Geoscientific
Model Development
Discussions

Interactive comment on "Validation of lake surface state in the HIRLAM NWP model against in-situ measurements in Finland" *by* Laura Rontu et al.

Laura Rontu et al.

laura.rontu@fmi.fi

Reply to reviewer 1

Thank you for your careful reading of the manuscript, leading to helpful remarks and suggestions, which we mostly agree with. We have made modifications throught the whole text, but the kept the line numbers of the original manuscript in this reply. Please find our detailed response below. The difference between our original and corrected manuscript versions is provided in an attached diffpdf file.

General comments: The paper presents results of HIRLAM (v7.4) model integrated to Flake model, lake surface state validation against in-situ observations of lake water temperature and ice cover during the period of 2012-2018 in Finland. In general, the

paper structure is good and it is mainly written well. Same validation results against these in-situ observations have not been published earlier, eventhough some earlier papers have dealt with lake temperature and ice cover observation use in the HIRLAM. However, the noticed bug related to ice cover modelling is rather fundamental in physical way, and dominating the results, and makes me consider revising results with proper snowfall calculations on ice. It seems that in the future the HIRLAM model is no longer used and will be substituted with a new model. In that aspect, erroneous calculation could be docu- mented in this article. The figures and tables could be improved and should be made more visual and reader-friendly; I will provide some specific comments on them. Espe- cially figures and tables should run better in line with the text. Now, some figures are mentioned many pages before that they appear.

Concerning the snow-on-ice bug, it has now been corrected in the operational HIRLAM system, that continues running at FMI. The coming spring will provide material to check if the melting of lake ice is better handled. Also, in earlier experiments described e.g. in Kheyrollah Pour et al., 2014 and Eerola et al., 2014, this bug was not present. However, the results in those experiments were not validated against the ice and snow thickness, even the ice dates were used to a limited extent. In these circumstances, we do not consider it useful to run new HIRLAM experiments for checking the impact of the correction. Please note that in the new operational NWP at FMI, based on HARMONIE-AROME, no lake observations are analysed but Flake runs as it would in a climate model, i.e. continuing directly from the previous short forecast.

We will come back to the figures and tables when replying the specific comments. We agree that they should be improved. To correct the setup of figures at distant pages (caused by use of the default latex with template in the manuscript mode) we will ask advice from the GMD typesetting specialists if needed.

Specific comments: 1. Introduction, first paragraph (page 1, lines 16-19): Please pro- vide some references.

We have first of all added references to papers describing the influence of lakes on weather forecasting in general, then influence on NWP and finally importance of describing the existence of ice correctly. We have selected the references so that they contain further relevant references.

2. You have used observations data for the year 2018 eventhough it is current year, usually provisional data. Is the in-situ data used in the analysis quality controlled? When the in-situ data was uploaded? And until which date the year 2018 data are used?

The operational analysis uses LWT observations from SYKE in real time. Those are quality controlled by the HIRLAM optimal analysis system: 1) excluding each station and comparing interpolated to its location nearby observations and 2) comparison against first guess. We use these quality-checked values from analysis feedback files in this study. Possible corrections by SYKE, made afterwards, were not used. The LID data and ice and snow thickness observations were obtained from SYKE open data base for this study, relying on their quality control: LID was fetched 15.8.2018, snow depth 17.9.2018 and ice thickness 16.10.2018 from http://rajapinnat.ymparisto.fi/api/Hydrologiarajapinta/1.0/odataquerybuilder/

We added a sentence about the quality control and mention how the SYKE data was obtained.

3. Page 3, Figure 1. I would like to have it more visual-friendly. Is there certain meaning with arrow line thickness, if not then harmonize.

We now mention that the thin arrows are related to data flow between the HIRLAM analysis-forecast cycles while the thick ones describe processes within each cycle. We made also another correction to the Figure as suggested by reviewer 2.

4. Page 5, line 16. Please make reference to SYKE network, which year status it is? (There are 34 sites in the network in year 2018 according to the SYKE database) We

C3

explain this better in section 3.2.1. , i.e. that there are 34 stations now from which we use in the operational HIRLAM the original year 2011 selection that has never been changed since that. Originally, we excluded rivers and a couple of stations that then seemed to send data less regularly. The list needs to be updated for HARMONIE-AROME if LSWT analysis will be introduced there in the future.

5. Chapter 3.2.2. Freezing and melting dates. Article Korhonen (2006) has introduced terms for freezing and break-up in English, please use those. See: Korhonen, J. 2006. Long-term changes in lake ice cover in Finland. Nordic Hydrology 37(4-5): 347–363. Thank you, we are aware of this terminology but selected freezing and melting according to the suggestion by our native linguistic advisor Emily Gleeson. In our earlier papers written together with our Canadian colleagues, we have used consistently the terms freeze-up and break-up. Now we did not like the suggested mixture of freezing and break-up, but perhaps there are good reasons to use this combination. We would like to leave the last word to our native British GMD editor of the current manuscript Jason Williams.

6. Please state little bit more why these lakes were chosen. Were they only ones large enough to HIRLAM grid or were there other criteria? The main criteria of selecting just these lakes for LID was the data availability: the most complete time series during the selected years, and a reasonable size that provided the best fraction of lake in HIRLAM grid. We now mention this.

7. I suggest combining figures 3 & 4 to same gridded figure with four graphs. Remove from temperature scale dots after the kelvin numbers. In figure caption open up meaning of fc an fob, an. We kept two figures, that we consider to be more clear in the final two-column setup of the journal. The fc-ob-an were added to captions.

8. Chapter 4.2. is little bit hard to read/understand. Try to rewrite it more clear. Thank you, we tried to clarify. This chapter is re-written totally to make it easier to read.

9. Page 10: Text paragraph, it is not clear what are different percentage categories.

C4

Rewritten

10. Table 2: What are the units in this table? Thank you, units added

11. I suggest combining Figures 5 and 6 to same figure (a and b) Done

12. Page 12, last paragraph: make more clear in the text if you are talking about HIRLAM (analysed/forecast) or observed freezing and melting days. Rewritten

13. Chapter 4.3. Make a reference to where lake area/depth records are taken. GLDB perhaps? We renewed the list in Table A3 based on updated material for GLBD v.3 (not yet available at the Flake site but received by courtesy of Margarita Choulga), made the reference and mentioned it more clearly in the text.

14. Figures 7-10 could be combined to one gridded figure (a, b, c, d) Remove dots after Kelvin scale numbers. We created 2 figures and removed the dots.

15. Figure 11. & 12. Add variable name and Unit in Y-scale. Just one legend could be below graphs since they are all same. For codes 28 and 29 use verbal definitions, please. It seems data until early 2018 is used? Done. Data till summer is used 2018 (see above).

16. Figure 13. Add variable name and Unit in the Y-scale. In headings, use only lake name and years: Lappajärvi 2012-2013, Kilpisjärvi 2012-2013, Simpelejärvi 2012-2013. Done

Technical/typo corrections:

1) Abstract: line 3 "integrated to HIRLAM" -> integrated into HIRLAM Done

2) Use wording "in-situ" or "in situ" through whole text. Now there are both versions in the text. Done

3) I would use "lake ice freezing and "lake ice melting" instead of lake freezing and melting (all text) (e.g. page 2, line 21) Done

4) Page 2, line 31: I would consider revising wording "became available" "were obtained"

5) Page 4, line 31: I would consider revising wording "basic material" "is the basis of this study"

6) Figure 2. Page 6: Please note that abbreviation LID has not been introduced in the text before. This a setup problem, now we repeat the definition in the caption, too.

7) Chapter 3.2.2 "codes 27-30" should not be used in the text or figures, use instead verbal definitions. Codes are so called database codes and not normally used as definitions. They are irrelevant as code numbers. Replaced

8) Please check through the text that LWT and LSWT are used coherently. Page 13, line 1: LWST -> LSWT, Page 18, line 13 SYKE LSWT?

This is a bit problematic. Our idea was to call SYKE observations LWT because they are taken at the depth of 20 cm, not exactly at the surface that the satellite would have seen. However, Flake and HIRLAM analysis are dealing with LSWT even when the analysis is based on observed LWT. Perhaps the easiest solution is to call everything LSWT and mention the small difference when introducing the SYKE observations. We now did this.

9) Chapter 3.2.3 Ice thickness and snow depth on lakes Done

10) Page 7, line 8: typo Simpelejärvi Corrected

11) Chapter 3.3.1. Lake surface water temperature (LSWT) Corrected

12) Page 8, line 2: Use verbal definition instead of category 29. Same in line 3 for category 28. Done

13) Page 8, line 9: SYKE LWT observations See 8

14) Page 8, line 21: typo known Corrected

15) Page 15, line 9: 125 Wm-2 (superscript) Corrected

16) Page 15, line 19: 2012-2018 ? This is LaTex's work ...

17) Page 18, line 1: wrong -> incorrect/erroneous Corrected

18) Page 18, line 17: ice thickness and snow depth Corrected, also elsewhere

19) References: Please check that all references are formatted same way. For example, if many initial letters using space between or not in consistent way. I noticed some typos: Thank you, corrected as suggested

Page 24,1. Potes, M. -> Potes, M.

Page 24, line 22. Gandin, L. missing :

Page 25, line 5. Remove ++ after Hydrology Research.

Page 25, line 11. co authors -> write all names

Page 25, line 33 et al > write all names

Page 25, line 33. Yang, Y., coauthors -> write all names and put the year in the end

Reply to reviewer 2

Thank you for your helpful remarks and suggestions, which we mostly agree with. We have made modifications throught the whole text, but the kept the line numbers of the original manuscript in this reply. Please find our detailed response below. The difference between our original and corrected manuscript versions is provided in an attached diffpdf file.

General comments:

Rontu et al. utilize archived forecast data (2012-2018) from the NWP model HIRLAM to validate the analysed and forecasted state of lakes with respect to observations within a model domain. Due to unfortunate circumstances this specific HIRLAM version

included a bug which prevented snow to accumulate on the lake ice. Due to this bug the model data related to ice behaviour and spring LSWT temperature became unrealistic and therefore the corresponding results and discussions are of very limited interest. Okay, it illustrates the importance of representing snow on ice when simulating lakes in cold climate conditions.

Indeed, this bug was not present in our earlier experiments, e.g. Eerola et al., 2014 nor is it there in the latest (development) version of HIRLAM reference code. Now it is corrected also in the FMI operational version, that will allow to check the situation during the coming spring.

The manuscript is in general carefully written and can be considered as a useful guidance on how to validate the state of lakes in a NWP or climate model when corresponding in-situ observational data are available. The authors carefully describe uncertainties with respect to representativeness of observations and representation of lakes in a model domain. Also, they describe how ice conditions may be estimated based on other data. All this information can be valuable for scientists planning similar exercises for other combinations of model and lake observations.

Thank you for the positive comment, nice to hear that our methods are considered useful!

As the authors say it is a well known behaviour of FLake to overestimate summer LSWT. This is also seen in the presented results. However, it can not be excluded that part of those biases presented may be explained by for example any biases in near-surface temperature conditions in general. After all, the lakes represent only some 10% of the land area even in Finland so a bias in near-surface air temperature due to discrepancies in representation of land processes can also contribute to the presented biases. Thus, I would like to see a comment on the general near-surface temperature bias in this HIRLAM setup. The authors do comment on the quality of simulated down-welling short-wave irradiance but a comment on long-wave would also be relevant.

FLake works over the fraction of lake in each gridbox, driven by the average radiative and specific over-lake turbulent fluxes at the lake-atmosphere interface. The lake water and ice temperatures and other in-lake prognostic variables result from the Flake prognostic parametrizations. The resulting (diagnostic) LSWT represents the lake surface temperature in each gridbox, while the land-surface tile is taken care by other parametrizations (ISBA land-surface scheme), which also essentially solve the surface temperature from the equation of surface energy balance, taking into account also the heat conduction in ground. The grid-average screen-level temperature, that is regularly verified against observations, results from intelligent interpolation between the surface (e.g. LSWT) and the lowest model level temperature. In practice, the latter seems to dominate, but in principle, T2m could be wrong due to wrong LSWT but not vice versa. While there is no direct connection between the average (dominated by land surface) predicted surface temperature and LSWT, both might be inaccurate due to inaccurate atmospheric forcing. Wrong radiation transfer in the model, for example due to the cloudiness or incorrect handling of cloud-radiation interactions, biased near-surface air temperatures (at the lowest model level) or wrong turbulent fluxes in the atmospheric boundary layer could be sources of such inaccuracies.

Presumably, the shortwave radiation is the dominating factor for the lake water and ice thermodynamics during the year. In the equation of surface energy balance, the radiation fluxes are net fluxes over specific surface types, and these depend also on the prescribed surface properties, in our case e.g. on the lake ice and snow albedo. It would be worth while to perform sensitivity studies, e.g. with a single-column version of a NWP model, to reveal how Flake parametrizations would react to the inaccuracies of the atmospheric forcing and to quantify the related uncertaincies. This could be left for a further study for example in the framework of HARMONIE-AROME NWP system.

We added a sentence "Most importantly, FLake provides HIRLAM with the evolving lake surface (water, ice, snow) temperature, that influences the HIRLAM forecast of the grid-average near-surface temperatures." into the Flake description (Section 2.1). We

also discuss the uncertainties related to atmospheric forcing where only the shortwave flux is now mentioned in the conclusions. We come back to the temperature aspect in the reply of your Kilpisjärvi comment.

Detailed comments:

Page 2, line 3: Sounds a bit strange to combine observed LSWT and simulated ice thickness to estimate fractional ice: " Fractional ice cover (lake ice concentration in each grid-square of the model) is estimated separately based on the analysed LSWT and the ice thickness predicted by Flake." We improved our unfortunate formulation that allowed misunderstanding and relocated the explanations into their proper sections. There are two cases and two ways to estimate ice cover extent: in analysis, only LSWT exists, so it is used there in similar way that is done for SST – full ice concentration if the grid-average temperature is -0.5oC, none when it is 0oC and linearly in between. In the forecast by Flake, only ice thickeness is available. When it is larger than a small treshold value, the diagnostics decides that lakes existing in this gridbox are all frozen, i.e. the ice concentration is 1. There is a fraction of lakes in each gridbox, so the grid-scale ice fraction is obtained by multiplying the ice concentration with lake fraction. Thus, 'separately' meant: based on LSWT for analysis and based on ice thickness for forecast.

Page 5, line 15 19: Here you refer to Figure 2 for the first time but in the caption of Figure 2 you use the abbreviation LID which is defined later in the text. Please, e.g., introduce "lake ice dates" also in the figure caption for clarity. Done

Page 8, lines 1-2: A bit strangely formulated sentence: "including in the comparison data over all months". Please make it more clear. Done. The idea was that in the LSWT (obsa file) comparisons the winter months were excluded but here we used all data.

Pages 9-12, Section 4.2: The bug which prevents snow to accumulate on ice in this HIRLAM version will seriously affect all results presented in this section so I think it

would be fair to the reader to comment on this in the beginning of this section although it has been mentioned in previous sections. We now discuss the reasons for too early melting when showing the results here. This section has been largely rewritten.

Page 13, line 5: You say that "Lake Kilpisjärvi is an Arctic lake at the elevation of 473 m". This is a complex terrain area so the height difference between the real lake and the model lake might contribute to estimated biases in temperature. What is the corresponding height of the HIRLAM grid box here? Would a height correction of tem- perature make any difference for the results? The mean surface elevation of this gridbox where Lake Kilpisjärvi occupies around 40% of the area, is 614 m that is higher than the lake elevation because the lake is located in a valley surrounded by mountains. The diagnostic screen-level temperature, to which a heigh correction of temperature could be applied, plays no role in the model's air-surface energy exchange. To our understanding, there is no way in Flake to apply height corrections as part of the prognostic calculations or diagnosis of lake surface (snow, ice, water) temperature, also we are not aware of studies related to this issue.

The observed LSWT is evidently measured on the lake at the correct height. During the objective analysis, Kilpisjärvi LSWT is influenced by the observation on the lake and possibly on the nearby lakes, which are probably too far from here to really influence the analysis result. Differences in lake elevations could in principle be taken into account in the optimal interpolation, but this is not currently done. More detailed discussion about the objective analysis of LSWT can be found in the paper by Kheyrollah Pour et al. 2017.

We now mention the difference in Kilpisjärvi and grid-average elevations.

Figure 1: In the text it says that (page 2, line 33 – page 3, line 1) "the prognostic Flake variables are not corrected using the analysed LSWT, which would require advanced data assimilation methods" but in the figure it says "INDEPENDENT LAKE DATA AS-SIMILATION IN AN INTEGRATED NWP + LAKE MODEL". I suggest to remove "DATA

ASSIMILATION" here since that is not done according to the text. And ice cover is simply 0 or 1 when ice is present or not, right? So, this is not really a diagnostic estimation I would say. Or does this include something else which is not yet clear from the text. . .? Okay, becomes clear on page 4. Maybe better to refer to Figure 1 a bit later when the background to the figure is clear from the text. We agree with the suggestion about "INDEPENDENT LAKE DATA ASSIMILATION" and replaced it with "OBJECTIVE ANALYSIS OF LSWT" in the figure. We also moved the figure and reference to it into Section 2.2.

Figure 11: Colour indications of freezing and melting dates in the caption (blue and red) do not fit with the figure (orange and magenta). Corrected

Reply to reviewer 3

Thank you for your helpful remarks and suggestions, which we mostly agree with. We have made modifications throught the whole text, but the kept the line numbers of the original manuscript in this reply. Please find our detailed response below. The difference between our original and corrected manuscript versions is provided in an attached diffpdf file.

General comments:

The paper presents the detailed validation of the FLake model implemented in the HIRLAM NWP system, focusing mainly on the lake surface state and utilizing in situ measurements. The validation period is considerably large spanning over six years and a large number of lakes are included in the investigation. The validation area covers only Finnish lakes, consequently results are referring to arctic conditions and might not be generalized to other climate regimes. The technical properties of the modelling system as well as the observational dataset are described properly. A lake water temperature assimilation scheme is also presented, however, it is mentioned that this is only a diagnostic product. Perhaps, the application areas of this product could be highlighted so that the purpose of it is clearer for the reader. We added a sentence

about the possible use of the diagnostic analysis into section 2.2.

During the validation, lake surface water temperature (LSWT), freezing and melting dates and ice thickness are investigated. Regarding LSWT results are in line with previous studies, namely an overestimation by FLake is pointed out. Freezing dates are simulated by an adequate precision, however, melting dates are poorly forecasted. The cause of this problem is enlightened during the investigation of the ice and snow thicknesses, namely due to a coding error snow is not accumulated on the ice surface. Physical consequences of this bug (missing insulation in winter and different albedo in spring) are well described.

Detailed comments:

1. Page 5 line 18: it is mentioned that water temperature is measured at 20 cm below water surface. Could the authors comment, whether this depth was used also in previous validation studies they are referring to (e.g. Kourzeneva 2014). Also, are there any difficulties in the validation when water is already frozen, but ice thickness is not reaching 20 cm? Yes, we have always used the same SYKE observations in our papers. These observations are only available during the ice-free period as mentioned in Section 3.2.1. and were used only then. There may be gaps between the observed freezing and melting dates and the dates when LSWT observations are made. Also, the locations of LID observations and LSWT measurements are not necessarily the same at the lakes where both types of observation are available. We added a couple of sentences about this into Section 3.2.2

2. Page 10, line 8: "with an area of 1 kmËĘ-2" should be "with an area of 1 kmËĘ2" Corrected

3. Page 13 line 14: "common to the majority of lakes" is a bit vague, "similar to the results averaged over all lakes" might more precise. Corrected as suggested

4. Page 15, line 9: "125 Wm-2": "-2" should be superscript as one line above. Cor-

C13

rected

5. Perhaps the authors could shortly comment, whether the bug revealed had any detectable impact on the forecasts of atmospheric variables (e.g. 2 m temperature) in the HIRLAM model in the six year period. The problem is that we do not know, because there is no way to compare the results with Flake (containing the bug) to those without FLake or with correct FLake as operationally only the parametrization with the bug was applied. The coming spring may show something because now the bug has been corrected while everything else remains unmodified in the operational HIRLAM system. Another problem is that there are not too much SYNOP stations making screen-level temperature observations in the immediate vicinity of the lakes so it is not easy to detect the impact in the verification statistics – these aspects where discussed by Eerola et al., 2014. Case studies might help, though. We mention this now shortly in the concluding section.

6. The use of in-situ observations is definitely of great value in the validation of lake surface state, however, when describing plans the authors might comment on the usability of satellite products as well. We added into the conclusions a sentence about the perspectives of using satellite products in the future.

Please also note the supplement to this comment:
https://www.geosci-model-dev-discuss.net/gmd-2018-270/gmd-2018-270-AC2-supplement.pdf

―――――――――――――――――――――

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-270, 2018.