1. The abstract is too general. I would like to see more concrete results of the study in the abstract.

Will be considered in revision.

2. Page 1 line 29. What does mean here the term "homogeneous"? How can the results be more homogeneous? In which sense?

In the mentioned article (Netzel and Stepinski, 2016) it was obtained that "we demonstrate that clustering-based classification results in climate types that are internally more homogeneous and externally more distinct than climate types in the KGC".
Homogenity was measured as "For each class the entropy of a histogram of it constituent types measures a level of class homogeneity with respect to types. Homogeneous
classes (like A) have small values of entropy and inhomogeneous classes (like D) have large values of entropy"

3. Page 2 lines 3-5. This sentence is a bit unclear and confusing to me. It is written the loadings of components are the coefficients that define indices. I have an imagination that loadings of principal components show correlation between time series of the components and observed variables. Can you explain this?

"The loadings of chosen principal components are the coefficients that define the newly created indices, which then describe the main features of climate."

We have chosen to use most common terminology (at least this is mentioned as most common terminology in Wilks, 2006. Statistical methods in the atmospheric sciences. This book is not mentioned in references, as this terminology is also used in book by Jolliffe, 2002, that we use).

Values of principal components are calculated as P = Xe, where X is data matrix of initial data, e are eigenvectors or loadings and P are principal components. Explained variance is calculated from eigenvalues.

This means, that correlation coefficients are: corr.coef = loadings * sqrt(eigenvalues).

Also we will specify used terminology in revision.

4. Page 2 line 24. Should it be the reference to de Castro et al. 2007?

"RCM models are continuously improving and correspond rather well to climate observations (Castro et al. 2007)."

My mistake, while working on the manuscript the sentence had changed too much. Original quote: "Making use of the Köppen-Trewartha (K-T) climate classification, we have found that a set of nine high-resolution regional climate models (RCM) are fairly capable of reproducing the current climate in Europe" (Castro et al. 2007).

I will restructurize references in revision, as there are other references I have used that correspond to this statement. For example, Tapiador et al. 2011.

5. The introduction is lacking of the description of similar studies. PCA is widely used in climatology and also for determining of various climate indices. I suggest a literature overview where is shown how the current study is fitting into other similar studies. The novelty of this study should be clearly indicated.

Will be considered in revision.

6. The description of the use of model data could be more precise. Does the ensemble data mean that the averages of 22 model runs were calculated? What was the spatial resolution of the ensemble data? I suppose that the resolution was different for every single model run.

We will revise to make it accurate.

7. It was not clear why the model data were used instead of station data. The density of meteorological stations is rather high in the study area. Therefore, the results of the PCA of station data would be compared with the results of the RCM-based data.

Main reason was because we have model data for future period. Also model data was bias corrected based on station data so statistics for each case (and also PCA result) should coincide.

8. Page 3 line 9. It is not indicated from which source the observation data (Fig.1) were obtained.

Bias-correction was fully described in (Sennikovs and Bethers 2009). We are considering to remove the illustration of loacations and put more emphasis on reference.

9. Page 3 line 18. Usually, it is written "…as it is done by Malmgren et al. (1999) and Forsythe et al. (2015)".

Will be considered in revision.

10. Page 3 lines 22-23. A very strong correlation in winter precipitation is detected. Is it so that correlations are calculated using the data from the same year? In that case there is not any time lag. I don't believe that there is a correlation above 0.8 between January and December of the same year. I don't believe the statement that winters are either dry or humid. There should be something wrong. I did some calculations with station data of monthly precipitation and did not find any significant correlations. Correlations presented in Fig. 4 are inadequate. Such high correlations in monthly precipitation are not possible at all. All other correlation coefficients seem also suspicious. The reason for presenting the correlation Matrix is not clear.

We use <u>climatic</u> variables – 30 year average.Data matrix that is used for both correlation calculation and PCA consists of 24 variables (12 temperature and 12 precipitation values) and 7143 cases (grid points). It means that we are looking for spatial patterns, not temporal. This upproach is less used, but there are similar applications in literature (for examples, Fovell and Fovell 1993).

The reason for correlation matrix was to show that there is redundant information that should be reduced through PCA.

11. Page 4 line 18. I suggest the word "them" instead of "then".

We will revise to make it accurate.

12. Standardisation of climatic data is a trivial procedure. It is not clear why the variances on tables 2 and 3 are presented in the study. Were they spatial mean variances?

Standardization of data is an important part of PCA that can influence acquired result, and as we are using a bit different approach from the standard (standard meaning subtracting the mean of variable and dividing by square root of variance) it is important to both clarify what we're doing

and why. Main reason for tables 2 and 3 was to show the variances for each variable and the mean variance per variable category (temperature or precipitation) after we had performed or standardization procedure. Table 2 illustrates what we tried to accomplish (and succeeded) in our standardization procedure. Table 3 illustrates differences in future data (from reference period data in Table 2) in regard to variance of data. Table 3 shows increased precipitation data variance and reduced temperature variance in comparison to reference period. Also Table 3 shows that pattern of variances between different months haven't changed (I was considering illustrating this point, but this would just duplicate information in Tables 2 and 3).

13. I think that more information about the PCA procedure is needed. Was it rotated or non-rotated PCA? There are different modes of PCA: T-mode, S-mode etc. How the matrix was performed? Which were variables and which were cases? What were loadings and what were scores? This information is needed for the interpretation of results.

Will be elaborated in revision. We used non-rotated PCA, and matrix is described in point 10 of this document. Terminology we used is shortly summarized in point 3 of this document. (Principal components = scores. Loadings = eigenvectors). Also we will consider specifying used terminology and technique in revision.

14. Why the loadings of the three first components are presented in Table 4. What do they show?

They define approach of using PCA results as climate indices. (This is elaborated a bit more in point 19 of this document).

15. The spatial patterns of three first components are very interesting and informative. I think that they could be wider and better interpreted. It is clear that PC1 represents the influence of the Baltic Sea. It is the main factor causing spatial climatic differences in the Baltic countries. It is directly related to higher temperature and precipitation in autumn and winter, and lower temperature and precipitation in spring and early summer in the coastal regions. In the hinterland far from the sea the spatial coefficients (scores ?) are negative. In conclusion, PC1 reflects continentality of climate.

Useful input, will be taken into account in revision. It is important to note that values itself (negative, positive, etc.) don't have a meaning associated with them! For example our standardization process of subtracting mean resulted in negative values for precipitation. If we used a different approach, for example, subtracting minimum value, we would get different values of principal components (scores). However that wouldn't change interpretation (correlation coefficients) of PCA results or spatial pattern. What we can compare are regions (or their lack) with similar values.

16. PC2 reflects the second main factor in formation of climate – i.e. latitude. The pattern shows positive scores in Lithuania and negative scores in Estonia. The southern region is characterised by higher temperature, especially in spring and autumn, comparatively higher precipitation from April to June and lower precipitation during the rest of a year.

Useful input, will be taken into account in rewrite. About positive/negative values see point 15 of this document.

17. The spatial pattern of the PC3 is very similar to the mean annual distribution of precipitation in the study region (Jaagus et al. 2010). Two regions with higher precipitation are described by areas of negative coefficients - one in western Lithuania and Latvia, and another in the western part of continental Estonia and central Vidzeme upland in Latvia. Positive areas correspond to coastal regions with lower precipitation in Estonia and Latvia. But I cannot understand why spatial coefficients

(loadings) on Fig. 6 are negative but temporal coefficients (scores) in Table 5 are positive. I cannot fully understand the results of PCA.

Useful input, will be taken into account in rewrite. Due to standardization some values of precipitation are negative, that can result in negative coefficients. About positive/negative values see point 15 of this document.

18. I suggest that the authors do not interpret the results fully and not always adequately. If I understand correctly, interannual variations of temperature and precipitation are not reflected in the results of PCA. There are presented only mean monthly variability. Consequently, the results of current PCA reflect spatio-temporal variability of monthly mean values. Therefore, the interpretation of the results on page 5 is not valid in the following sentences: "This means that higher values of PC1 correspond to warmer winters …" (lines 15-16), "In general, higher values of PC2 correspond to earlier phenological processes" (lines 28-29), "This means that high PC3 values correspond to overall high precipitation and warm spring, or in other words – overall wetter year" (lines 31-32). If interannual variability is not included into the analysis, the relationships with phenological phases are not appropriate. Anyway, here are many problems to be clarified.

We will rephrase our interpretations in revision to avoid confusion.

19. It is not clear how the loadings were used to calculate climate indices for the future. I am not sure but it would be correct to realise PCA for the modelled mean values for 2071-2100 and analyse the results of past and future analyses.

One of the aims of this work was to see how our components (climate indices) change in future.

I will try to better explain reasoning for method used in this paper. So the idea is that we perform PCA and acquire principal components, and then once we have some kind of interpretation of principal component, we can just assume that it's a climate index. Try to discard for a moment the principal componets part and just think about the climate index

$$\boldsymbol{CI = a_1 T_1 + \cdots + a_{12} T_{12} + a_{13} P_1 + \cdots + a_{24} P_{12}}$$

We have these coefficients a_1 … a_{24} and we can just calculate this index for present and future.

An analogy would be for example growing degree days. If we want to compare change of growing degree days in present and future we would want to use same base temperature and calendar period to calculate the sum of growing degree days. Even though there might be an argument that in future there is shift in seasons or plant adaptation that would affect methodology. Similarly in our case, we wish to use the same coefficients a_1 … a_{24} for reference period and future, because only then can we make conclusions about the change in value of chosen climate index (It's important to note that currently we can make conslusions about general increase of climate index value or identify regions that have similar values. However we can't make any conclusions about numerical increase, for example, what is the meaning of increase of 5 or 10 units) . Because of the issue about comparison also there were some considerations about standardization process (especially application to future).

20. The section of discussion is pure. I recommend to restructure the paper. In the section of results, there could be only the description of the results of PCA. All interpretations might be included into the section of discussion.

We acknowledge the problem and will consider in it revision.