**Response to Reviewer #1 of the manuscript**

# "Recent trends of groundwater temperatures in Austria"

**by Benz et al. submitted to *Hydrology and Earth System Sciences*.**

Manuscript Number: hess-2017-663

Revision due before: 4 May 2018

**Reviewer comments:**

This paper addresses an important and interesting topic regarding the influence of atmospheric warming on groundwater temperature (GWT) in shallow systems. The authors used temperature records from 229 wells located in Austria and climatic data from weather stations installed nearby the wells. The positioning of the paper within the framework of studies devoted to the impact of climate change on hydrological system is well presented. The authors found that nationwide temperatures of groundwater increase and correlate statistically well with surface air temperature (SAT). Additionally, authors have used linear and step-wise models to describe the evolution of temperatures. Based on the step-wise approach (which seems to be more accurate than the linear model) the authors have identified that groundwater respond to climate regime shifts with sudden increase in temperature. This paper has been carefully prepared and is well written. The conclusions will definitely trigger the attention of the scientific community and the readers of HESS. Nevertheless I believe that some points need to be clarified before publication.

**Reply:** Thank you very much for your kind words and constructive comments. Your concerns are addressed in the following.

**Rev #1 General comments:**

**Rev #1, Comment # 1:** Some aspects of the methodology are not clear or absent. More details on how the 229 wells investigated in this study have been selected is required. More information regarding the type of sensors used to monitor GWT would be helpful to appreciate the quality of the data analyzed. More information regarding the regression approaches is also needed. How the shifts in regimes are determined in the step-wise model (mathematically speaking)? I also raise some additional points regarding the methodology in the specific comments.

**Reply:** We agree. While no specific information on the type of sensors used for monitoring is available to us, the information on the well selection strategy was extended: *"In Austria, GWTs up to Dec 2013 are provided by the Austrian Federal Ministry of Sustainability and Tourism Directorate-General IV. - Water Management (BMNT, former Federal Ministry of Agriculture, Forestry, Environment and Water Management (BMLUFUW)) in 1138 wells. Here, we focus on all wells with a measurement depth of less than 30 m, a record of at least 20 years and no major*

*breaks (> 3 month) in the last 20 years of the time series. Hence, all studied wells are monitored at least since Jan 1994, and some already since 1966 (see Fig. S1a for more information). Additionally wells impacted by geothermal hot springs were excluded. Overall in this study …"* (lines 109ff).

The paragraph discussing the linear analysis now reads (lines 152ff): "*Equivalent to the work by Lee et al. (2014), a linear temperature change was determined for all 227 wells. For this, a linear regression model of the annual mean temperature data was determined in Matlab 2016b. Because all wells in our dataset were continuously monitored between January 1994 and December 2013, only this timeframe was analysed.*"

Information on the step-wise model was also extended (lines 159ff). The mathematical basis of the method can be found in the given references: *"… in recent years the method by Rodionov (2004) became standard. It identifies the significance of each possible shift by calculating the so-called Regime Shift Index (RSI): the cumulative sum of the normalized differences between the observed values and the long term mean of the assumed regime. Only shifts with a positive RSI are considered significant, and a higher value of RSI denotes a more pronounced CRS. The entire algorithm is described in detail by Rodionov (2004). This sequential analysis is data driven and requires no prior knowledge of the timing of possible shifts. It was updated to further include prewithening in order to reduce background noise (Rodionov 2006) and is available online as a Microsoft Excel add-in (NOAA). In this study we applied the method to the complete timeseries of all 227 wells and 12 weather stations. Because the algorithm cannot handle gaps within the analysed series, gaps in our data were filled using a linear fit …"*

**Rev #1, Comment #2:** I believe that there is a discrepancy between the original objective of the paper, which aims at highlighting impact of climate change at regional (country) scale (Line 12), and the description of potential local effects for (some) specific wells and locations. Indeed, the authors describe potential factors which could explain uncorrelated data locally. Local information that are made available to the reader are to my opinion not sufficiently detailed to support the arguments. The conclusions are consequently difficult to trust. I would recommend to separate the description of local factors from the result of the regional statistical analysis (which to my opinion constitutes the novelty of this study). The local impacts could be introduced in a separate discussion section. In this specific section, the authors could provide an exhaustive list of potential factors that could explain uncorrelated data along with some examples from specific sites to illustrate the hypothesis.

**Reply:** We agree and separated local and countrywide results more clearly in the Results chapter. Now all subsections (correlation, linear fit, and climate regime shift) have a countrywide discussion first and a more local discussion second. Additional detail for each specific location such as the immediate surrounding, land use, or similar is now given where applicable.

**Rev #1, Comment #3:** I believe that the conclusions of this paper could be strengthened by performing a more robust multivariate statistical analysis (Principal Component Analysis for example) considering more factors which might have an influence on GWT, integrating not only SAT but also geology, land cover evolution, water level variation, precipitation, population dynamic, length of the temperature time series:

**Reply:** Thank you very much for this comment, this would be very interesting indeed and we hope to implement this in future studies. However, the suggested analysis is far beyond the scope of this manuscript and most of the mentioned parameters such as geology, land cover evolution, water level variation and population dynamic, are not available to us yet. If you are interested in this topic please feel free to contact us at susanne.benz@kit.edu.

**Rev #1 Specific comments**

**Rev #1, Comment # 4:** Line 68: "...over decades". Please be more precise here.

**Reply:** The text was changed to *"...GWTs of 227 wells in Austria, measured in part since 1966, are analysed ...*" (line 80).

**Rev #1, Comment # 5:** Line 73: "... step-wise increases between the regimes". This is not clear to me. What regimes? Please clarify.

**Reply:** We agree. Text was changed to: "*These control atmospheric temperatures as well and are often described as sudden, step-wise temperature changes separating stable periods, called climate regimes."* (lines 85ff).

**Rev #1, Comment # 6:** Figure 1 b. needs clarification. The presence of 3 curves is confusing. Could you, for example, make the inner percentile filled with transparent colors?

**Reply:** We agree. The dashed line was used to show the $95^{th}$ percentile, but transparent color is a better idea. We changed it accordingly:

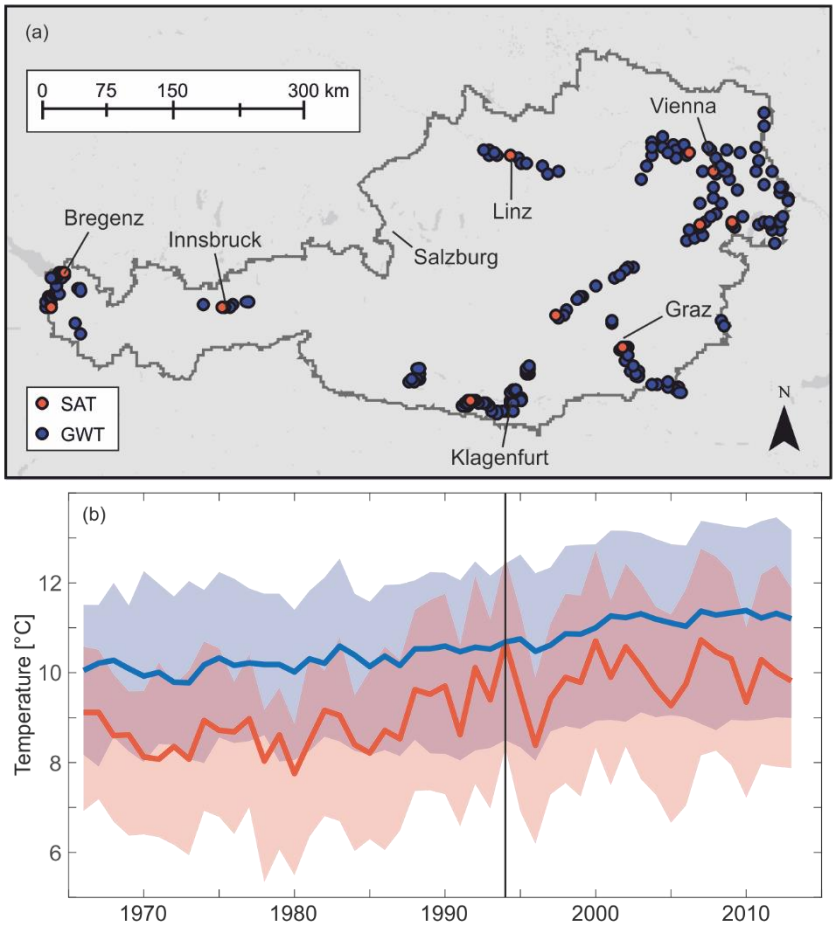


**Figure 1. (a) Location of all analysed groundwater temperature (GWT - 227 wells) and surface air temperature (SAT - 12 weather stations) measurement points; (b) temporal evolution of the spatial median, annual mean temperatures for groundwater (blue) and air (red). The inner 90 percentiles are marked in lighter colours. All time series were monitored since at least 1994.**

**Rev #1, Comment # 7:** Line 98: How the wells have been selected? What proportion of wells has been excluded from the database? See general comment.

**Reply:** Information on how the wells where selected is now given in the manuscript. See Rev #1, Comment #1 for details.

**Rev #1, Comment # 8:** Line 128. Please clarify why you choose 1994 as initial time for fitting.

**Reply:** A clarifying sentence was included: "*Because all wells in our dataset were continuously monitored between 1994 and 2013, only this timeframe was analysed*." (lines 152f).

**Rev #1, Comment # 9:** Line 129: Knowing which software you used is not informative here...

**Reply:** We agree. It is not necessary to know, but as it is common curtsey to give this information, we would like to keep it.

**Rev #1, Comment # 10:** Line 132: Please justify the choice of using the Spearman correlation coefficient and provide references.

**Reply:** Spearman correlation was chosen as it is more robust to outliers than other correlation measures. A sentence was added for clarification (lines 145f): *"Within this study, the Spearman correlation coefficient was determined, as it is especially robust to outliers caused for example by heat waves, which impact air temperatures but have only minor effect on groundwater temperatures."*

**Rev #1, Comment # 11:** Line 133: Taking annual mean values calculated with 8 months of data only may introduce some bias... Considering only years with full year of data would be more robust to my opinion. Otherwise, please discuss the limitations in the text. It is also not clear why yearly averages are used in the correlation analysis while the linear regressions are performed on monthly mean temperature (Line 129).

**Reply:** We agree, the linear analysis was changed and is now working with annual mean data as well. Interestingly, this decreased the determined temperature change for both GWT and SAT by about 0.1 K per 10 years compared to the analysis with monthly mean data. So far we are not certain what causes this discrepancy.
Additionally, the process of getting annual mean data was also revised, and gaps in the time series are now filled before determining the annual mean. The procedure is described in the chapter "Groundwater Temperatures" (lines 114ff): *"Overall, in this study annual mean data of 227 individual wells from all over the country (Fig. 1a) are analysed. Years with less than 9 months of data are excluded. For the timeframe 1994 and 2013, this amounts to 74 excluded data points in 60 wells. Additionally, only 9-11 months of data were available for 260 data points in 122 wells. To minimize the associated bias, these small gaps in the time series were filled using a linear fit. Hence small errors for years without a full set of monthly mean data have to be expected."*

**Rev #1, Comment # 12:** Line 132-136: It would be interesting to perform complementary correlation analysis accounting for other parameters such as depth of the wells, depth to the water table, geology, vegetation and land use. This could be assess with multivariate methods such as PCA. This could add valuable picture of the factors influencing the results.

**Reply:** We generally agree, however this analysis is beyond the scope of this manuscript. Please find our previous reply to Comment # 3 for more information.

**Rev #1, Comment # 13:** Line 145: "Breaks within the data were filled using linear fit". This is not clear... Please provide more information why you have to fill gaps for this analysis (and not for the other analysis?).

**Reply:** More Information on of the climate regime shift analysis is now given in the manuscript. See Rev #1, Comment #1 for details. This includes further information on why gaps have to be filled: "*In this study we applied the method to the complete time series of all 227 wells and 12 weather stations. Because the algorithm cannot handle*

*gaps within the analysed series, gaps in our data were filled using a linear fit.*" (Lines 166ff).

**Rev #1, Comment # 14:** Lines 160-164 and Figure 2a and b: This part require clarifications. As the authors stated, it seems that the shape of Austria (political boundary) might influence the results. Also the topography, with E-W strike orientation, might also have an influence. It is not so surprising that the correlation is better E-W that N-S (same latitude and orientation of topography). I am wondering if the figures are really informative...the decreasing correlation with distance in the figure a) is not obvious with the sharp increase at 550 km... Does this distance correspond to a decrease of the number of wells considered in the calculation?

**Reply:** We agree. Hence, we changed the entire paragraph to clarify this issue (lines 183ff): "*Additionally, the correlation between two wells seems to be anisotropic: correlation coefficients between two wells decrease faster with north-south distance than with west-east distance (Fig. 2b), which can be explained by the dominant striking direction of the geology and the resulting topography in Austria, where valleys generally run from west to east. Hence, larger rivers typically follow this direction and wells at the same latitude experience similar temperature signals."*

Additionally the number of pairs of wells is now also given for each distance in Figure 2a):

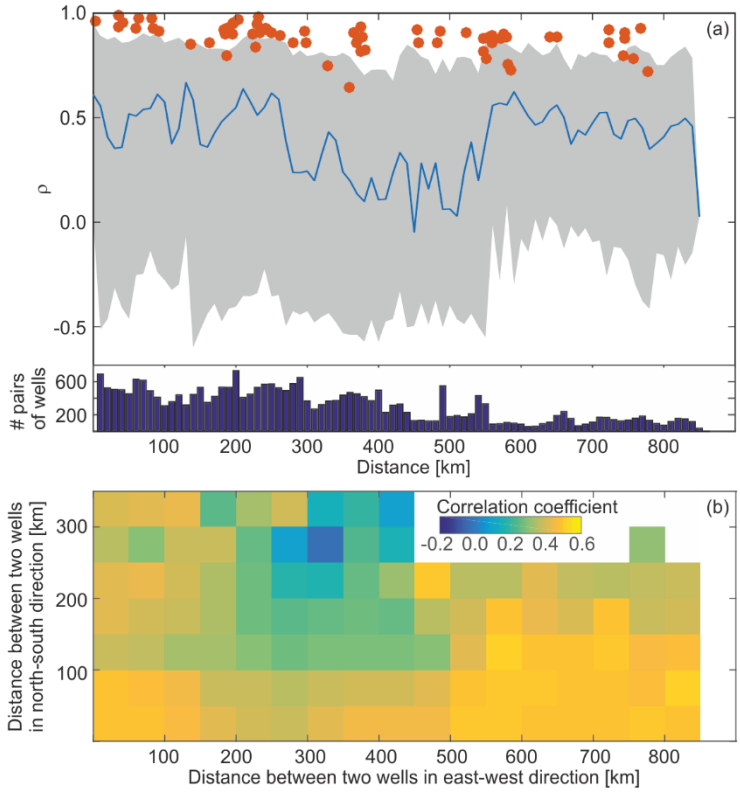


**Figure 2. Influence of distance on the correlation between the annual means of two measurement points. a) Correlation between SAT time series is given in red, median correlation between GWT time series is given in blue. The inner 90 percentile are coloured in grey, the number of pairs of wells per distance is shown in dark blue below. b)**

**The colour gives the median correlation between GWTs of two wells in relation to their absolute distance to each other in east-west direction (x-axis) and in north-south direction (y-axis).**

**Rev #1, Comment # 15:** Figure 3 is interesting but difficult to read. Would it be clearer if you display the relative change in temperature for all the wells? What are the p values here (not introduced in the text)?

**Reply:** We agree, the figure gives now relative temperature change, and p-values are provided:

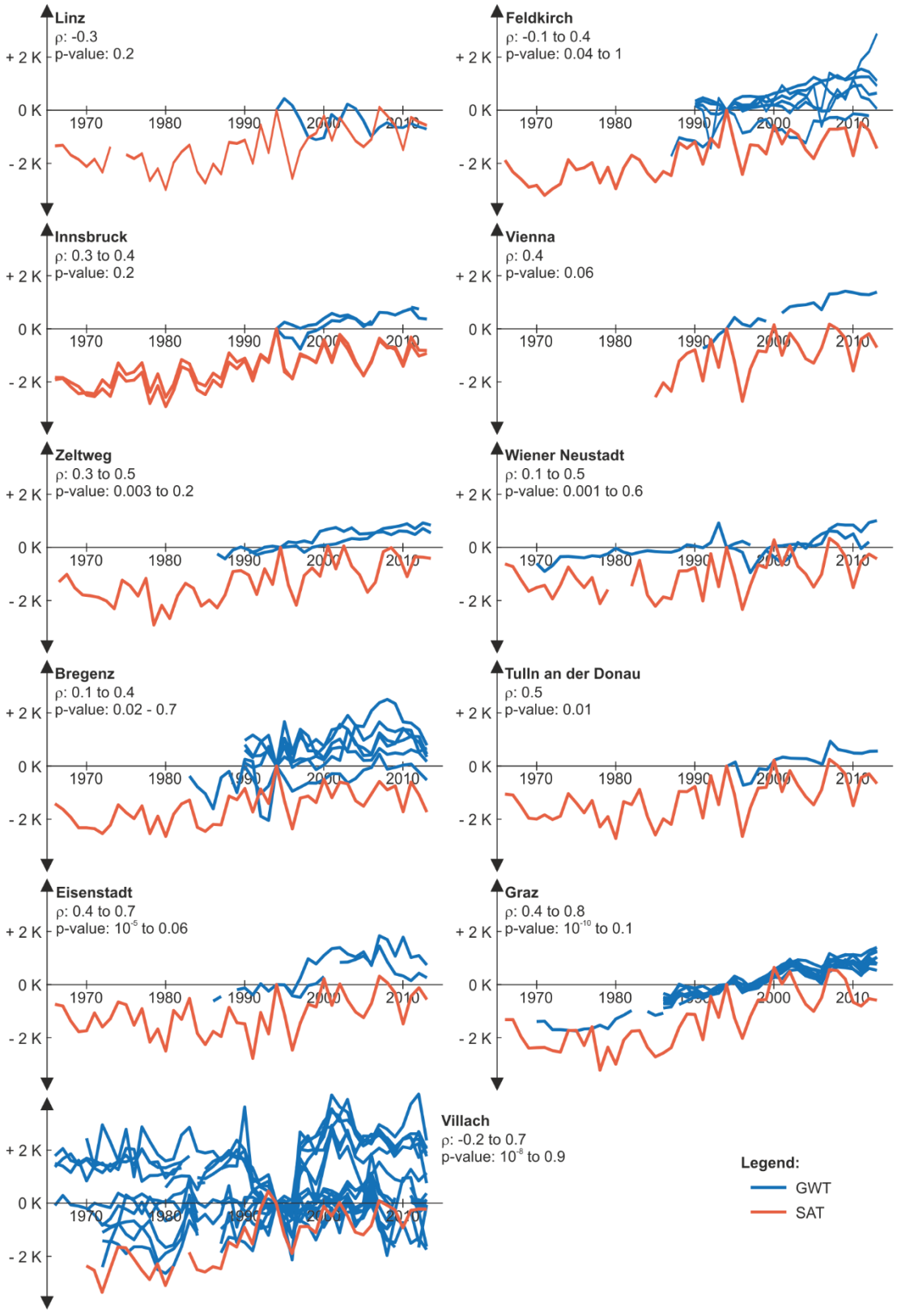


**Figure 3. Change from 1994 in surface air temperature (SAT) and groundwater temperatures (GWTs) of all wells within 5 km of the analysed weather station. See Fig. S3 for an overview of the locations. Minimum and maximum correlations and p-values between individual wells and weather stations are given.**

**Rev #1, Comment # 16:** Line 175 -176: To what coefficient are you referring to? The p values in the figure 3?

**Reply:** It is the Spearman correlation coefficient. For clarification, we changed the sentence as follows: *"… and Spearman correlation coefficients are < 0.5 …"* (Lines 201f).

**Rev #1, Comment # 17:** Lines 187-190: Here it seems that the length of the time series is critical in the interpretation of the correlation analysis... Please discuss this point.

**Reply:** The discussion of the length of the time series was extended as follows: "*The well with the highest correlation of 0.80 to SAT is located less than 1 km from the weather station close to the airport parking lot next to suburban housing. It is continuously monitored since 1970 and the longest time series in the area. The well with the lowest correlation (0.45) to the weather station here is located slightly to the east near a dog-park and suburban housing. Here observations started in 1994, it is the shortest time series in this area. At all other wells, measurements began in 1986 and show correlations between 0.6 and 0.7 to SAT indicating that the duration of the measurements play a significant role for local comparisons. In contrast, duration of the time series appears to be of minor importance on a countrywide scale. For example, the long time series in Wiener Neustadt (Fig. 3), which started measurements in 1970 and is located near a mineral extraction site, has a correlation of 0.48 and is therefore comparable to the short time series in Graz, starting in 1994 located in a suburban area."* (Lines 212ff).

**Rev #1, Comment # 18:** Table 1: What does p-value mean here? Not introduced in the text or the caption....

**Reply:** Sorry for being unclear. Those are the p-values of the correlations. An explanation was added in the table caption: "*Correlation coefficient and corresponding p-value between spatial median SAT and spatial median GWT for all analysed SAT locations and additional information* "

**Rev #1, Comment # 19:** Lines 205-206: Reference to table is missing. It is actually not a big difference of correlation coefficient 0.36 vs 0.24... The comparison with population density is not obvious to me from these values. Please clarify. The influence of city center and development of urban area is actually critical. Could it be possible that the increases in temperatures are partly related to urban development? Identifying the correlation with such factors could be assessed with a multivariate correlation methods (PCA).

**Reply:** We agree. Hence, a reference to the table was added and the paragraph (lines 234ff) was changed following also the recommendation by Rev. #2:
*"In addition, the data indicates that city size or rather population of the city does not necessarily influence the correlation between GWT and SAT (Table 1). For example, both locations Graz (population of more than 250,000) and Eisenstadt (population of 13,000) have similar correlation coefficients despite their vastly different population. Meanwhile, Bregenz and Feldkirch have a similar population (~30,000) and number of wells (six), but different correlation coefficients (0.52 and 0.19). However, it is also important to note that not all wells analysed here are located in the city centre, still all of them are within close proximity (< 250 m) of build-up and urban areas (Fig. S3)"*

Regarding the second part of your comment: The comparison of temperature increase and land cover type in Figure S2b indicates that there is no link between land cover (such as artificial surface and forest) and temperature change on the scale analysed here.

**Rev #1, Comment # 20:** Line 214-215: This difference in average changes in temperature with higher values for GWT than SAT is surprising… Could it reflect the effect of urban development or other anthropogenic activities (pumping, injection, heating system...).

**Reply:** Yes, it is not as expected. However, it is not only observed in urban areas, but also in rural areas. Our current hypothesis for slightly higher GWT increase than SAT increase are due to the chosen timeframe 1994 to 2013. In 1994 there was a heat wave over Central Europe and annual SAT was considerable higher than the long-term average at that time. The text was extended to include this discussion (lines 248ff): *"During the time between 1994 and 2013, GWTs have changed on average by +0.36 ± 0.44 K per 10 years and SAT on average by +0.24 ± 0.13 K per 10 years. The lower changes in SAT are most likely due to the chosen timeframe: A heat wave in summer 1994 led to extraordinary high annual mean SAT in this year (Figure 1b) and thus impacts the determined linear temperature change."*

**Rev #1, Comment # 21:** Line 226: Please provide a reference to the figure supporting the statement that spatial pattern of temperature changes is visible...

**Reply:** Reference to Fig. 4 was added (line 266).

**Rev #1, Comment # 22:** Lines 226-235: Too few information are available on the effect of this flood event. What was the difference in temperature between the river and GW during the event? Did it cover the entire well area? Estimated volume? Please provide more information or I would recommend to remove this paragraph.

**Reply:** We agree and removed any discussion of the flooding as insufficient information is available to prove a link between both events.

**Rev #1, Comment # 23:** Line 236-249: It is somehow surprising and confusing how local effects are introduced again... I believe that it should be discussed in a dedicated section discussing potential hypothesis that may explain uncorrelated data with eventually some examples of local factors from specific sites as examples.

**Reply:** We agree, the Results chapter was restructured, and countrywide and local factors are now discussed separately. See our reply to your comment #2 for more information.

**Rev #1, Comment # 24:** Line 247-249: Do you mean that the hot springs appeared suddenly?... I imagine that they were active before and constitute a constant temperature boundary...

**Reply:** The hot springs are known since the roman ages and their touristic use goes back to much earlier times than the beginning of the herein used monitoring data. We assume

there was some construction work or something similar going on and hydrogeological conditions changed. However, we could not find concrete evidence of this. Following Rev #2 Comment #2, all wells dominated by this hot spring are now taken out of the analysis.

**Rev #1, Comment # 25:** Line 262: I do not understand what the authors mean by "spatial median annual mean"... please clarify.

**Reply:** We agree. The sentence was simplified as follows: *"All detected climate regime shifts (CRS) of the spatial median temperature time series are shown in Fig. 6a.."* (Line 293).

**Rev #1, Comment # 26:** Lines 263 - 266: I am confused here. How do you explain that the shift in GWT occurs earlier than for the SAT? If the "CRS method (do you mean step-wise method) cannot be used to determine the precise timelag between GWT and SAT" why do you use it?

**Reply:** Yes, the original sentence was unclear. The CRS method was previously used to determine the time lag between GWT and SAT, but our results indicate that it is actually not precise enough to do so. The sentence was therefore changed following a suggestion also by Rev #2 Comment #25: *" ... the shift in the late 90s appears earlier and is more significant in GWTs. However, because SATs are the drivers of GWTs and not vice versa, the fact that the GWT change precedes the SAT change suggests that this method does not have the necessary resolution to determine short time lags between SATs and GWTs. Accordingly ... "* (Lines 298ff).

**Rev #1, Technical corrections:**

**Rev #1, Comment # 27:** Line 29-31: Reference is missing.

**Reply:** Two references were added: "*While, already at depth of a few meters, the amplitudes of periodic diurnal and seasonal temperature trends are strongly attenuated (Taylor and Stefan, 2009), long term non-periodic changes of air temperature permanently influence the subsurface down to greater depths of several tens to hundreds of meters (Beltrami et al., 2005)."* (Lines 30ff).

Beltrami, H., Ferguson, G., and Harris, R. N.: Long-term tracking of climate change by underground temperatures, Geophysical Research Letters, 32, 1–4, 2005.
Taylor, C. A. and Stefan, H. G.: Shallow groundwater temperature response to climate change and urbanization, Journal of Hydrology, 375, 601–612, doi:10.1016/j.jhydrol.2009.07.009, 2009.

**Rev #1, Comment # 28:** Line 58: Reference style for Menberg et al. (2014).

**Reply:** Was updated from "*(Menberg et al., 2014)"* to "*Menberg et al. (2014)"* (line 69).

**Rev #1, Comment # 29:** Line 72: Check reference style.

**Reply:** Was changed to "*e.g. Minobe (1997) and Rodionov (2004)"* (line 84).

**Rev #1, Comment # 30:** Line 128: should be "Equivalent to the work by Lee et al. (2014)".

**Reply:** Sentence was changed from *"(Lee et al., 2014)"* to *"Lee et al. (2014)"* (line 151).

**Rev #1, Comment # 31:** Labels of figure 2b could be changed by Northing and Easting.

**Reply:** To avoid any confusion labels of both axis were changed to "*distance between two wells in east-west / north-south direction [km]*".

**Rev #1, Comment # 32:** Figure 6. Please add legends to your figures.

**Reply:** Legend was added to Figure 6. Additionally the bar plot was changed to transparent following Rev #2, Comment #27.

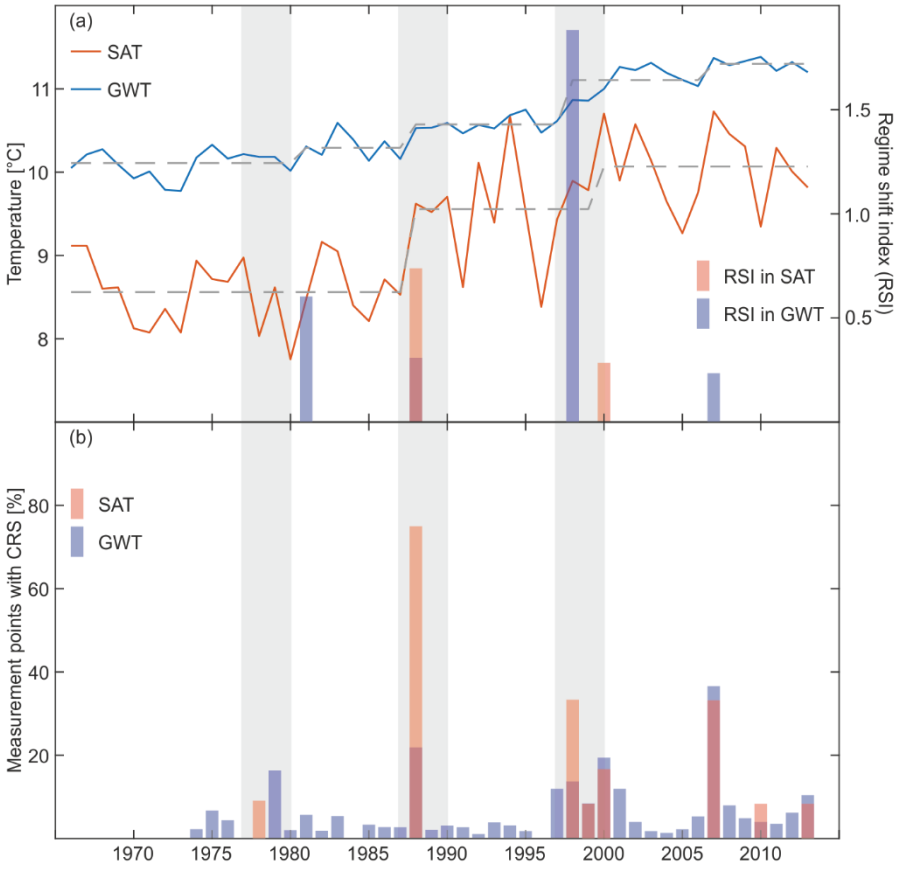


**Figure 6. (a) Median groundwater temperature (blue) and surface air temperature (red) of all wells or rather weather stations as well as the corresponding climate regime shifts (CRS) in form of the regime shift index (RSI). (b) Percentage of measurement points in GWT (blue) and SAT (red) that show a CRS in each year. The analysis of global temperatures data indicates a regime shift at the end of the 70s, the 80s and the 90s which are shown here in as grey bars.**