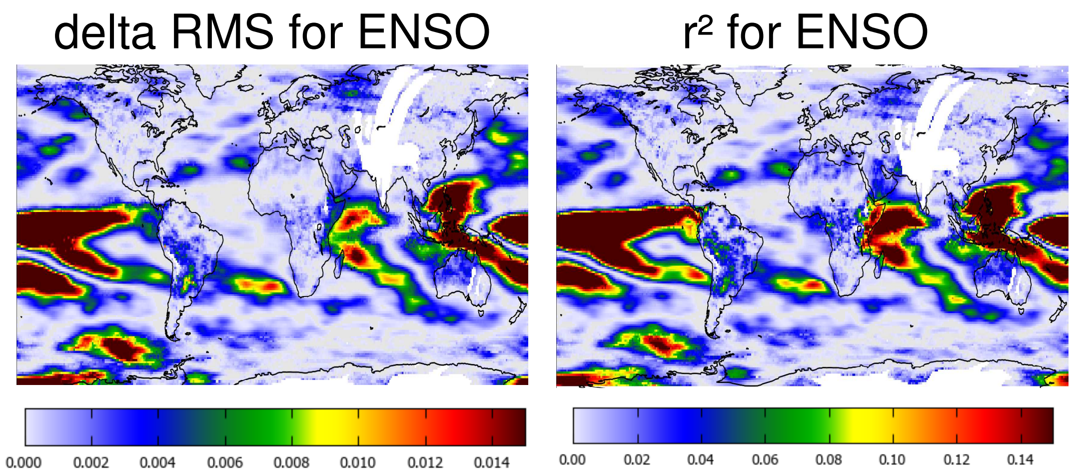


Fig. A8: Delta RMS (left) and r^2 values (right) for the fit of the ENSO index to the TCWV derived from satellite observations.

Reviewer comment:
5. Table 2 and Figure 5 captions.

Author reply:
Corrected, many thanks!