Author Response to reviewer 1:

Thank you very much for the very detailed and helpful review. In our response we will address to your suggestions on the manuscript structure (GENERAL COMMENTS) and your SPECIFIC COMMENTS. Regarding your TECHNICAL COMMENTS the manuscript will be checked again by a professional editor considering your comments.

Reviewer's comment	**Authors' comment**	**Suggestion of changes in the manuscript**
GENERAL COMMENTS		
1. Appendices A and B are more suited to be included as part of the Supplementary material, forming two separate technical reports, given that the focus of the paper and of the journal is on the glmGUI package and not on the development of the case studies.	Your assessment is very valuable for us and will follow you suggestions	We will convert the Appendices A and B in 2 separate 2 Supplementary Materials
2.Appendices C and D should be moved to the main text, the former be-cause having to look up each time many pages forward in the paper to understand the meaning of model parameters makes understanding passages troublesome for people unfamiliar with GLM, the latter because the counterpart figure for Lake Ammersee (Fig.9) is already part of the main text. The paper would strongly benefit from reorganizing the material in the Appendices, receiving a more compact outlook.	We wanted to avoid to overload the manuscript with large tables figures. Figure D1 is erroneously the same as Fig. 9, it was supposed to be the same plot type for Lake Ammersee. As you stated above the manuscript's focus should remain on the glmGUI, the presentation of an second plot of the same type in the main part would not have a great benefit.	
SPECIAL COMMENTS		
1. Autocalibration routine- this is perhaps one of the most value features in the GUI as the calibration process can be difficult and time consuming. However, no information is given on how the calibration is actually performed, what are the objective functions, how is the parameter space sampled, what are the stopping	Example: If the user chooses to calibrate the model for the two parameters P1 (default value 1.0) and P2 (default value 2.0), she/he has to choose the upper and lower calibration limit by a percentage range. Let's say she/he wants to vary P1 by 10% only and P2 by 50%. → P1 will be in the range of 0.9 to 1.1 and P2 between 1.0 and 3.0.	

criteria, how flexible is the routine to user definitions, and other issues. Furthermore, how good is the calibration tool in relation to manual calibration? I would like to see a much more extensive description, testing and discussion of the calibration process.	Then the user chooses the resolution of the space between these ranges. Let's say the user wants to have only 4 samples in the percentage range plus the default value. So P1 will get the following values: 0.9, 0.95, 1.0, 1.05, 1.1 P2 gets: 1.0, 1.5, 2.0, 2.5, 3.0 Now all possible combinations of P1 and P2 are tested: 5 x 5 combinations So the GLM.exe is run 25 times by the glmGUI and the resulting RMSE of lake level and water temperature is saved to two CSV-files. So the user can see, which combination results in which RMSE. This process is already described in the manuscript as follows: "*The user can choose out of these parameters those which are to be included in the calibration process and define a percentage range, by which the upper and lower limit of every parameter is changed from the value in the glm2.nml-file. The resolution of the increase/decrease of the parameters within the defined limits can be set as well. According to these settings, model runs of GLM are executed with all possible combinations of the selected parameters ("brute-force"). The overall RMSE of the lake level or water temperature is calculated and saved for every parameter combination to a csv file, so the "best fit" is indicated*. "	

	- Objective function is RMSE (as described) - Sampling of the parameter space is described - There is no stopping criterium, as all possible combinations are calculated(as described) - The flexibility is given by selecting: 1) The set of parameters 2) The percentual ranges of upper/lower limits for each parameter 3) The resolution of the increase/decrease of the parameters within the defined limits Furthermore, saving the RMSE of **all parameter combinations** gives the experienced user the possibility to choose not the "best fit" parameter set, if the user thinks that one parameter of this set might have an unrealistic value. The comparison to a manual calibration is very subjective, as it depends on the user and her/his skill. We can just state, that the autocalibration function can simplify the calibration process, as the user does not have to re-run the GLM.exe for each parameter change manually. Additionally, using this tool can save time, e.g. when running on a server in advance or as a "background task".	
2. Sensitivity analysis- This too is very useful however there is insufficient information on how it is actually conducted. How is SI	Yes, we totally agree on that. We should have added the formula and the selection options for the sampling of the parameter space.	We will add the following at pg. 6 ln. 7: "*The widely used approach after Lenhart etal. (2002) is implemented in the GUI.*

calculated? Is the analysis conducted by changing one parameter (or variable) at a time or changing all at the same time? How is the parameter space sampled for the analysis? While some of the meteorological variables are included in the SA I would also expect shortwave and longwave radiation to be included as they can be difficult to measure accurately, especially the later.		SI= $\frac{(y_1-y_2)/y_0}{2\Delta x/x_0}$ *The Sensitivity Index (SI) is calculated for each selected parameter separately, as only one parameter is changed at a time. The parameter with the value x_0 is increased and decreased by Δx. The resulting outputs y_1 and y_2 (either water temperature, lake level or the respective RMSEs) are subtracted and normalized by the output y_0, which results from using the unchanged parameter value x_0. Δx can be set to four different values in the GUI (5%, 10%, 20%, 50%).*"
3. Along these lines, including quantifiable indices for the goodness of fit of the model to lake-based data is critical and the authors have included RMSE and MBE. I think the authors should include a range of indices which the user can select from when conducting the analysis.	Adding more quantifiable indices is relatively easy to implement in future Releases (V1.1).	
4. In the Lake Baratz lake level results (fig. 5) there is a period during which the fit between the model and lake data is not good in contrast to the other periods. I would like to see discussion of this and possible explanations. Similarly for Lake Ammersee. You mention the issue in lines 10-13 (pg 16) but don;t attempt to explain the discrepancy. I think the large discrepancies that are obvious on Fig. 8 need to be explained.	For Lake Baratz (Fig. 5)	We will add to description of these results in section 3.2 the following at pg 11, ln 5): "*The uncertainty related to the lake level at the beginning of 2014 could be derived from discharge events difficult to simulate. The basin in that period of the year is still in an intermediate status of soil moisture. Probably the model overestimated the discharge on the base of rainfall events in January 2014 which in reality did not produce a significant lake level variation (see Fig. 2 in Giadrossich et al., 2015).*

		Moreover, the missing data in these two months doesn't allow understanding well the process that shift the estimation of lake lever of about 20 cm higher. Notice that the two lines in Fig. 5 have the same trend, but they keep the error accumulated in January, for the whole year."
	For Lake Ammersee (Fig. 8):	We will insert to the presentation of the results (Section 3.4) pg16 ,L9: "No obvious explanation for these trend shifts could be found, despite a detailed investigation of the existing hydrological data was conducted. An impact of a highly complex groundwater inflow system is likely to have a key role in the water balance of the lake, which is not considered by the applied input data. Furthermore it cannot be excluded, that unknown alterations or errors in the observation system of the gauges cause these "turning points" as some of them correspond to flood events, which implied problems with the measurements."
5. Conclusions section- I think this section requires significant strengthening in order to better convey the key points. The way it is currently written does not touch on allthe important points and mentions issues that are not necessary.	We will rewrite this section to strengthening the key points of our goals.	
6. Input data for Lake Baratz- you mention a 5 month gap in met data (pg 20 lines12-	The entire preprocessing of the meteorological input including the dealing of the observation	To clarify we could add an indication after line 13 (pg 20):

13) how did you deal with this gap?	gap of the lake station is described in detail for each parameter in Section A2.	"*A detailed description of the source and required processing steps of the respective parameters is given in the section A2.*"
Fig. A1- the air temperature data from Fertila station does not look like continuous data. What type of data were these?	The data at Fertilia station were observed with a precision of 1 degree, which cause the distribution in the graph.	We will to section A2.1 "*Values at Fertilia station were available in a precision of 1 degree*".
Fig. A6- Isn't it possible that the unique water transparency event in 2017 affected the relationship shown in this figure and that a different equation is required for that period? Please discuss.	We rephrased the paragraph form line 1 to line 5 at page 9 because it was not clear. There is not a unique water transparency event in 2017 that affected the relationship. The sensitivity of Kw is low for the whole period and doesn't change significantly, giving an average light extinction coefficient value Kw = 0.57. Thus, we considered 0.57 to be representative of the whole period. It has been obtained dividing the Secchi-disk constant ranging from a minimum value of 1.44 to a maximum of 1.80 (Hornung, 2002; Holmes, 1975; Chapra, 1997) divided by Secchi-disk depth ranging from 2.50 to 3 meters. In these cases, the Kw values range between 0.48 and 0.72. The value of 0.57 has been adopted, because lake had a higher depth I the period between 13.07.2011 to 31.12.2016. If we would apply the constant = 1.44 and secchi-depth=2.5, and constant 1.7 (as suggested by Poole and Atkins, 1929), secchi-depth = 3, we will obtain the same value of 0.57.	NEW PARAGRAPH: line1-5 P9 "*The simulation period for Lake Baratz is determined to be 13.07.2011 to 31.12.2016. We assume the light extinction coefficient value Kw = 0.57 is representative of the whole study period. Kw have been calculated dividing the Secchi-disk constant assumed to have a minimum value of 1.44 (Hornung, 2002;Holmes, 1975; Chapra, 1997, ranging from 1.44 and 1.80) divided by the mean Secchi-disk depth 2.50 meters (data from June 2016 to June 2017). Similar value can be obtained considering Secchi-disk depth of 3 meters (assumed when the lake had a higher water level) and Secchi-disk constant of 1.70 (Poole and Atkins, 1929).*" **REFEENCES to be added:** Holmes, R. W.: The secchi disc depth in turbid coastal water. Limnology and Oceanography 15, 688–694,1975. Chapra, S. C.: Surface Water-Quality Modeling, international edn. McGraw-Hill, 1997.

		Hornung, R.: Numerical Modelling of Stratification in Lake Constance with the 1-D hydrodynamic model DYRESM, http://dx.doi.org/10.18419/opus-163, 2002 Poole, H. H.andAtkins,W. R. G.,Photo-electric measurements of submarine illumination throughout the year.Journal of the Marine Biological Association of theUnited Kingdom,16,297–324, 1929.
7. Section A4.2 field data- in the main text you mention that mixing occurs in the winter however here you state that you assume isothermal conditions from 24.09.13. Do isothermal conditions develop as early as September?	Isothermal conditions establish usually during early autumn. In this year the lake had already a very small temperature gradient before the station was out of action.	We will add at the end of the paragraph (pg. 29, line 17): "*… and the vertical temperature gradient of 0.91° C on 24.09.2014 was already low indicating no stable thermal stratification.* "
8. English- the MS needs to be edited by a native English speaker or professional editor. Currently there are many sections/sentences that need rewriting.	The entire manuscript will be checked again by a professional editor.	
9. The shutdown button in the GUI is in German and not English. Better to have it in English like the rest of the GUI.	Button label depends on language settings of R. One solution could be to name the button with an unique string. gbutton ("Close window", …)	
10. Pg 5 ln 18: erroneously - what do you mean?		We will revise the sentence(s): "*…water temperature plots taking into account the range of temperatures and also erroneously the range of lake depth. This method is adopted in glmGUI, but discarding the consideration of the lake depth, and the temperature range…*"

11. pg 13 ln 6- outflow or inflow?	Yes, we are talking about the outflow	We reformulate: "*The lake has a catchment area of about 994 km² and its outflow in the north (Stegen gauge station).*"
12. Pg 16 ln 4- the RMSE reduced significantly- under which conditions? Why?	The reason for the reduction of the RMSE is due to the application of the above mentioned inflow factors.	
13. Fig 9- isn;t the lake 83 m deep? If so, why is only 9 m shown?	Thank you for your meticulous review! This is erroneously the Figure for Lake Baratz and will be replace by the right plot for Lake Ammersee	Figure shows erroneously results for Lake Baratz and will be replaced by the plot for Lake Ammersee
14. Pg 19 ln11- "This includes a data quality assessment..."- That is not the case. The GUI allows visualization but does not include, as far as I understand, QA tools.		We reformulate here:" *The GUI includes tools to check the quality of the input data. This includes the option of a visual detection of errors, missing values and plausibility.*"
15. Pg 19 ln 15- sentence not clear.		We would rewrite this sentence in the following manner: "*The GUI allows a high level of interoperability due to the option of combining with other operating systems. Furthermore, we designed the software with the aim of a high flexibility for the application of other scenarios, for various study areas or with diverse time steps.*"
16. Pg 20 ln 17- R2 between which two sets of data?	The information in brackets is misleading here and can just be removed	Information in brackets will be removed
17. Pg 22 ln 7- why correct only data after 21.6.2016 and not the earlier data if they are much lower than Fertilia station	Fertilia station is quite in a distance to the lake and the data measured at the lake station before June 2016 are reliable, which is confirmed by . observations at the closer Grifone station (systematically lower for the available period until 2014). Hence only data after June 2016 were corrected	We will add sentence at pg21 ln 8: "*Observations at Grifone station are in the range of measurements taken at the lake station befor e21.06.2016 and hence these data were taken as reference for correction.*"