

Response to Reviewer #2 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:

Thank you for the opportunity to review “Recent trends of groundwater temperatures in Austria” by Benz, Bayer, Winkler, and Blum. I enjoyed reading the manuscript and appreciate the work it represents. I have outlined my specific primary suggestions for improvement below. I’ve also included minor comments, along with typographical suggestions as requested by the Journal. Only the primary comments rise to the level of serious consideration and response. The authors should feel free to contact me if anything is unclear at rjhunt@usgs.gov.

Reply: Thank you very much for your kind words.

Rev #2 Specific/Primary Comments:

Rev #2, Comment # 1: Overall manuscript: It strikes me that a focus on annual air temperature misses a fundamental process important for this discussion. The temperature of the groundwater system reflects the temperature of groundwater recharge. Groundwater recharge, however, is variable over time, thus annual temperature changes are likely too coarse to capture the temperature effects of inter-annual recharge process. That is: snowmelt recharge will be near 0 degrees C; rain-derived recharge will be warmer. Perhaps there is a shift in recharge from less snowmelt to warmer rain sources that is driven by air temperatures. A groundwater recharge approach means that the simple relation of air temperature to groundwater temperature is more indirect, and this additional “noise” to the signal is perhaps why the correlations are not higher.

Reply: We agree that a more detailed consideration of the annually changing recharge temperature would improve our general understanding of the vertical heat transport process in the unsaturated zone. However, such data is not available to us and previous studies have indicated this to be of minor importance, and thus it is does not included in our analysis. Further discussion of recharge processes, citing relevant sources, was added to the introduction in lines 44ff:

“When dynamic groundwater flow conditions exist, then advective heat transport can substantially affect the thermal regime in the subsurface [...]. Additionally, recharge processes, including snowmelt and rain-derived recharge, might impact the thermal regime of the shallow subsurface. Previous studies, however, indicate that in many cases their influence can be neglected. Ferguson and Woodbury (2005) and Bense and

Kurylyk (2017) demonstrated that it is possible to estimate groundwater recharge by using temperature-depth profiles based on the common assumption that the mean annual groundwater recharge temperature is equal to the mean annual surface air temperature. Menberg et al. (2014) showed in their study that the contribution of snowmelt-induced recharge with low temperature is minor in comparison to the overall recharge. Finally, Molina-Giraldo et al. (2011) investigated the impact of seasonal temperature signals into an aquifer upon bank infiltration including also varying groundwater recharge temperatures. They showed that the convective heat transfer by groundwater recharge compared to conduction through the unsaturated zone and convection within the aquifer is of minor impact. Still, the interplay of long-term climate variations, land use change and groundwater produces a complex transient system, which is difficult if not impossible to accurately understand based on a few borehole measurements.”

Rev #2, Comment # 2: Section Groundwater Temperature/Figure S1: Similar to comment #1, groundwater basins have a residence time, with multiple ages and potential lags. There is an assumption that all groundwater reflects current air temperatures (e.g., line 221) but this may not be the case. Given the importance of other factors such as residence time, and the unsaturated zone buffering that dampens the climatic drivers, it seems worthwhile to include well statistics relating to:

- Depth to water table
- Well open interval
- Distance the well’s open interval is below land surface
- Distance the well’s open interval is below the water table
- Estimated position in the groundwater flow system (e.g., uppermost, middle, discharge; near groundwater divide versus near flow system end; urban versus rural agriculture versus forest; high elevation versus low elevation)

Reply: We agree and have now included a discussion of temperature measurement depth in the chapters discussing correlation and linear temperature change. All wells are observation wells and are open all the way through. Unfortunately there is no information on the position in the groundwater flow system. Depth to the water table is also monitored in all wells, but currently not available to us.

The chapter discussing correlation between SAT and GWT now includes a discussion of measurement depth (lines 219ff): “*Additionally, measurement depths of GWT can have an impact on the correlation between SAT and GWT. While it is generally assumed that a measurement depth closer to the surface results in a better correlation with SAT as there is less of a shift between both datasets, this is only the case for some of the here analysed locations such as Villach (Figure S4a). In contrast, correlation increases with GWT measurement depth for other locations such as the one in Graz. This might be related to local underground heat sources such as sewage systems impacting GWT near the surface more than temperatures at greater depth. However, as the depth of the wells analysed here varies only slightly, no definite conclusions can*

be drawn without further inspection of specific cases.”

Additionally, information on the GWT measurement depth for all wells next to weather stations is now included in Table 1.

Table 1. Correlation coefficient and corresponding p-value between spatial median SAT and spatial median GWT for all analysed SAT locations, and additional information.

Location	Number of wells	Measurement depth GWT [m below surface]	Number of weather stations	Spearman correlation	p-value	Population ¹
Linz	1	10	1	-0.31	10 ⁻¹	192,000
Feldkirch	6	4 to 17	1	0.19	10 ⁻¹	31,000
Innsbruck	2	10	2	0.37	10 ⁻¹	123,000
Vienna	1	12	1	0.41	10 ⁻²	1,740,000
Zeltweg	2	6 to 7	1	0.48	10 ⁻³	7,000
Wiener Neustadt	2	9 to 20	1	0.51	10 ⁻⁴	42,000
Bregenz	6	4 to 10	1	0.52	10 ⁻³	28,000
Tulln an der Donau	1	7	1	0.54	10 ⁻²	15,000
Eisenstadt	2	4 to 5	1	0.67	10 ⁻⁴	13,000
Graz	9	4 to 12	1	0.73	10 ⁻⁸	266,000
Villach	17	3 to 11	1	0.80	10 ⁻¹¹	60,000

¹ Register-based Labour Market Statistics 2014, municipality level (Statistik Austria).

The following paragraph was added to the chapter on linear temperature change (lines 259ff): *“Temperature change decreases slightly with GWT measurement depth by approximately 0.015 K per 10 years per meter (Fig. S4b). This relationship can be related to deeper temperatures corresponding to earlier temperatures, when temperature increase was less severe. However, because the vast majority of temperatures are monitored at a depth of less than 15 m and show a high variability in linear temperature change, this number must be taken with caution. R² of the fit is only 0.02 and RMSE is 0.4 K.”*

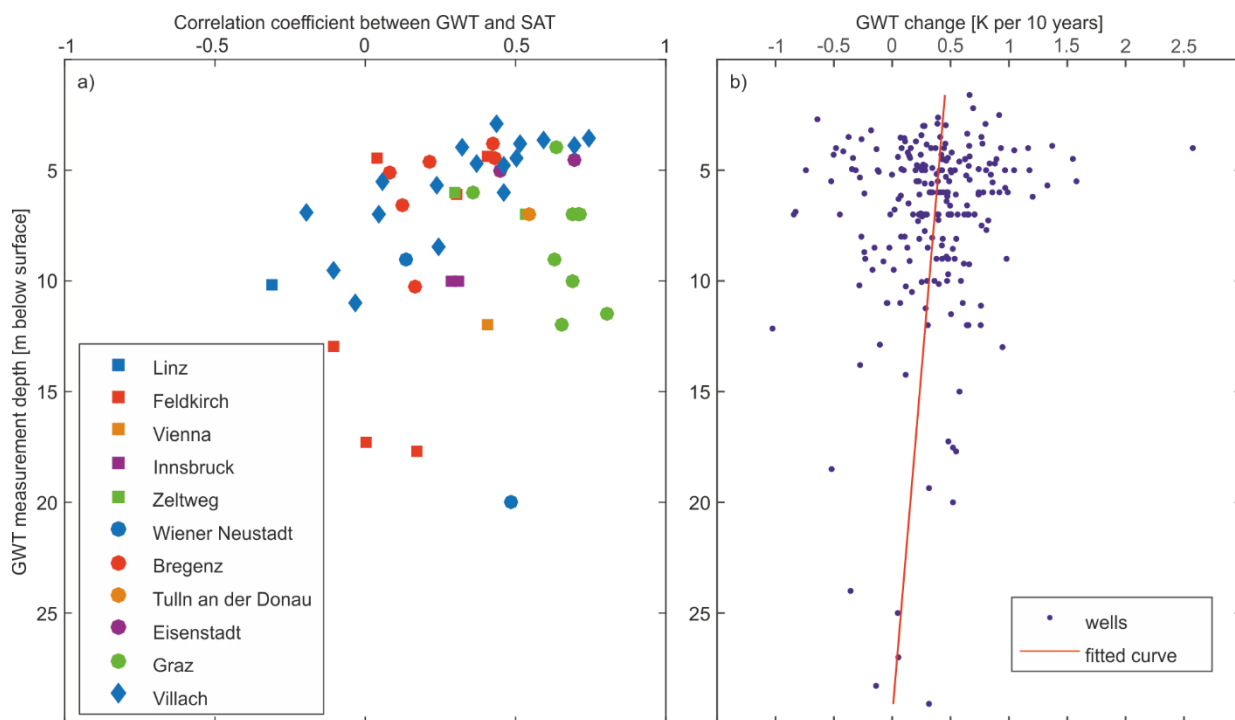


Figure S4: a) Influence of measurement depth on Spearman correlation between GWT and nearby SAT measurements. Shown are results for all wells depicted in Fig. S3. b) Influence of measurement depth on observed change in temperature. The best fit implies a linear temperature change of 0.48 K per 10 years for a depth of 0 m and a decrease in temperature change by 0.015 K per 10 years for each additional meter between measurement point and surface.

Rev #2, Comment # 3: Lines 104-106: It seems that only focusing on annual averages may limit the applicability of the insights. For example, for cold water fisheries it is usually the temperatures in the late summer – late fall that are important.

Reply: We agree. However, in order to keep the data analyzed in this study consistent, we would like follow advice from Rev. #1 and focus solely on one set of temperature data (annual or monthly means). Because the climate regime shift analysis works best with annual mean data, we prefer working with this. A more detailed analysis of seasonal variation would require an extensive investigation of the data and this is beyond the scope of this study.

Rev #2, Comment # 4: Figure 2: The shaded area and short-duration blue line dipping below $y=0.0$ is interesting – can you say something about what conditions would cause the GWT to be inversely correlated with SATs?

Reply: Fig 2a) shows correlation between different wells on the y-axis and the distance between those wells on the x-axis. All correlation coefficients close to or below zero all have a p-value of close to one, meaning these wells do not correlate. It is likely that at least one well in each of these pairs is influenced by other, local heat sources and not by surface temperatures. A sentence was added to the manuscript to clarify this (lines 183ff): “*For the weather station, each individual pair is shown by a red point, for GWTs, as there are many possible pairs of wells, the line gives the moving median (± 25 km) correlation of all pairs at the corresponding distances. The inner 90 percentiles are shown in grey, and correlation coefficients close to or below zero are*

determined for several pairs of wells. However, here p-values are generally also close to one and GWTs do not correlate. This is most likely due to local heat sources impacting at least on well in these pairs.”

Rev #2, Comment # 5: Figure 3: It appears that the annual averaging is hiding important relations. That is, if surface air temperature (SAT) is the driver of groundwater temperature, it does not follow that the summary groundwater system temperatures would be warmer than the SATs at every location. Is it not likely winter periods skew the annual SAT, but the groundwater system is buffered from these colder temperatures? Therefore, might it be more insightful to look at SATs during non-winter conditions?

Reply: Yes, you are right; the annual average is simplifying the complicated short-term relationship between above ground and subsurface temperatures. One example, as you mention, are cold air temperatures in winter: several studies have shown that snow cover insulates groundwater temperatures during that time and annual mean GWT are therefore warmer than annual mean above ground temperatures. Still, as discussed in response to your comment #3 we would like work with annual mean data only for this study.

Rev #2, Comment # 6: Lines 176-177: For this sentence: “This indicates that GWTs are often influenced by local causes and not necessarily solely by surface temperatures.”, the correlation is between the weather station that is measuring surface temperatures correct? Then wouldn’t it follow that the correlation is between groundwater temperatures and local SATs?

Reply: Correct. Surface temperature was the wrong term and therefore this was changed to “*local SATs*” (line 202).

Rev #2, Comment # 7: Lines 220-225: Can you provide reasons (and citations for the interested reader) for why there are different levels of change with land use?

Reply: We meant to say that groundwater temperatures do not change significantly differently for the different land cover classes. We changed the paragraph to make this message clearer: “*There appears to be no significant influence of land cover on the observed temperature change (Fig. S2c). Median temperature change is approximately 0.4 ± 0.4 K per 10 years for groundwater under artificial surfaces and forest areas, and 0.3 ± 0.5 K per 10 years under cultivated areas.*” (Lines 256ff).

Rev #2, Comment # 8: Line 248: Did the hot spring suddenly appear or was it always there and something else changed? It was not apparent to me in Figure 5 what is the hot spring effect that I should be seeing in IIB and IIC in Figure 5. It does seem these outlier wells that have known atypical perturbations make the narrative hard to follow because they pop up every time a point is being made, and cause two sets of statistics to be reported – one with them and one without them (e.g., Villach wells, lines 265-359, wells near the Drau River). Because you know they are not representative of the larger scale climate driver would it not

be clearer to just state this in the beginning and say you are not going to include them when reporting the subsequent statistics?

Reply: Yes, you are right, we have decided to exclude the two wells that are dominated by these hot springs and now only 227 wells are analyzed in total. The wells nearby and not impacted by the hot springs are still included.

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. But some changes in the hydrogeological conditions must have happened when additional wells started to pick up the signal. However, we could not identify any concrete evidence of these changes.

Rev #2, Comment # 9: Please describe briefly the technique of Menberg et al. (2014) and define “regime shift index” used to save the reader from having to find it.

Reply: The paragraph describing the method was extended accordingly to include a brief description (lines 158ff): “... *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 [Climate Regime Shift].*”

Rev #2, Comment # 10: Lines 296-297 and 313-314: There are other statistical tests that beyond linear and regime shift methods (such as autoregressive integrated moving average techniques). Were any of these tried? The difference in RMSE is reported here is so small that it seems a stretch to say one performs superior than the other, and maybe other methods would perform better.

Reply: No, we did not apply any ARIMA models yet. However, this is a very interesting thought that we will consider it in the future.

Rev #2, Comment # 11: Is there something we can learn about the fact that nationwide correlation is higher than any of the individual weather station / well combinations? Would it be worth including a sentence in the manuscript pointing out that if researchers simply used the nationwide relation they could potentially hurt their ability to solve their more local problem?

Reply: Yes. This is definitely the main message here that groundwater temperatures are dominated by local features. While a national average can give us important information on large scale trends and problems, local temperatures might behave differently. A sentence was therefore added to the conclusion: “*This reveals the extent in which groundwater temperatures are dominated by local events, groundwater flow, and the thermal properties of the surrounding. When solving local problems we can*”

therefore not recommend relying on average relationships valid on a nation scale.”
(Lines 334ff).

Rev #2, Minor Comments / Technical Corrections:

Rev #2, Comment # 12: Line 19: It would be nice to relate the locations to features transferable to other parts of the world (e.g., high topographic relief/mountainous versus less topographic relief/less mountainous).

Reply: We agree, a sentence discussing the need for future work transferring these results to other regions has been added to the conclusion (lines 343ff): *“However, further research dedicated to other climate parameters such as permafrost and snowfall is necessary to validate these findings. Additionally, our observations made in Austria should be transferred to similar regions in the world testing the transferability of the presented results.”*

Rev #2, Comment # 13: Lines 47-67: Bill Selbig used a regression of historical groundwater and air temperatures for the purpose of forecasting what future groundwater temperatures would be given expected changes calculated by GCMs. Not sure if your work would benefit from an application of how groundwater temperature trends influences societally relevant endpoints such as trout. There are others as well, but this work can be found in: Hunt, R.J., Walker, J.F., Selbig, W.R., Westenbroek, S.M, and Regan, R.S., 2013, Simulation of Climate-Change Effects on Streamflow, Lake Water Budgets, and Stream Temperature Using GSFLOW and SNTMP, Trout Lake Watershed, Wisconsin: U.S. Geological Survey Scientific-Investigations Report 2013-5159, 118 p., <http://pubs.usgs.gov/sir/2013/5159/>.

Reply: Thank you for your suggestion. This study is very interesting. Thus, we added this reference (line 60).

Rev #2, Comment # 14: Figure 1: the dashed line is not defined in the figure or in the caption.

Reply: The dashed line represented the 95th percentile of SAT. However, following the suggestion by Rev #1, Comment # 6, a transparent color is now used to show the inner percentiles.

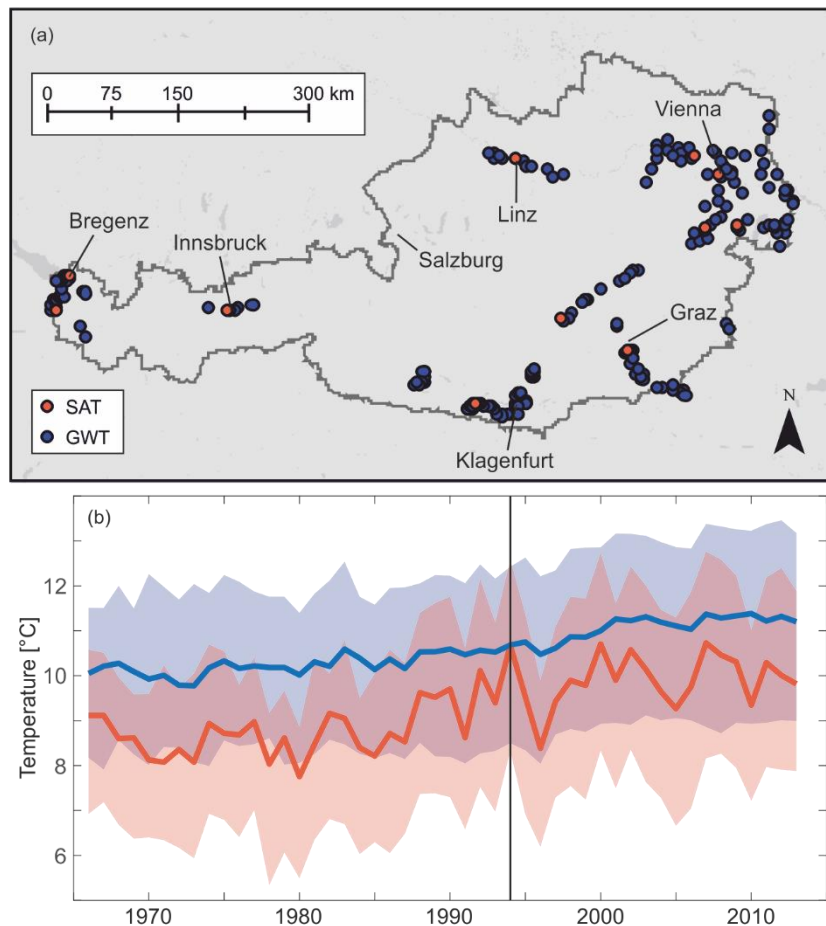


Figure 1. (a) Location of all analysed groundwater temperature (GWT - 227 wells) and surface air temperature (SAT - 12 weather stations) measurement points; (b) temporal evolution of the spatial median, annual mean temperatures for groundwater (blue) and air (red). The inner 90 percentiles are marked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

Rev #2, Comment # 15: Lines 158-159: It would be clearer to state exactly what is meant when stating “...the distance in the north-south direction of two wells has more influence on the correlation...” As written the influence can be augmenting (more correlation) or degrading (less correlation).

Reply: We agree. The sentence was changed to: “*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.*” (Lines 182ff).

Rev #2, Comment # 16: Line 176: I don’t think figure 3 shows “pairs of wells” but wells within 5 km of a weather station.

Reply: Correct. The sentence was therefore changed to “*Here correlations vary greatly and Spearman correlation coefficients are < 0.5 for about half of all wells within 5 km of a weather station.*” (Lines 199f).

Rev #2, Comment # 17: Lines 205-206: It seems Vienna may not be the best example to state as it only has one well included in its calculation of correlation.

Reply: We agree, another example was added: *“For example, both locations Graz (population of more than 250,000) and Eisenstadt (population of 13,000) have similar correlation coefficients despite their 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).”* (Lines 238ff).

Rev #2, Comment # 18: Line 240-241: I am not sure I followed the sentence construction – what is meant by “...but in one sudden drop or rather rise in temperatures.”?

Reply: Sentence was change for clarification: *“In general, most of the extreme changes in temperature appear to be linked to local causes and do not happen gradually, but rather rapidly over the short time span of one or two years.”* (Lines 280f).

Rev #2, Comment # 19: Lines 222-224: In the beginning of this paragraph the topic is rate of change and then in these lines it is absolute change over a period, then the next paragraph goes back to rate of change. Perhaps better to start out with the differences in absolute temperatures then stay with changes in temperature. Also, the period 1990-2012 stated in these lines is not the same as reported in the caption of Figure S2 (01/1994 – 12/2013).

Reply: It appears that the wording of the paragraph was misleading. The last part discussed not temperature change over time, but the absolute difference between different time series. This discussion is completely independent of our main results and was now moved to the Materials and Methods section when introducing the groundwater dataset (lines 124ff):

“Following the CORINE Land Cover (CLC) data from 2012 (Fig. S2a), 45 % of all wells are under artificial surfaces, 46 % under agricultural areas, and 9 % under forest following the 100 m × 100 m classification. In addition, CLC from 1990 was consulted, however, no land cover changes near any of the analysed wells are observed. Overall, for the time period 1994 – 2013 when all wells were monitored, absolute GWTs under artificial surfaces are on average 1.5 ± 0.3 K warmer than GWTs under forest; GWTs under agricultural areas are on average 0.6 ± 0.2 K warmer than GWTs under forest (Fig. S2b). This validates previous findings by Benz et al. (2017b) for GWTs in Germany, who identified even larger differences of up to 3 K between the individual land cover classes.”

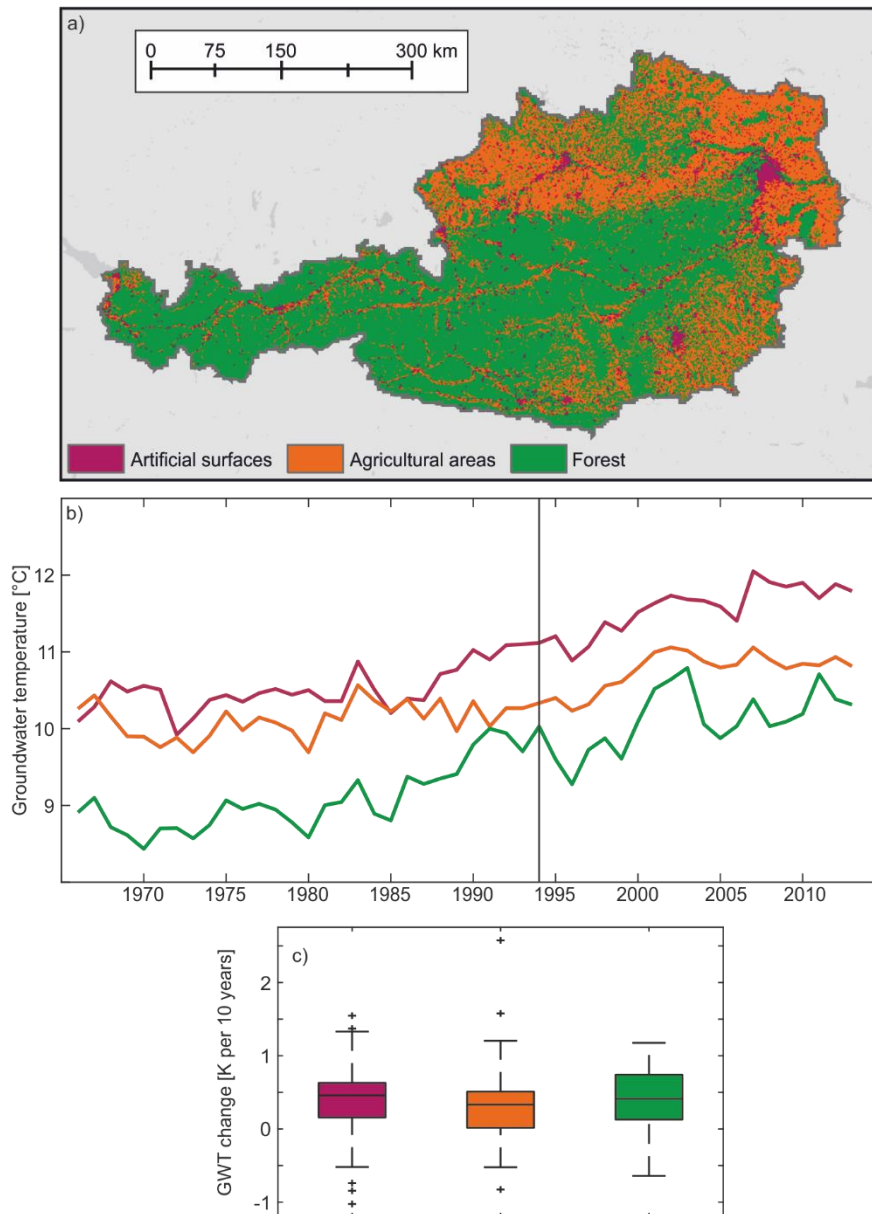


Figure S2. a) Corine Land Cover 2012 of Austria. None of the analyzed wells and weather stations experienced a land cover change since 1990. b) Spatial median GWTs for each of the individual land cover classes. All wells are monitored since at least 1994. c) Relationship between land cover and groundwater temperature (GWT) change between 1994 and 2013. There appears to be no significant influence.

Rev #2, Comment # 20: Line 239: Here is perhaps an opportunity to reinforce the importance of including groundwater flow when trying to interpret groundwater temperature (as opposed to dry borehole temperatures mentioned in the introduction). Same with line 304 in the Conclusions.

Reply: We agree. The following sentence was added (lines 278ff): *“These wells seem to be affected by the new drinking water supply (four wells with a total pumping rate of about 100 l/s) located about 1 km in the south. This demonstrates the importance of including groundwater flow when trying to interpret groundwater temperature.”* and in lines 331f: *“This reveals the extent in which groundwater temperatures are dominated by local events, groundwater flow, and the thermal properties of the surrounding.”*

Rev #2, Comment # 21: Line 240: Are there other cases of extreme changes not discussed in the text?

Reply: Interesting point. We checked this issue and found 57 wells with extreme changes (> 400% of average change per year). However, a more detailed analysis is beyond the scope of this study.

Rev #2, Comment # 22: Line 243: The word “extend” should be “extent”.

Reply: Word was changed (line 283)

Rev #2, Comment # 23: Line 236-249: The discussion starts with the <5% cases then includes the >95% then concludes again with <5%.

Reply: The chapter was slightly restructured, it now starts with the <5% and ends with the >95% cases (lines 266ff).

Rev #2, Comment # 24: Line 261: My PDF had an odd “extend” tacked onto the end of the line.

Reply: We checked this issue, however cannot find any anomalies in our PDF version.

Rev #2, Comment # 25: Line 265-266: I think this sentence is less clear than it could be. I think the point is that if SATs are the driver of GWTs the former cannot lag behind the latter. The fact that GWT changes precede the SAT driver suggests this method does not have the resolution to determine short lags between SATs and GWTs.

Reply: Thanks for your suggestion. The sentence was changed: “... , *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 #2, Comment # 26: Line 303: “instalment” should be “installment”, or even better, “installation”

Reply: Changed (line 333).

Rev #2, Comment # 27: Figure 6: I am not sure what to make of the checkerboard bar around 2006.

Reply: In 2007 bars for SAT and GWT were at the exact same height. However, this changed when two wells in Villach were excluded (your comment #8). Still, we changed the bars to be transparent in order to clarify this issue. In addition, a legend was added to the figure following Rev #1, Comment #32.

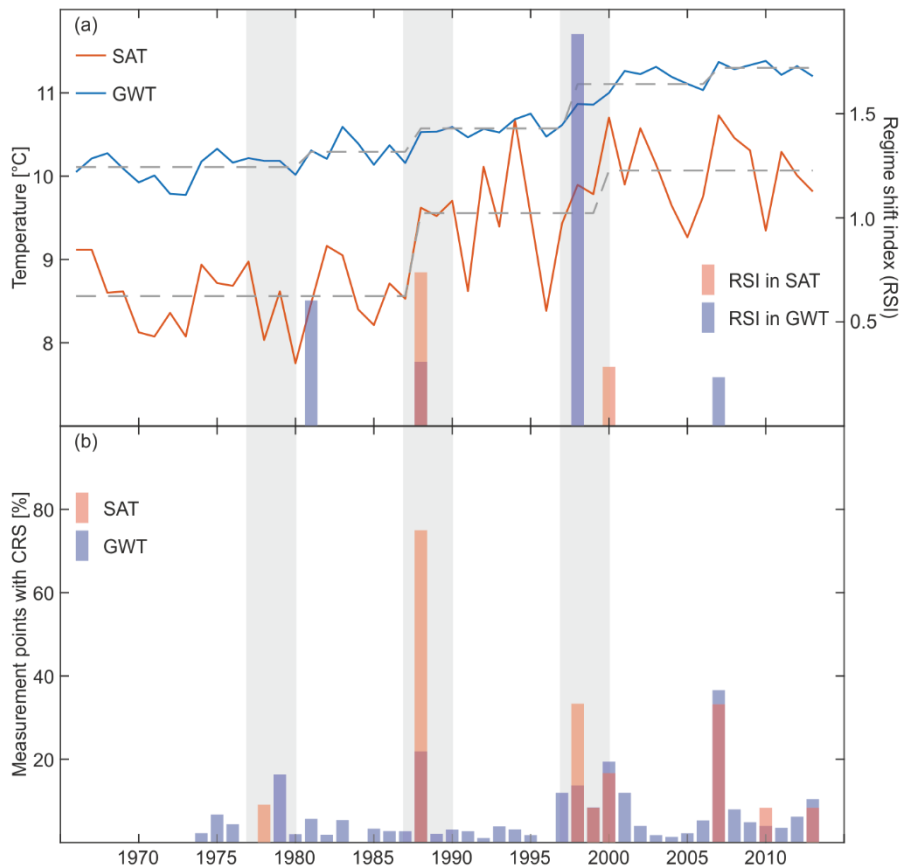


Figure 6. (a) Median groundwater temperature (blue) and surface air temperature (red) of all wells or rather weather stations as well as the corresponding climate regime shifts (CRS) in form of the regime shift index (RSI). (b) Percentage of measurement points in GWT (blue) and SAT (red) that show a 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.

Rev #2, Comment # 28: Figure S5: Perhaps add a vertical line to the figure to help the reader identify the exact date of the July 2007 flood.

Reply: Following a suggestion by Reviewer #1 all references to the flood were deleted. There is currently not enough information available (e.g. extend of the flooding) to proof our hypotheses that flood and drop in temperatures are related. Still, the year in question was marked in Figure S5 to make the figure clearer.

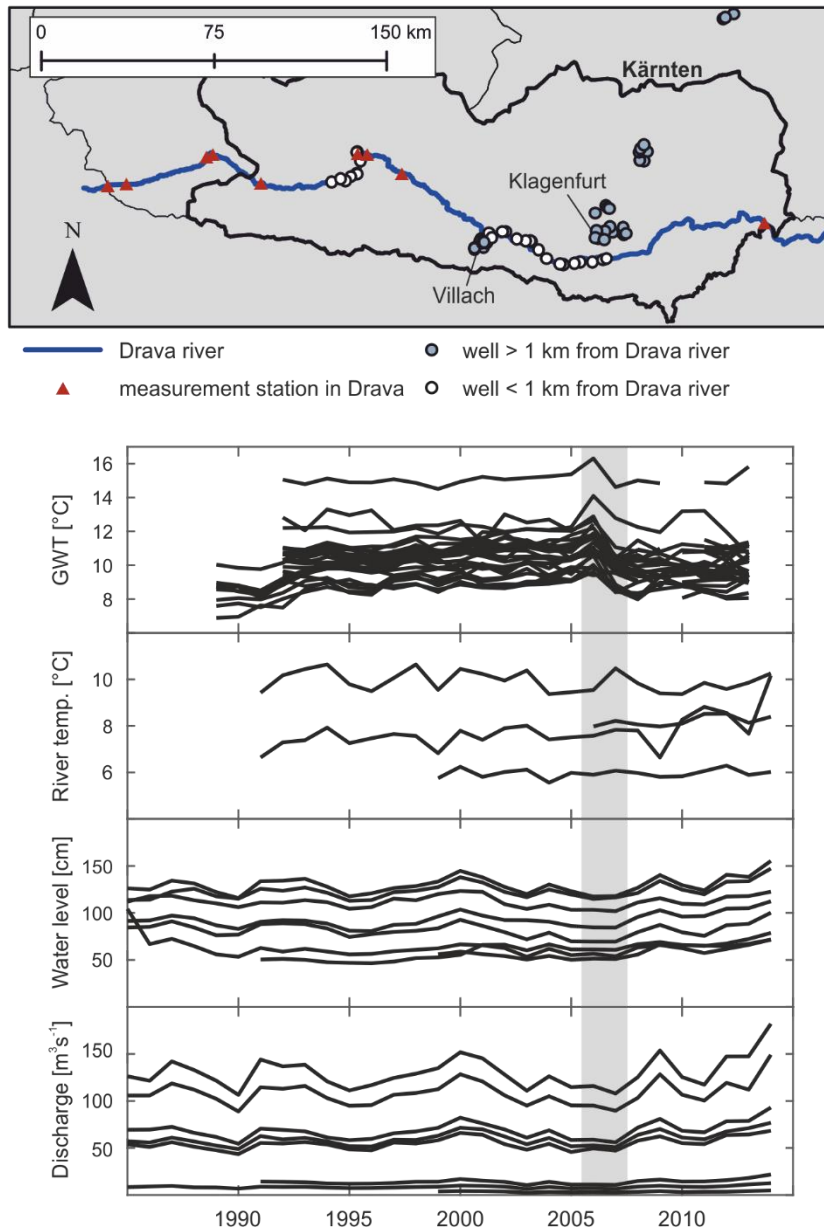


Figure S5. Location of the Drava river, the groundwater monitoring wells around it and measurement stations within the river (EHYD, 2017). Also shown is the groundwater time series of all wells within 1 km of the river and all measured river parameters. While GWTs show a sudden drop in 2007 (marked in red), observed river parameters give no indication of an abnormal event around that time.