

Review of gmdd-5-3993-2012

“A one-dimensional model intercomparison study of thermal regime of a shallow turbid mid-latitude lake ”

General comments

This manuscript is a useful and important intercomparison of different lake models. It provides some insight into their utility for weather and climate modelling based on their performance at a midlatitude shallow lake, as compared with research observations. This work helps the modelling community both to distinguish between models, and to anticipate the performance characteristics of any one model.

I believe that there are some areas where the manuscript requires improvement in the presentation of results, the focusing of conclusions, and in some particular points of analysis. These are detailed in the next section.

Specific comments

sections 1–4

The abstract and sections 1 (Introduction), 2 (Lake models), 3 (Observations) and 4 (Experimental set-up) are in general clear and fit for the purpose, except that I am curious as to what precipitation data are used for (p3998, line 28)?

section 5.1

The presentation and discussion of results begins in section 5.1.1. The division of the timeseries into two periods should be more clearly identified with the shorthand labels of “summer” and “autumn” which are also used later.

The data in figure 2 are difficult to distinguish, and perhaps could be expanded in the supplementary material.

Please provide a reference for the observations at midlatitude lakes during the warm period (p4008, line 1).

There is an unsatisfactory overlap in the presentation of data between sections 5.1.1 and 5.1.2. The data for the FLake model with and without the bottom-sediment effect are presented in figure 3 (and subsequent graphs of this type), but not those data for the other two models for which deactivation of the bottom-sediment effect is tested in section 5.1.2. The presentation style should be unified to show all three models in this way, or none. In my opinion the inclusion of data from both FLake runs in the figures is a useful addition and I would prefer to see all three models thus displayed. Perhaps this could be achieved e.g. by grouping results for each model together and shading the deactivated-sediment results.

Minor related points are that the graph labels of “FLake active sediments” and “FLake passive sediments” do not seem to be fully explained anywhere, and change to other terminology in figure 6. Also that the entry in table 3 for FLake, period 1 shows 1.13 but the central panel of figure 3 shows 1.12, and if I am right in assuming that these numbers should be the same then this needs to be checked.

Statistical quantities are defined in the last paragraph of p4005 and I assume that the bias (referred to here as DM) follows the convention of being modelled variable *minus* observed variable.

Throughout sections 5.1.1 and 5.1.2 there are statements about the importance of the bottom-sediment effect:

p4006, line 9, “this suggests that the differences in DMs may be largely due to the treatment of water-sediments exchange and surface flux schemes in these models”

p4006, line 20, “this result hints that, in summer when stratification is typically stable, heat exchange with sediments is not strong”

p4007, line 5, “that again hints to the lack of heat supply from bottom sediments in these models”

p4008, line 4 “these DM changes are relatively small, and this supports the statement that lake-

bottom heat exchange typically is of minor importance for the thermal regime of upper water layers during summer”

However, section 5.1.2 could do with some sort of summing up at the end indicating exactly the conclusions that are drawn from this analysis. Are the effects regime dependent? Model dependent? Is Flake less sensitive to the deactivation of the bottom sediment than the other models, and if so can the authors speculate as to why?

In section 5.1.3, figure 4 is useful in interpreting the data shown in figure 5. Attention should be drawn to figure 4 earlier on in this section.

section 5.2

At the bottom of p4012, the divergence of the LAKE model from other models in late autumn is attributed to the large thickness (10m) of the soil layer underneath this model. This begs the question as to how water-sediment heat exchange is parametrized in this model. Is heat exchanged with the entire mass of sediment, or only through the interface by some parametrization of a diffusive process? If the latter, then the presence of a very deep soil layer ought not to significantly alter results compared to having a layer of only moderate depth. If the former, then this would seem to be a sub-optimal feature of the model which should be highlighted.

section 5.3

Figures 7 and 8 appear to be identical and I assume that this is a mistake (the alternative being that the lake models are all perfect, which is unlikely!). The discussion of these results in the text would suggest that the proper data are similar, or that the authors have not noticed that the graphs are identical. However, if the correct form of one figure is actually missing, the assessment of the results by the reader is not possible.

The possible underestimation of fluxes by the eddy-covariance measurements is discussed on p4016 as an explanation of the positive bias in the turbulent-flux models. While this may be a contributing factor, other land-based studies comparing surface-flux schemes with such measurements do not describe such a clear systematic bias (the most recent example I can find is Flerchinger et al., J.Hydromet. 2012 volume 13). The Nordbo et al. (2011) reference which is cited indicates that the low-frequency deficit is possibly caused by a lake-forest internal boundary layer and may be minimised in wind directions of maximal open-water fetch, however the present study does not observe any substantial reduction of bias by such a fetch restriction (p4014, lines 14–17). Nordbo et al. (2011, section 2.7) also indicate that “tube attenuation” in their gas sampling system is the dominating cause of LH-flux underestimation, whereas the LI-7500 in the present study is an open-path gas analyzer which does not suffer from this effect. Is it not also the case that Nordbo et al. (2011) obtain their energy balance residual *after* correcting their flux measurements?

There are ways of estimating or correcting eddy-covariance flux underestimation due to the spectral cut-off. It would be useful if the likely mean error in the covariance measurements could be estimated independently, to see if this is similar to the observed bias, or to the 20–30 Wm^{-2} underestimate proposed on p4016, line 22. Without this, the superiority of the flux models over the flux measurements is debatable. It would also be helpful to estimate the random error in the flux measurements for comparison with the RMSE values presented.

On p4016 the discussion of the term R representing advective and diffusive contributions to the heat budget is slightly speculative. The possible existence of seasonal circulations and their potential for providing a net heat flux is discounted, although one could imagine possible scenarios where this occurs, particularly for a marginal location like that of the observations. It is also unclear what the flux rate of water through this lake is. It may not have a major inlet, but it has an outlet. Is this gauged, or is there a gauge further downstream that may be used to estimate the outflow? What is the ratio of the lake area to the lake catchment area? Either of these may help provide a quantitative estimate of advective strength.

Technical corrections

I suggest some corrections as follows.

p3996, line 3: it would be better to use "shore" rather than "coast", as elsewhere in the manuscript.

p3999, line 21: "estimates of the effects of certain model features" would be clearer.

p4000, line 13: should this read "for running climate models on long timescales"?

p4000, line 16: "the models of this category use"

p4002, line 3: "The original data"

p4003, line 1: "used in the Stefan-Boltzmann law"

p4003, line 3: for "unified", should this be "uniform"?

p4003, line 13: "were run"

p4003, lines 15–16: I do not understand "were caused solely by the differences between these schemes and of observation errors"

p4004, line 15: "were run"

p4004, line 16: "with the sediment layer either included or neglected"

p4004, line 17: "No numerical experiments included calibration"

p4008, line 23: "The perturbed experiment is that"

p4009, line 5: replace "lower index" by "subscript"?

p4011, line 15: "is close to that of the lower"

p4015, equation 4: What does the subscript h on F signify?

p4016, line 16: I do not understand "The lake has no inlet so that an inflow may be caused only by brooks. . .". Are brooks not inlets? Should this be "no major inlet"?

p4018, line 12: Should this be "without sufficient knowledge of the hydrogeology of lakes"?

p4018, line 14: I do not understand "using different lake depths, i.e., its bathymetry". Does this mean "using different lake depths based on its bathymetry"?