Dear Referee #1,

Thank you very much for your comments on our manuscript that will certainly be considered in a revised version and that will help to improve our contribution. At the current stage, we would like to reply and address a few of these comments:

"(2) *The methods appear rather complicated, except for the PCA which is well established and applicable in this context. Are the other methods also established or are they applied for the first time here? I do not understand why and how these methods were chosen. Further, I do not understand the benefit of investigating the persistence; and the added value of the behavioural measure analysis over the PCA."*

Authors reply:

Our main intention using all of the presented methods is to strengthen the reliability of our results. The first part on "persistency measures" is to our knowledge novel and our own contribution in terms of a new methodology of spatial data exploration. By applying these persistency methods we are able to confirm the existence of spatially and temporally consistent patterns within the time series of images. This finding supports the application of a principle component analysis (PCA) where the most dominant patterns (in the form of independent principle components) within the time series are extracted and information on explained variance by that PC is given.

It could be argued, that a PCA resulting in PCs with a significant high percentage of explained variance would be sufficient to confirm pattern persistency. However, there might be situation where 2 (or more) PCs with a high percentage of explained variance exist, but where e.g. some oscillating landscape behavior might result in non-persistent time series.

The part on behavioral measure analysis is a new approach to classify the dataset into functional units (or hydrological response units, here only under radiation driven condition, see Zehe et al. 2014 for an extended discussion). We assume that different loading values derived from PCA are related to a dominance of a different PC and therefore a different control on land surface temperature (LST) (and hence related to the functioning of the land and sub-surface as a reaction to the differing meteorological short time history and surface states). In this way we can choose a limited set of LST-images showing most distinct patterns. The derived classification by using the 5 most distinct LST-images is a representation of the spatio-temporal dynamics of LST and therefore of the "real landscape functioning". We are currently not in the situation to evaluate this procedure as superior to other classification methods (e.g. using the first 5 PCs, deviding them into a number of classes and intersecting them). Such an approach would involve a catchment scale hydro-meteorological modelling exercise, where different classification methods are compared with regard to effectiveness of parameterization and the quality of modelling results. While this is beyond the scope of this paper, it is motivation for current research and we will briefly add that in the outlook part of the paper.

Overall, we belief that the persistency analysis is a very helpful additional tool needed to avoid biased handling of the dataset. The behavioral measure is used to complete the PCA to spatially classify the catchment concerning the compartments' functioning.

"(1) Whereas the approach is useful as mentioned above, I am missing information on its novelty. Has anybody done this before? If not, why not outlining clearly that this is a novel approach. The introduction references studies by Anderson and Steenpass but differences and similarities to the present study remain unclear."

Answer:

As of our knowledge, neither in thermal remote sensing nor in catchment hydrology where the delineation of hydrological response units or functional units including their parameterization is subject to research is there any publication on the use of complex time series analysis of TIR data in combination with PCA as used here. In this sense, our approach is new. However, we recently got aware of the application of empirical orthogonal functions (EOF) that seem to be frequently used in oceanography and atmospheric research (e.g. Denbo & Allen, 1984; Hamlington et al., 2011; Lorenz, 1956). The approach is similar to PCA with an adjustment considering the extent of a single spatial data point/model output due to calculations on a global coordinate system and therefore occurring contortions. Nevertheless, the used data and the suggested applications differ largely. We will add a short section on the EOF in a revised version of the manuscript.

Anderson et al. (2011) and Steenpass et al. (2010) are both using similar thermal RS data within their work. However, Anderson et al. focus on the translation of thermal data into evapotranspiration data and, therefore, are limited to real data transformation based on the knowledge of physical processes. Steenpass et al. use the data to derive hydrological properties by the use of inversion. These two approaches differ largely from ours. They are mainly quoted to note **different** appliances of TIR data, as noted.

Further, we tried to address all specific comments in a way the problems are solved. Also, we will address for a language check by a native speaker.

We also improved the remarked images. You can find them on the following pages.

For any other changes, we want to wait for the second nominated referee's comments.

We acknowledge for the comments and we are disposed to add your suggestions. All issues will be revised in a new version of the manuscript.

B. Müller on behalf of all authors

References:

Anderson, M. C., Kustas, W. P., Norman, J. M., Hain, C. R., Mecikalski, J. R., Schultz, L., ... & Gao, F. (2011). Mapping daily evapotranspiration at field to continental scales using geostationary and polar orbiting satellite imagery. Hydrology and Earth System Sciences, 15(1), 223-239.

Denbo, D. W., & Allen, J. S. (1984). Rotary empirical orthogonal function analysis of currents near the Oregon coast. Journal of physical oceanography, 14(1), 35-46.

Hamlington, B. D., Leben, R. R., Nerem, R. S., Han, W., & Kim, K. Y. (2011). Reconstructing sea level using cyclostationary empirical orthogonal functions. Journal of Geophysical Research: Oceans (1978–2012), 116(C12).

Lorenz, E. N. (1956). Empirical orthogonal functions and statistical weather prediction. Sci. Rep. No. 1, Statist. Forecating Proj., Dept. Meteor., MIT, 49 pp.

Steenpass, C., Vanderborght, J., Herbst, M., Šimůnek, J., & Vereecken, H. (2010). Estimating soil hydraulic properties from infrared measurements of soil surface temperatures and TDR data. Vadose Zone J., 9(4), 910-924.

Zehe, E. and Ehret, U. and Pfister, L. and Blume, T. and Schröder, B. and Westhoff, M. and Jackisch, C. and Schymanski, S. J. and Weiler, M. and Schulz, K. and Allroggen, N. and Tronicke, J. and Dietrich, P. and Scherer, U. and Eccard, J. and Wulfmeyer, V., & Kleidon, A. (2014). HESS Opinions: Functional units: a novel framework to explore the link between spatial organization and hydrological functioning of intermediate scale catchments. Hydrology and Earth System Sciences Discussions, 11(3), 3249-3313.

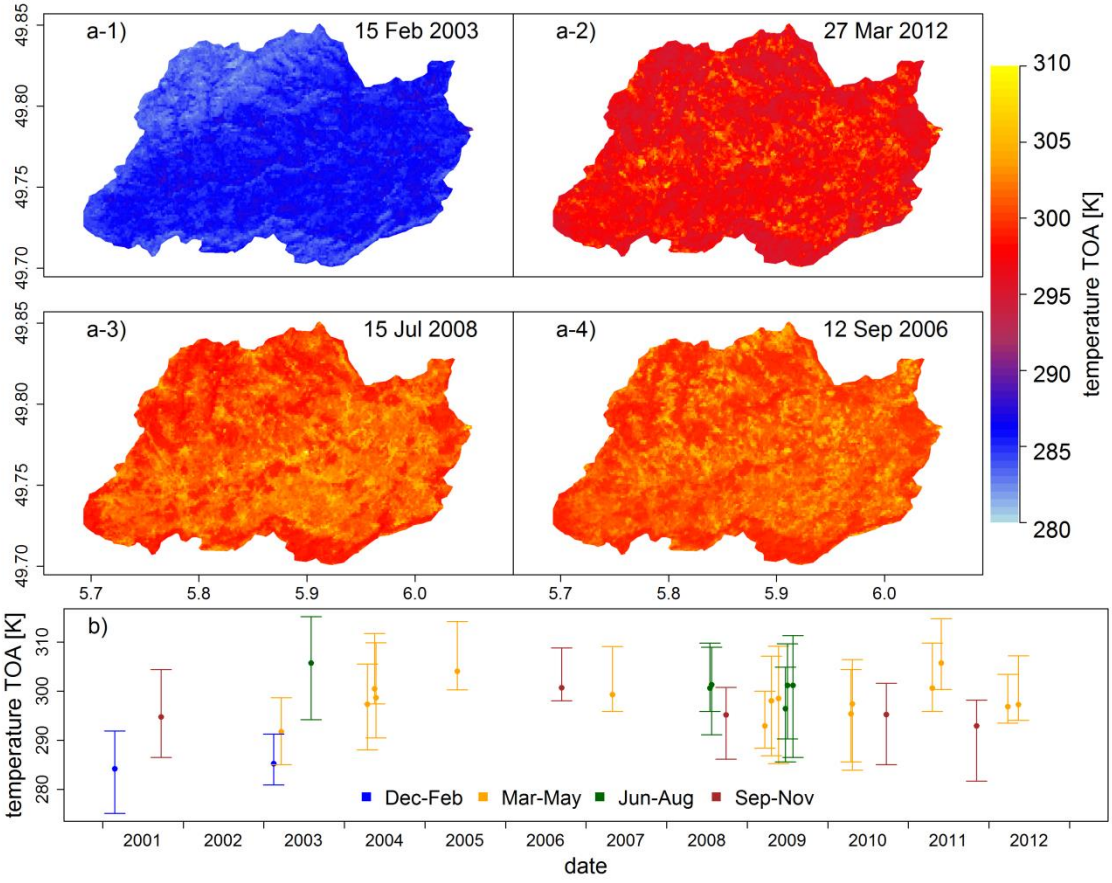

Figure 2: a) Examples of single band top-of-atmosphere (TOA) temperature time series covering winter (1), spring (2), summer (3) and autumn (4). b) Basic temporal and statistical information (mean, ranges) of the image time series.

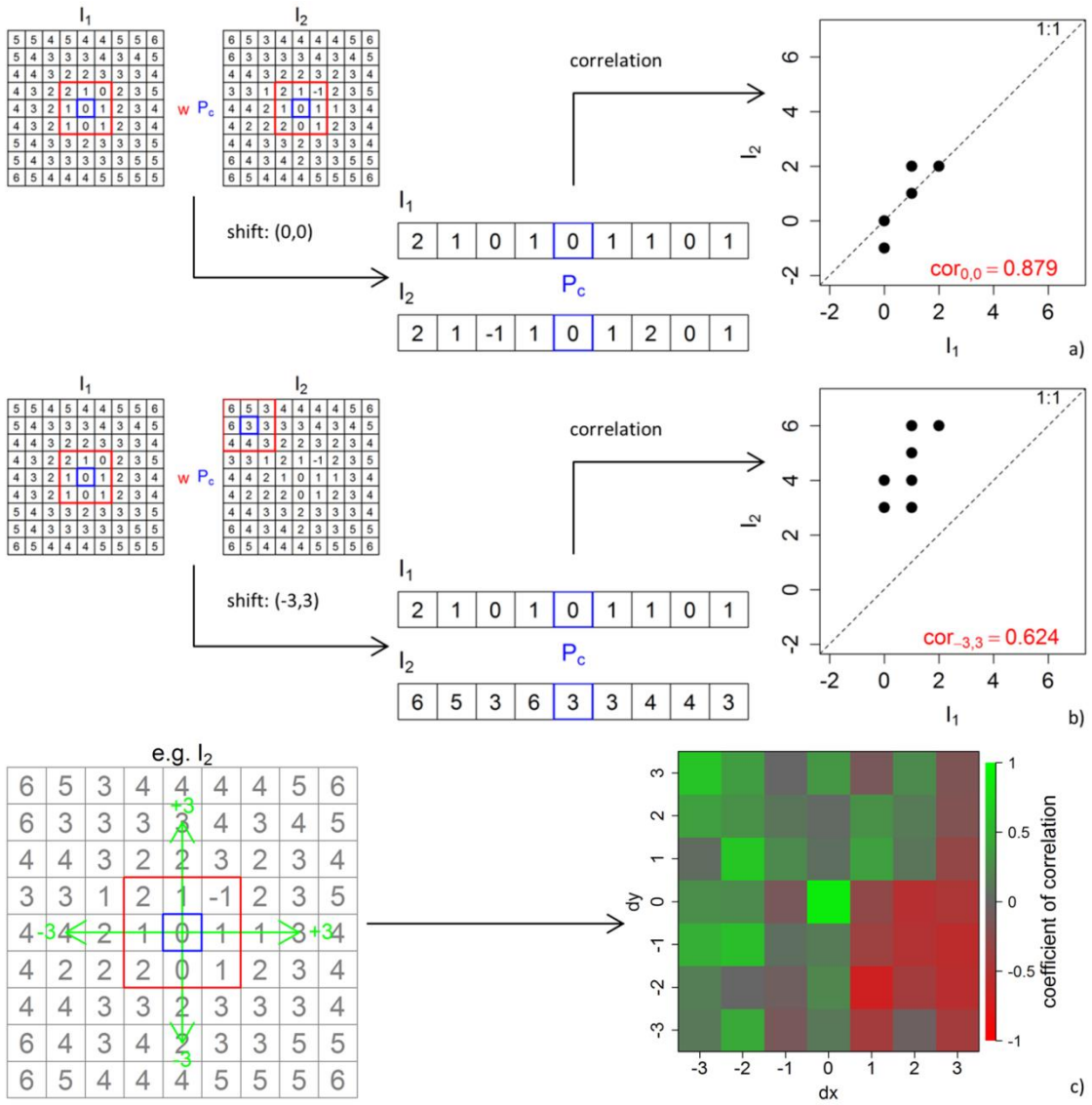

Figure 3: Analysis for the coefficient of correlation for a designed spatial dataset. We added small normal distributed noise to a concentric spatial pattern I_1 to construct I_2 and show the correlation for an extracted window w (red) around the central pixel P_c (blue) in the same position (a), in different positions (b) and for the whole image I_2 within the maximum ranges [-3,+3] (c).

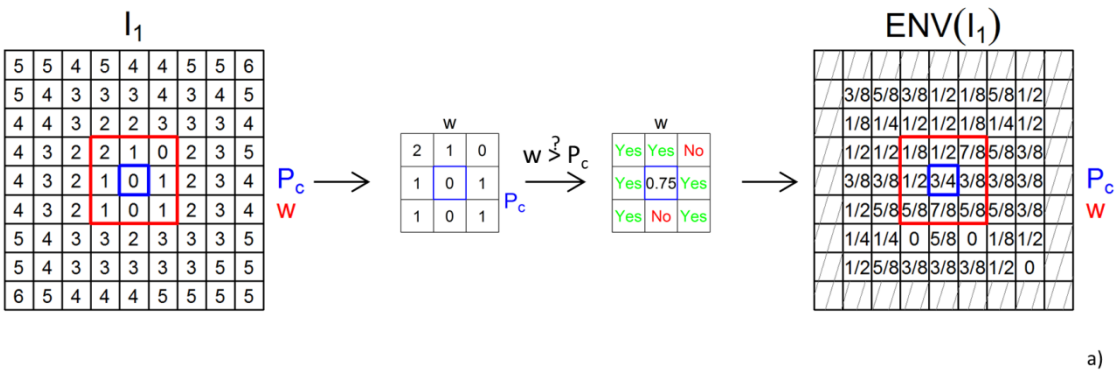

Figure 5: Analysis of the coefficient of variation via an “environment assessment” for a designed dataset. The data are generated in the same way as in the previous analysis (see Fig. 3). Subfigure (a) illustrates the derivation of a single summary value for the central pixel P_c (blue) from the data of the surrounding environment *w* (red). The example here investigates how many values within the environment are larger than the central value. This is repeated for all image pixels (except for boundary pixels) resulting in the leftmost picture.

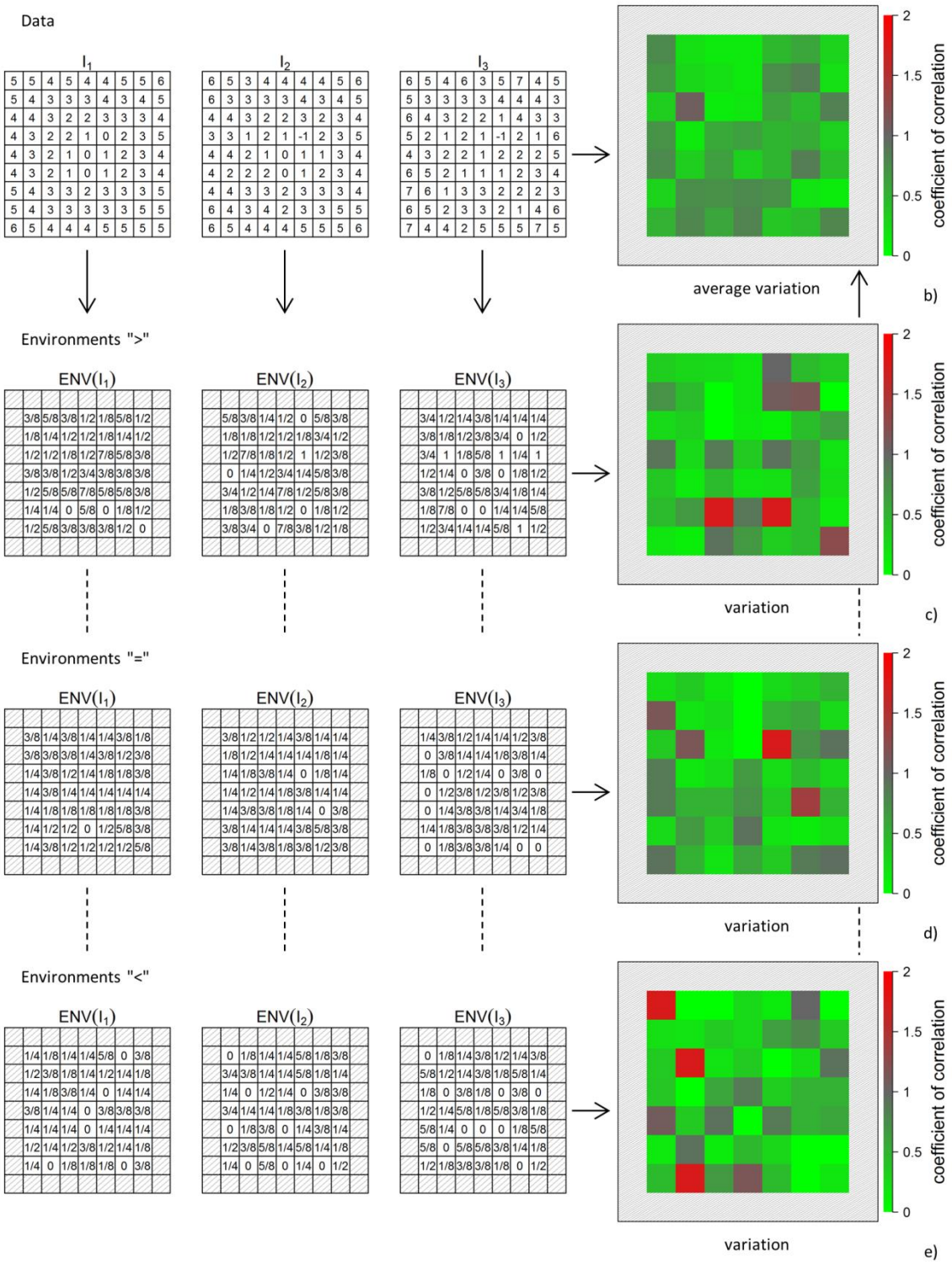

Figure 5 continued: Subfigures (b-e) illustrate the procedure from dataset (b, left) to the environment measures (c-e, left), to the coefficients of variation for different environments (c-e, right) and to the final describing average pattern (b, right).

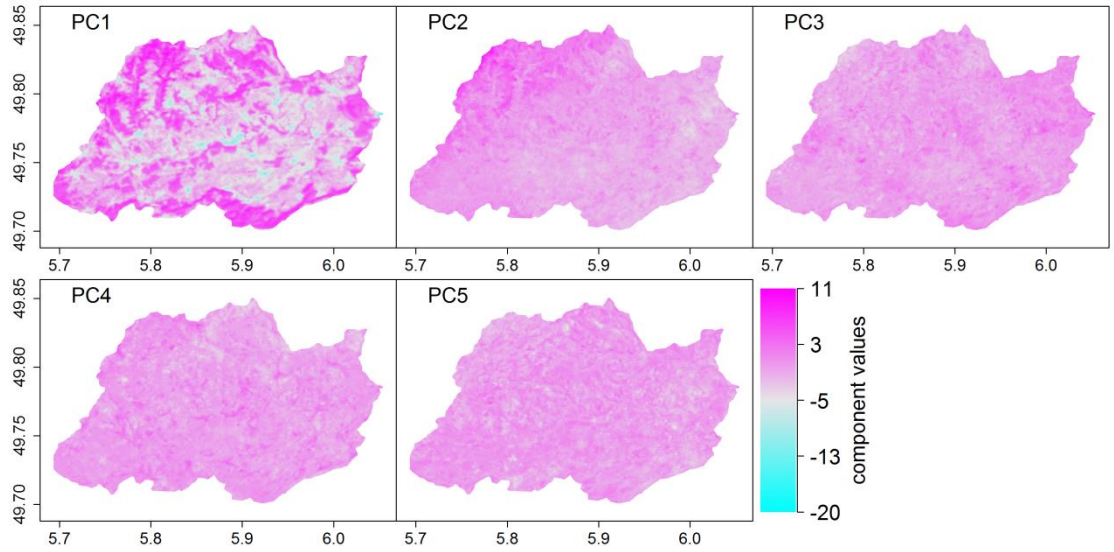

Figure 8: The first 5 components of the PCA for the LST time series data.

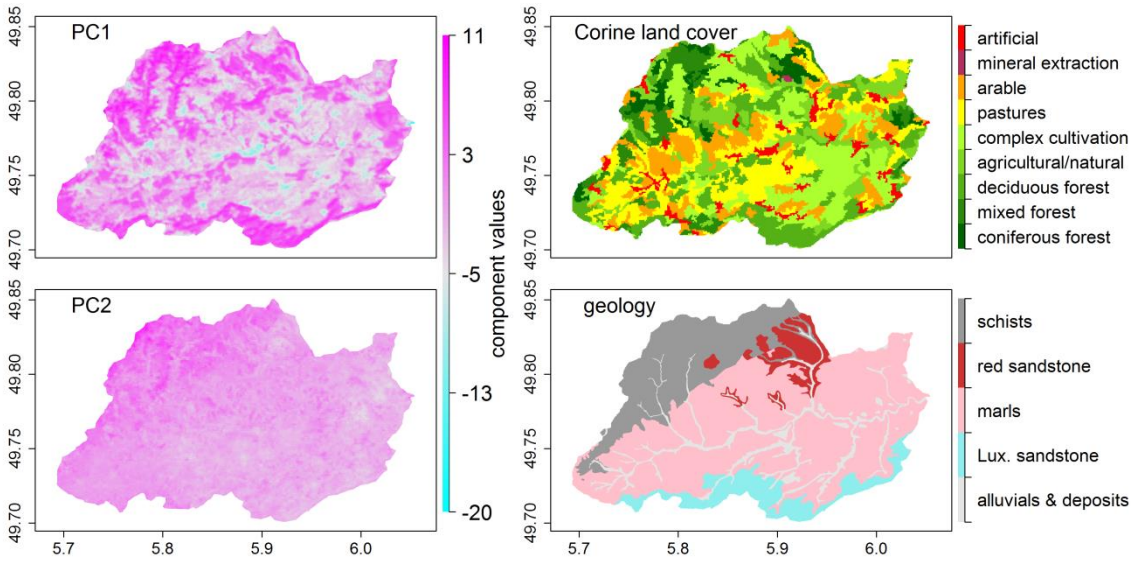

Figure 9: The first and second component of the PCA for the LST time series data (left) next to the patterns of the illustration of Corine land cover and geology data (right) of the Attert catchment.

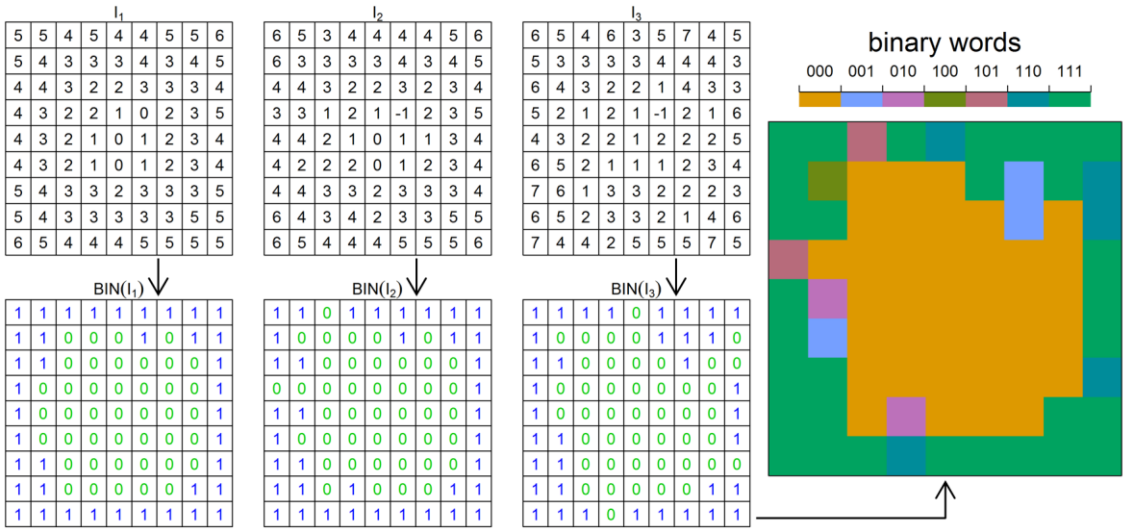

Figure 11: Construction of “binary word” classification for a designed dataset. The data are the same as for Fig. 5. On the left, the three images are binarized (BIN) from the upper to the lower panel. Values larger than the median are converted to 1 (blue), values lower are converted to 0 (green). The right panel shows the aggregated words for the three datasets. Not every possible occurrence of words is produced (maximum: 2^3=8).