Validation of lake surface state in the HIRLAM v.7.4 NWP model against *in-situ* measurements in Finland

Laura Rontu[1], Kalle Eerola[1], and Matti Horttanainen[1]
[1]Finnish Meteorological Institute, P.O. Box 503, 00101 Helsinki, Finland
Correspondence: laura.rontu@fmi.fi

Abstract. High Resolution Limited Area Model (HIRLAM), used for the operational numerical weather prediction in the Finnish Meteorological Institute (FMI), includes prognostic treatment of lake surface state since 2012. Forecast is based on the Freshwater Lake (FLake) model integrated into HIRLAM. Additionally, an independent objective analysis of lake surface water temperature (LSWT) combines the short forecast of FLake to observations from the Finnish Environment Institute (SYKE). The resulting description of lake surface state - forecast FLake variables and analysed LSWT - was compared to SYKE observations of lake water temperature, freeze-up and break-up dates as well as the ice thickness and snow depth for 2012-2018 over 45 lakes in Finland. During the ice-free period, the predicted LSWT corresponded to the observations with a slight overestimation, with a systematic error of + 0.91 K. The colder temperatures were underrepresented and the maximum temperatures were too high. The objective analysis of LSWT was able to reduce the bias to + 0.35 K. The predicted freeze-up dates corresponded well the observed dates, mostly within the accuracy of a week. The forecast break-up dates were far too early, typically several weeks ahead of the observed dates. The growth of ice thickness after freeze-up was generally overestimated. However, practically no predicted snow appeared on lake ice. The absence of snow, presumably be due to an incorrect security coefficient value, is suggested to be also the main reason of the inaccurate simulation of the lake ice melting in spring.

1 Introduction

Lakes influence the energy exchange between the surface and the atmosphere, the dynamics of the atmospheric boundary layer and the near-surface weather. This is important for weather forecasting over the areas where lakes, especially those with a large yearly variation of the water temperature, freezing in autumn and melting in spring, cover a significant area of the surface (Kheyrollah Pour et al., 2017; Laird et al., 2003 and references therein). Description of the lake surface state influences the numerical weather prediction (NWP) results, in particular in the models whose resolution is high enough to account for even the smaller lakes (Eerola et al., 2014 and references therein). Especially, the existence of ice can be important for the numerical forecast (Eerola et al., 2014; Cordeira and Laird, 2008).

Validation of lake surface state in the HIRLAM v.7.4 NWP model against *in-situ* measurements in Finland

Laura Rontu[1], Kalle Eerola[1], and Matti Horttanainen[1]
[1]Finnish Meteorological Institute, P.O. Box 503, 00101 Helsinki, Finland
Correspondence: laura.rontu@fmi.fi

Abstract. High Resolution Limited Area Model (HIRLAM), used for the operational numerical weather prediction in the Finnish Meteorological Institute (FMI), includes prognostic treatment of lake surface state since 2012. Forecast is based on the Freshwater Lake (FLake) model integrated into HIRLAM. Additionally, an independent objective analysis of lake surface water temperature (LSWT) combines the short forecast of FLake to observations from the Finnish Environment Institute (SYKE). The resulting description of lake surface state - forecast FLake variables and analysed LSWT - was compared to SYKE observations of lake water temperature, freezing and melting dates as well as the ice thickness and snow depth for 2012-2018 over 45 lakes in Finland. During the ice-free period, the predicted LSWT corresponded to the observations with a slight overestimation, with a systematic error of + 0.91 K. The colder temperatures were underrepresented and the maximum temperatures were too high. The objective analysis of LSWT was able to reduce the bias to + 0.35 K. The predicted freezing dates corresponded well the observed dates, mostly within the accuracy of a week. The forecast melting dates were far too early, typically several weeks ahead of the observed dates. The growth of ice thickness after freezing was generally overestimated. However, practically no predicted snow appeared on lake ice. The absence of snow, found to be due to a technical error in HIRLAM, is suggested to be also the reason of the inaccurate simulation of the lake ice melt in spring.

Copyright statement. ©Authors 2018. This work is distributed under the Creative Commons Attribution 4.0 License.

1 Introduction

Lakes influence the energy exchange between the surface and the atmosphere, the dynamics of the atmospheric boundary layer and the near-surface weather. This is important for weather forecasting over the areas where lakes, especially those with a large yearly variation of the water temperature, freezing in autumn and melting in spring, cover a significant area of the surface (Kheyrollah Pour et al., 2017; Laird et al., 2003 and references therein). Description of the lake surface state influences the numerical weather prediction (NWP) results, in particular in the models whose resolution is high enough to account for even the smaller lakes (Eerola et al., 2014 and references therein). Especially, the existence of ice can be important for the numerical forecast (Eerola et al., 2014; Cordeira and Laird, 2008).

1

In the Finnish Meteorological Institute (FMI), the High Resolution Limited Area Model HIRLAM (Undén et al., 2002; Eerola, 2013) has been applied since 1990 for the numerical short-range weather forecast. In the beginning, the monthly climatological water surface temperature for both sea (sea surface temperature SST) and lakes (Lake Surface Water Temperature LSWT) was used. Since 2012, HIRLAM includes a prognostic lake temperature parameterization based on the Freshwater Lake Model (FLake, Mironov et al., 2010). An independent objective analysis of observed LSWT (Kheyrollah Pour et al., 2017 and references therein) was implemented in 2011. The fractional ice cover (lake ice concentration in each gridsquare of the model) is diagnosed from the analysed LSWT.

FLake was designed to be used as a parametrization scheme for the forecast of the lake surface state in NWP and climate models. It allows to predict the lake surface state in interaction with the atmospheric processes treated by the NWP model. The radiative and turbulent fluxes as well as the predicted snow precipitation from the atmospheric model are combined with FLake processes at each time-step of the model integration in the model grid, where the fraction and depth of lakes are prescribed.

FLake has been implemented into the other main European NWP and regional climate models, first into COSMO (Mironov et al., 2010) then into ECMWF (Balsamo et al., 2012), Unified Model (Rooney and Bornemann, 2013), SURFEX surface modelling framework (Masson et al., 2016), regional climate models RCA (Samuelsson et al., 2010), HCLIM (Lindstedt et al., 2015) and REMO (Pietikäinen et al., 2018), among others. Description of lake surface state and its influence in the numerical weather and climate prediction has been validated in various ways. Results of case studies, e.g. Eerola et al. (2014) and shorter-period NWP experiments, e.g. Eerola et al. (2010); Rontu et al. (2012); Kheyrollah Pour et al. (2014); Kheyrollah Pour et al. (2017) as well as climate model results, e.g. Samuelsson et al. (2010); Pietikäinen et al. (2018), have been compared with remote-sensing satellite data and *in-situ* lake temperature and ice measurements as well as validated against the standard weather observations. In general, improvement of the scores has been seen over regions where lakes occupy a significant area. However, specific features of each of the host models influence the results of the coupled atmosphere-lake system as FLake appears to be quite sensitive to the forcing by the atmospheric model.

The aim of the present study is to validate the lake surface state forecast by the operational HIRLAM NWP model using the *in-situ* LSWT measurements, lake ice freeze-up and break-up dates and measurements of ice and snow thickness by the Finnish Environment Institute (Suomen Ympäristökeskus = SYKE). For this purpose, HIRLAM analyses and forecasts archived by FMI were compared with the observations by SYKE over the lakes of Finland from spring 2012 to summer 2018. To our knowledge, this is the longest available detailed dataset that allows to evaluate how well the lake surface state is simulated by an operational NWP model that applies FLake parametrizations.

2 Lake surface state in HIRLAM

FLake was implemented in the HIRLAM forecasting system in 2012 (Kourzeneva et al., 2008; Eerola et al., 2010). The model utilizes external datasets on the lake depth (Kourzeneva et al., 2012a; Choulga et al., 2014) and the lake climatology (Kourzeneva et al., 2012b). The latter is only needed in order to provide initial values of FLake prognostic variables in the very first forecast (so-called cold start). The use of real-time *in-situ* LSWT observations by SYKE for 27 Finnish lakes was

In the Finnish Meteorological Institute (FMI), the High Resolution Limited Area Model HIRLAM (Undén et al., 2002; Eerola, 2013) has been applied since 1990 for the numerical short-range weather forecast. In the beginning, the monthly climatological water surface temperature for both sea (sea surface temperature SST) and lakes (Lake Surface Water Temperature LSWT) was used. Since 2012, HIRLAM includes a prognostic lake temperature parameterization based on the Freshwater Lake Model (FLake, Mironov et al., 2010). An independent objective analysis of observed LSWT (Kheyrollah Pour et al., 2017 and references therein) was implemented in 2011. The fractional ice cover (lake ice concentration in each gridsquare of the model) is diagnosed from the analysed LSWT.

FLake was designed to be used as a parametrization scheme for the forecast of the lake surface state in NWP and climate models. It allows to predict the lake surface state in interaction with the atmospheric processes treated by the NWP model. The radiative and turbulent fluxes from the atmospheric model are combined with FLake processes at each time-step of the model integration in the model grid, where the fraction and depth of lakes are prescribed.

FLake has been implemented into the main European NWP and regional climate models, first into COSMO (Mironov et al., 2010) then into ECMWF (Balsamo et al., 2012), Unified Model (Rooney and Bornemann, 2013), SURFEX surface modelling framework (Masson et al., 2016), regional climate models RCA (Samuelsson et al., 2010), HCLIM (Lindstedt et al., 2015) and REMO (Pietikäinen et al., 2018), among others. Description of lake surface state and its influence in the numerical weather and climate prediction has been validated in various ways. Results of case studies, e.g. Eerola et al. (2014) and shorter-period NWP experiments, e.g. Eerola et al. (2010); Rontu et al. (2012); Kheyrollah Pour et al. (2014); Kheyrollah Pour et al. (2017) as well as climate model results, e.g. Samuelsson et al. (2010); Pietikäinen et al. (2018), have been compared with remote-sensing satellite data and *in-situ* lake temperature and ice measurements as well as validated against the standard weather observations. In general, improvement of the scores has been seen over regions where lakes occupy a significant area. However, specific features of each of the host models influence the results of the coupled atmosphere-lake system as FLake appears to be quite sensitive to the forcing by the atmospheric model.

The aim of the present study is to use *in-situ* LSWT measurements, lake ice freezing and melting dates and measurements of ice and snow thickness by the Finnish Environment Institute (Suomen Ympäristökeskus = SYKE) for validation of the lake surface state forecast by the operational HIRLAM NWP model. For this purpose, HIRLAM analyses and forecasts archived by the Finnish Meteorological Institute (FMI) were compared with the observations by SYKE over the lakes of Finland from spring 2012 to summer 2018. To our knowledge, this is the longest available detailed dataset that allows to evaluate how well the lake surface state is simulated by an operational NWP model that applies FLake parametrizations.

2 Lake surface state in HIRLAM

FLake was implemented in the HIRLAM forecasting system in 2012 (Kourzeneva et al., 2008; Eerola et al., 2010). The model utilizes external datasets on the lake depth (Kourzeneva et al., 2012a; Choulga et al., 2014) and the lake climatology (Kourzeneva et al., 2012b). The latter is only needed in order to provide initial values of FLake prognostic variables in the very first forecast (so-called cold start). Real-time *in-situ* LSWT observations by SYKE for 27 Finnish lakes were obtained in 2011 introduced in 2011 into the operational LSWT analysis in HIRLAM (Eerola et al., 2010; Rontu et al., 2012). In the current operational HIRLAM of FMI, FLake provides the background for the optimal interpolation analysis (OI, based on Gandin, 1965) of LSWT. However, the prognostic FLake variables are not corrected using the analysed LSWT. This would require more advanced data assimilation methods based on e.g. the extended Kalman filter (Kourzeneva, 2014).

2.1 Freshwater lake model in HIRLAM

FLake is a bulk model capable of predicting the vertical temperature structure and mixing conditions in lakes of various depths on time-scales from hours to years (Mironov et al., 2010). The model is based on two-layer parametric representation of the evolving temperature profile in the water and on the integral budgets of energy for the layers in question. Bottom sediments and the thermodynamics of the ice and snow on ice layers are treated separately. FLake depends on prescribed lake depth information. The prognostic and diagnostic variables of HIRLAM FLake together with the analysed lake surface variables in HIRLAM are listed in the Appendix (Table A1).

At each time step during the HIRLAM forecast, FLake is driven by the atmospheric radiative and turbulent fluxes as well as the predicted snowfall, provided by the physical parameterisations in HIRLAM. This couples the atmospheric variables over lakes with the lake surface properties as provided by FLake parametrization. Most importantly, FLake provides HIRLAM with the evolving lake surface (water, ice, snow) temperature and radiative properties, that influence the HIRLAM forecast of the grid-average near-surface temperatures.

Implementation of FLake model as a parametrization scheme in HIRLAM was based on the experiments described by Rontu et al. (2012). Compared to the reference version of FLake (Mironov et al., 2010), minor modifications were introduced, namely, use of constant snow density = 300 $\mathrm{kgm^{-3}}$, molecular heat conductivity = 1 $\mathrm{Jm^{-1}s^{-1}K^{-1}}$, constant albedos of dry snow = 0.75 and ice = 0.5. Bottom sediment calculations were excluded. Global lake depth database (GLDB v.2, Choulga et al., 2014) was used for derivation of mean lake depth in each gridsquare. Fraction of lake was taken from HIRLAM physiography database, where it originates from GLCC (Loveland et al., 2000).

Lake surface temperature is diagnosed from the mixed layer temperature for the unfrozen lake gridpoints and from the ice or snow-on-ice temperature for the frozen points. In FLake, ice starts to grow from an assumed value of one millimeter when temperature reaches the freezing point. The whole lake tile in a gridsquare is considered by FLake either frozen or unfrozen. Snow on ice is accumulated from the model's snowfall at each time step during the numerical integration.

2.2 Objective analysis of LSWT observations

A comprehensive description of the optimal interpolation (OI) of the LSWT observations in HIRLAM is given by (Kheyrollah Pour et al., 2017). Shortly, LSWT analysis is obtained by correcting the FLake forecast at each gridpoint by using the weighted average of the deviations of observations from their background values. Prescribed statistical information about the observation and background error variance as well as the distance-dependent autocorrelation between the locations (observations and gridpoints) are applied. The real-time observations entering the HIRLAM surface analysis system are subject to quality control in two phases. First, the observations are compared to the background, provided by the FLake short forecast. Second, optimal

to be used for the operational LSWT analysis in HIRLAM (Eerola et al., 2010; Rontu et al., 2012). In the current operational HIRLAM at FMI FLake provides the background for the optimal interpolation analysis (OI, based on Gandin, 1965) of LSWT. However, the prognostic FLake variables are not corrected using the analysed LSWT. This would require more advanced data assimilation methods based on e.g. the extended Kalman filter (Kourzeneva, 2014).

2.1 Freshwater lake model in HIRLAM

FLake is a bulk model capable of predicting the vertical temperature structure and mixing conditions in lakes of various depths on time-scales from hours to years (Mironov et al., 2010). The model is based on two-layer parametric representation of the evolving temperature profile in the water and on the integral budgets of energy for the layers in question. Bottom sediments and the thermodynamics of the ice and snow on ice layers are treated separately. FLake depends on prescribed lake depth information. The prognostic and diagnostic variables of HIRLAM FLake together with the analysed lake surface variables in HIRLAM are listed in the Appendix (Table A1).

At each time step of the HIRLAM forecast, FLake is driven by the atmospheric radiative and turbulent fluxes provided by the physical parameterisations in HIRLAM. This couples the atmospheric variables over lakes with the lake surface properties as provided by FLake. 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.

Implementation of FLake model as a parametrization scheme in HIRLAM was based on the experiments described by Rontu et al., 2012. Compared to the reference version of FLake (Mironov et al., 2010), minor modifications were introduced, namely, use of constant snow density = 300 $\mathrm{kgm^{-3}}$, molecular heat conductivity = 1 $\mathrm{Jm^{-1}s^{-1}K^{-1}}$, constant albedos of dry snow = 0.75 and ice = 0.5. Bottom sediment calculations were excluded. Global lake depth database (GLDB v.2, Choulga et al., 2014) was used for derivation of mean lake depth in each gridsquare. Fraction of lake was taken from HIRLAM physiography database, where it originates from GLCC (Loveland et al., 2000).

Lake surface temperature is diagnosed from the mixed layer temperature for the unfrozen lake gridpoints and from the ice or snow-on-ice temperature for the frozen points. In FLake, ice starts to grow from an assumed value of one millimeter when temperature reaches the freezing point. The whole lake tile in a gridsquare is considered by FLake either frozen or unfrozen. Snow on ice is accumulated from the model's snowfall at each time step during the numerical integration.

2.2 Objective analysis of LSWT observations

A comprehensive description of the optimal interpolation (OI) of the LSWT observations in HIRLAM is given by (Kheyrollah Pour et al., 2017). Shortly, LSWT analysis is obtained by correcting the FLake forecast at each gridpoint by using the weighted average of the deviations of observations from their background values. Prescribed statistical information about the observation and background error variance as well as the distance-dependent autocorrelation between the locations (observations and gridpoints) are applied. The real-time observations entering the HIRLAM surface analysis system are subject to quality control in two phases. First, the observations are compared to the background, provided by the FLake short forecast. Second, optimal

Table 1. FMI operational HIRLAM

Domain	From Atlantic to Ural, from North Africa beyond North Pole
Model horizontal / vertical resolution	7 km / 65 levels
HIRLAM version	7.4
Model dynamics	Hydrostatic, semi-Lagrangian, grid-point
Atmospheric physical parametrizations	Savijärvi radiation, CBR turbulence, Rasch-Kristiansson cloud microphysics + Kain-Fritsch convection
Surface physical parametrizations	ISBA-newsnow for surface, FLake for lakes
Data assimilation	Default atmospheric (4DVAR) and surface (OI) analysis
Lateral boundaries	ECMWF forecast
Forecast	Up to +54 h initiated every 6h (00, 06, 12, 18 UTC)

is assumed fully ice-covered when LSWT falls below -0.5°C and fully ice-free when LSWT is above 0°C. Between these temperature thresholds, the fraction of ice changes linearly (Kheyrollah Pour et al., 2014).

The HIRLAM surface data assimilation system produces comprehensive feedback information from every analysis-forecast cycle. The feedback consists of the observed value and its deviations from the background and from the final analysis at the observation point. Bilinear interpolation of the analysed and forecast values is done to the observation location from the nearest gridpoints that contain a fraction of lake. In addition, information about the quality check and usage of observations is provided. Fractions of land and lake in the model grid as well as the weights, which were used to interpolate gridpoint values to the observation location, are given. This information is the basis of the present study (see sections 3.3 and 4).

3 Model-observation intercomparison 2012-2018

In this intercomparison we validated HIRLAM results against observations about the lake surface state. The impact of FLake parametrizations to the weather forecast by HIRLAM was not considered. This is because no non-FLake weather forecasts exist for comparison with the operational forecasts during the validation period.

Throughout the following text, the analysed LSWT refers to the result of OI analysis, where FLake forecast has been used as background (Section 2.2) while the forecast LSWT refers to the value diagnosed from the mixed layer water temperature predicted by FLake (Section 2.1). Observed LSWT refers to the measured by SYKE lake water temperature (Section 3.2).

3.1 FMI operational HIRLAM

FMI operational HIRLAM is based on the last reference version (v.7.4), implemented in spring 2012. (Eerola, 2013 and references therein). FLake was introduced into this version. After that the development of HIRLAM was frozen. Thus, during the years of the present comparison, the FMI operational HIRLAM system remains unmodified, which offers a clean time series of data for the model-observation intercomparison. The general properties of the system are summarised in Table 1.

is assumed fully ice-covered when LSWT falls below -0.5°C and fully ice-free when LSWT is above 0°C. Between these temperature thresholds, the fraction of ice changes linearly (Kheyrollah Pour et al., 2014).

The HIRLAM surface data assimilation system produces comprehensive feedback information from every analysis-forecast cycle. The feedback consists of the observed value and its deviations from the background and from the final analysis at the observation point. Bilinear interpolation of the analysed and forecast values is done to the observation location from the nearest gridpoints that contain a fraction of lake. In addition, information about the quality check and usage of observations is provided. Fractions of land and lake in the model grid as well as the weights, which were used to interpolate gridpoint values to the observation location, are given. This information is the basis of the present study (see sections 3.3 and 4).

3 Model-observation intercomparison 2012-2018

In this intercomparison we validated HIRLAM/FLake results against observations about the lake surface state. The impact of FLake parametrizations to the weather forecast by HIRLAM was not considered. This is because no non-FLake weather forecasts exist for comparison with the operational forecasts during the validation period.

Throughout the following text, the analysed LSWT refers to the result of OI analysis, where FLake forecast has been used as background (Section 2.2) while the forecast LSWT refers to the value diagnosed from the mixed layer water temperature predicted by FLake (Section 2.1). Observed LSWT refers to the measured by SYKE lake water temperature (Section 3.2).

3.1 FMI operational HIRLAM

FMI operational HIRLAM is based on the last reference version (v.7.4), implemented in spring 2012. (Eerola, 2013 and references therein). FLake was introduced into this version. After that the development of HIRLAM was frozen. Thus, during the years of the present comparison, the FMI operational HIRLAM system remains unmodified, which offers a clean time series of data for the model-observation intercomparison. The general properties of the system are summarised in Table 1. In the present study, a coding error in FLake implementation was revealed in the reference HIRLAM v.7.4. A too large critical value to diagnose snow existence prevented practically all accumulation of the forecast snowfall on lake ice in the FMI HIRLAM-FLake operational system.

3.2 SYKE lake observations

In this study we used three different types of SYKE lake observations: LSWT, lake ice dates (LID) and ice thickness and snow depth on lake ice. In total, observations on 45 lakes listed in Appendix (Table A2) were included as detailed in the following. The lake depths and surface areas given in Table A2 are based on he updated lake list of GLDB v.3 (Margarita Choulga, personal communication).

3.2 SYKE lake observations

In this study we used three different types of SYKE lake observations: LSWT, freeze-up and break-up dates and ice thickness and snow depth on lake ice. In total, observations on 45 lakes listed in Appendix (Table A2) were included as detailed in the following. The lake depths and surface areas given in Table A2 are based on he updated lake list of GLDB v.3 (Margarita Choulga, personal communication).

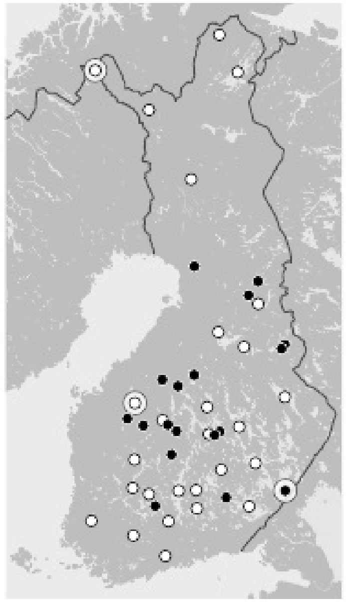


Figure 2. Map of SYKE observation points used in this study: lakes with both lake surface water temperature (LSWT) and lake ice date (LID) observations (white), lakes where only LID is available (black). On Lakes Lappajärvi, Kilpisjärvi and Simpelejärvi also ice thickness and snow depth measurements were used (Section 4.3), they are surrounded with a large white circle. List of the lakes with coordinates is given in Appendix A2.

3.2.1 Lake temperature measurements

Regular *in-situ* lake water temperature measurements are performed by SYKE. Currently SYKE operates 34 regular lake and river water temperature measurement sites in Finland. The temperature of the lake water is measured every morning at 8.00 AM

Table 1. FMI operational HIRLAM

Domain	From Atlantic to Ural, from North Africa beyond North Pole
Model horizontal / vertical resolution	7 km / 65 levels
HIRLAM version	7.4
Model dynamics	Hydrostatic, semi-Lagrangian, grid-point
Atmospheric physical parametrizations	Savijärvi radiation, CBR turbulence,
	Rasch-Kristiansson cloud microphysics + Kain-Fritsch convection
Surface physical parametrizations	ISBA-newsnow for surface, FLake for lakes
Data assimilation	Default atmospheric (4DVAR) and surface (OI) analysis
Lateral boundaries	ECMWF forecast
Forecast	Up to +54 h initiated every 6h (00, 06, 12, 18 UTC)

3.2.1 Lake temperature measurements

Regular *in-situ* lake water temperature measurements are performed by SYKE. Currently SYKE operates 34 regular lake and river water temperature measurement sites in Finland. The temperature of the lake water is measured every morning at 8.00 AM local time, close to shore, at 20 cm below the water surface. The measurements are recorded either automatically or manually and are performed only during the ice-free season (Korhonen, 2002; Rontu et al., 2012). Further, we will for simplicity denote also these data as LSWT observations although they do not represent exactly the same surface water temperature (skin temperature, radiative temperature) that could be estimated by satellite measurements. These data are available in the SYKE open data archive (SYKE, 2018). Measurements from 27 of these 34 lakes (Figure 2, white dots) were selected for use in the FMI operational HIRLAM in 2011, and the list has been kept unmodified since that. The set of 27 daily observations, quality-controlled by HIRLAM, were obtained from the analysis feedback files and used in all comparisons reported in this study.

3.2.2 Freezing and melting dates

Regular visual observations of freezing and melting of lakes have been recorded in Finland for centuries, the longest time series starting in the middle of the 19th century (Korhonen, 2005). Presently, dates of freezing and melting are available from SYKE (2018) on 123 lakes, but the time series for many lakes are discontinuous. Further, we will denote the melting and freezing dates together by "lake ice dates" (LID). For both freezing and melting the dates are available in two categories: for freezing "freezing of the visible area" (code 29 by SYKE) and "permanent freezing of the visible area" (code 30). For melting the dates are defined as "no ice visible from the observation site" (code 28) and "no ice on the outer open water areas" (code 27). LID observations aim at representing conditions on entire lakes. LID observations by SYKE are made independently of their LSWT measurements and possibly from different locations on the same lakes. The LSWT measurements may be started later than the date of reported lake ice melting or end earlier than the reported freezing date.

local time, close to shore, at 20 cm below the water surface. The measurements are recorded either automatically or manually and are performed only during the ice-free season (Korhonen, 2019). Further, we will for simplicity denote also these data as LSWT observations although they do not represent exactly the same surface water temperature (skin temperature, radiative temperature) that could be estimated by satellite measurements. These data are available in the SYKE open data archive (SYKE, 2018). Measurements from 27 of these 34 lakes (Figure 2, white dots) were selected for use in the FMI operational HIRLAM in 2011, and the list has been kept unmodified since that. The set of 27 daily observations, quality-controlled by HIRLAM, were obtained from the analysis feedback files and used in all comparisons reported in this study.

3.2.2 Freeze-up and break-up dates

Regular visual observations of freeze-up and break-up of lakes have been recorded in Finland for centuries, the longest time series starting in the middle of the 19th century (Korhonen, 2019). Presently, dates of freeze-up and break-up are available from SYKE (2018) on 123 lakes, but the time series for many lakes are discontinuous. Further, we will denote the break-up and freeze-up dates together by "lake ice dates" (LID). LID observations aim at representing conditions on entire lakes. For both freeze-up and break-up the dates are available in two categories (terminology from Korhonen, 2019): "freeze-up of the lake within sight" (code 29 by SYKE) and "freeze-up of the whole lake" (code 30). For break-up the dates are defined as "no ice within sight" (code 28) and "thaw areas out of the shore" (code 27). LID observations by SYKE are made independently of their LSWT measurements and possibly from different locations on the same lakes. The LSWT measurements may be started later than the date of reported lake ice break-up or end earlier than the reported freeze-up date.

LID from the 27 lakes whose LSWT measurements are used in HIRLAM were available and selected for this study. In addition, 18 lakes with only LID available (Figure 2, black dots) were chosen for comparison with HIRLAM LID.

3.2.3 Ice thickness and snow depth on lakes

In the period 2012-2018 SYKE recorded the lake ice thickness and snow depth on around 50 locations in Finland. (Archived historical data are available in total from 160 measurement sites). The manual measurements are done three times a month during the ice season. Thickness of ice and snow depth on ice are measured by drilling holes through snow and ice layers along chosen tracks, normally at least 50 m from the coast (Korhonen, 2019). The locations may differ from those of the LSWT measurement or LID observation over the same lakes.

3.3 Validation of HIRLAM lake surface state

3.3.1 Lake surface water temperature

LSWT by HIRLAM, resulting from the objective analysis or diagnosed from the forecast, was compared with the observed LSWT by SYKE using data extracted from the analysis feedback files (Section 2.2) at the observation locations on 06 UTC every day, excluding the winter periods 1 December - 31 March. The observations (ob) at 27 SYKE stations were assumed

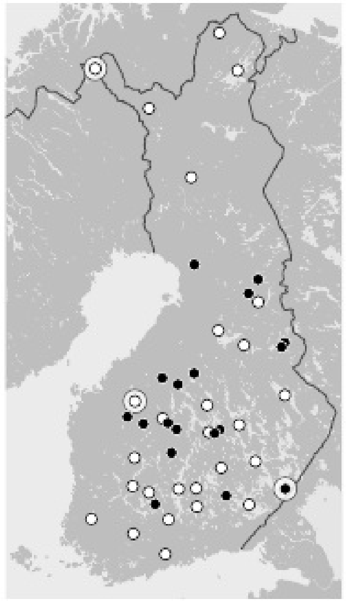


Figure 2. Map of SYKE observation points used in this study: lakes with both lake surface water temperature (LSWT) and lake ice date (LID) observations (white), lakes where only LID is available (black). On Lakes Lappajärvi, Kilpisjärvi and Simpelejärvi also ice thickness and snow depth measurements were used (Section 4.3), they are surrounded with a large white circle. List of the lakes with coordinates is given in Appendix A2.

LID from the 27 lakes whose LSWT measurements are used in HIRLAM were available and selected for this study. In addition, 18 lakes with only LID available (Figure 2, black dots) were chosen for comparison with HIRLAM/FLake LID.

3.2.3 Ice thickness and snow depth on lakes

SYKE records the lake ice thickness and snow depth on around 50 locations in Finland. Archived data are available in total from 160 measurement sites. The manual measurements are done three times a month during the ice season. Thickness of ice and snow depth on ice are measured by drilling holes through snow and ice layers along chosen tracks, normally at least 50 m from the coast (Korhonen, 2005). The locations may differ from those of the LSWT measurement or LID observation over the same lakes.

to represent the true value, while the analysis (an) is the result of OI that combines the background forecast (fc) with the observations. Time-series, maps and statistical scores, to be presented in Section 4.1, were derived from these.

3.3.2 Lake ice conditions

For this study, the observed LID, ice and snow thickness observations were obtained from SYKE open data base, relying on their quality control. The analysed LSWT as well as the predicted ice thickness and snow depth were picked afterwards from the HIRLAM archive for a single gridpoint nearest to each of the 45 observation locations (not interpolated as in the analysis feedback file that was used for the LSWT comparison). It was assumed that the gridpoint value nearest to the location of the LSWT observation represents the ice conditions over the chosen lake.

LID given by HIRLAM were defined in two independent ways: from the analysed LSWT and from the forecast lake ice thickness. Note that the ice thickness and snow depth on ice are not analysed variables in HIRLAM. In autumn a lake can freeze and melt several times before final freeze-up. The last date when the forecast ice thickness crossed a critical value of 1 mm or the analysed LSWT fell below freezing point was selected as the date of freeze-up. In the same way, the last date when the forecast ice thickness fell below the critical value of 1 mm or the analysed LSWT value crossed the freezing point was selected as break-up date. To decrease the effect of oscillation of the gridpoint values between the HIRLAM forecast-analysis cycles, the mean of the four daily ice thickness forecasts or analysed LSWT values was used.

LID by HIRLAM were compared to the observed dates during 2012-2018. In this comparison we included data also during the winter period. The category 29 observations ("freeze-up of the lake within sight", see Section 3.2.2) were used. In this category the time series were the most complete at the selected stations. For the same reason, the break-up observations of category 28 ("no ice within sight") were used for comparison. Furthermore, using a single gridpoint value for the calculation of LID also seems to correspond best the observation definition based on what is visible from the observation site. The statistics were calculated as fc - ob and an - ob. Hence, positive values mean that break-up or freeze-up takes place too late in the model as compared to the observations.

Lake ice thickness and snow depth measurements from lakes Lappajärvi, Kilpisjärvi and Simpelejärvi were utilised as additional data for validation of predicted by HIRLAM ice thickness and snow depth (Section 4.3). These lakes, representing the western, northern and south-eastern Finland, were selected for illustration based on the best data availability during the study years. They are also sufficiently large in order to fit well the HIRLAM grid.

4 Results

4.1 Analysed and forecast LSWT at observation points

Figure 3 shows the frequency distribution of LSWT according to FLake forecast and SYKE observations. It is evident that the amount of data in the class of temperatures which represents frozen conditions (LSWT flag value 272 K) was underestimated

3.3 Validation of HIRLAM/FLake lake surface state

3.3.1 Lake surface water temperature

LSWT by HIRLAM/FLake, resulting from the objective analysis or diagnosed from the forecast, was compared with the observed LSWT by SYKE using data extracted from the analysis feedback files (Section 2.2) at the observation locations on 06 UTC every day, excluding the winter periods 1 December - 31 March. The observations (ob) at 27 SYKE stations were assumed to represent the true value, while the analysis (an) is the result of OI that combines the background forecast (fc) with the observations. Time-series, maps and statistical scores, to be presented in Section 4.1, were derived from these.

3.3.2 Lake ice conditions

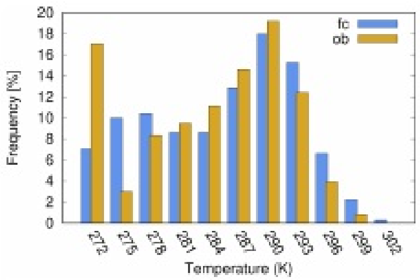
For this study, the observed LID, ice and snow thickness observations were obtained from SYKE open data base, relying on their quality control. The HIRLAM/FLake analysed LSWT as well as the predicted ice thickness and snow depth were picked afterwards from the HIRLAM archive for a single gridpoint nearest to each of the 45 observation locations (not interpolated as in the analysis feedback file that was used for the LSWT comparison). It was assumed that the gridpoint value nearest to the location of the LSWT observation represents the ice conditions over the chosen lake.

LID were defined in two independent ways: from the analysed LSWT and from the forecast lake ice thickness. Note that the ice thickness and snow depth on ice are not analysed variables in HIRLAM. In autumn a lake can freeze and melt several times before final freezing. The last date when the forecast ice thickness crossed a critical value of 1 mm or the analysed LSWT fell below freezing point was selected as the date of freezing. In the same way, the last date when the forecast ice thickness fell below the critical value of 1 mm or the analysed LSWT value crossed the freezing point was selected as melting day. To decrease the effect of oscillation of the gridpoint values between the HIRLAM forecast-analysis cycles, the mean of the four daily ice thickness forecasts or analysed LSWT values was used.

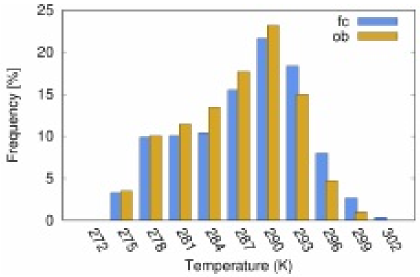
LID by HIRLAM/FLake were compared to the observed dates during 2012-2018. In this comparison we included data also during the winter period. The category 29 observations ("freezing of the visible area", see Section 3.2.2) were used. In this category the time series were the most complete at the selected stations. For the same reason, the melting observations of category 28 ("no ice visible from the observation site") were used for comparison. Furthermore, using a single gridpoint value for the calculation of LID also seems to correspond best the observation definition based on what is visible from the observation site. The statistics were calculated as fc - ob and an - ob. Hence, positive values mean that melting or freezing takes place too late in the model as compared to the observations.

In this study, lake ice thickness and snow depth measurements from lakes Lappajärvi, Kilpisjärvi and Simpelejärvi were utilised as additional data for validation of predicted by HIRLAM/FLake ice thickness and snow depth (Section 4.3). These lakes, representing the western, northern and south-eastern Finland, were selected for illustration based on the best data availability during the study years. They are also sufficiently large in order to fit well the HIRLAM grid.

by the forecast (Figure 3a). When subzero temperatures were excluded from the comparison (Figure 3b), underestimation in the colder temperature classes and overestimation in the warmer classes still remains.



(a) with all temperatures (also frozen conditions) included



(b) only open water temperatures included

Figure 3. Frequency of observed (ob, yellow) and forecast (fc, blue) LSWT over all 27 SYKE lakes 2012-2018. x-axis: LSWT, unit K, y-axis: frequency, unit %.

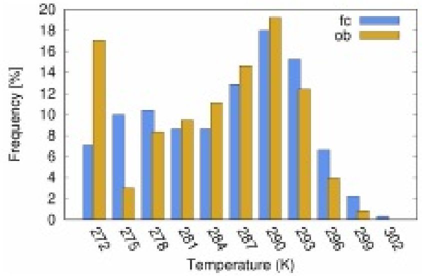
LSWT analysis (Figure 4) improved the situation somewhat but the basic features remain. This is due to the dominance of FLake forecast via the background of the analysis. In Section 4.3, we will show time-series illustrating the physics behind these LSWT statistics.

Table 2 confirms the warm bias by FLake in the unfrozen conditions. Similar results were obtained for all stations together and also for our example lakes Lappajärvi and Kilpisjärvi, to be discussed in detail in Section 4.3. There were three lakes with negative LSWT bias according to FLake forecast, namely the large lakes Saimaa and Päijänne and the smaller Ala-Rieveli. After the correction by objective analysis, a small positive bias converted to negative over 6 additional lakes, among them the

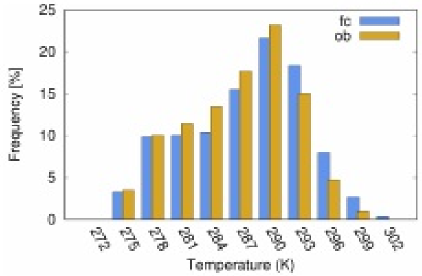
4 Results

4.1 Analysed and forecast LSWT at observation points

Figure 3 shows the frequency distribution of LSWT according to FLake forecast and SYKE observations. It is evident that the amount of data in the class of temperatures which represents frozen conditions (LSWT flag value 272 K) was underestimated by the forecast (Figure 3a). When subzero temperatures were excluded from the comparison (Figure 3b), underestimation in the colder temperature classes and overestimation in the warmer classes still remains.



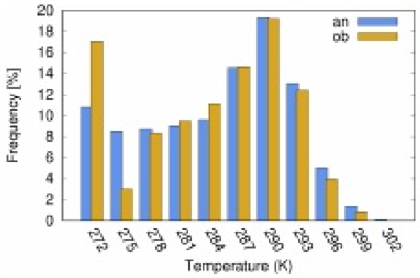
(a) with all temperatures (also frozen conditions) included



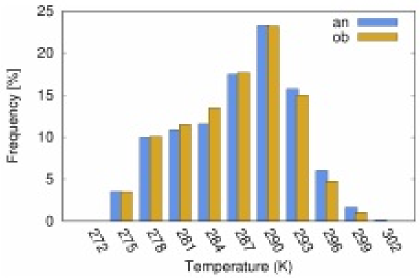
(b) only open water temperatures included

Figure 3. Frequency of observed (ob, yellow) and forecast (fc, blue) LSWT over all 27 SYKE lakes 2012-2018. x-axis: LSWT, unit K, y-axis: frequency, unit %.

LSWT analysis (Figure 4) improved the situation somewhat but the basic features remain. This is due to the dominance of FLake forecast via the background of the analysis. In Section 4.3, we will show time-series illustrating the physics behind these LSWT statistics.



(a) with all temperatures (also frozen conditions) included



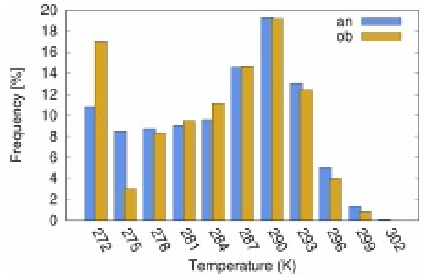
(b) only open water temperatures included

Figure 4. As for Figure 3 but for observed and analysed (an) LSWT.

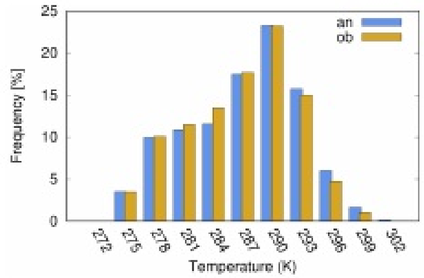
large lakes Lappajärvi in the west and Inari in the north. The mean absolute error decreased from forecast to analysis on every lake.

In the frequency distributions, the warm temperatures were evidently related to summer. For FLake, the overestimation of maximum temperatures, especially in shallow lakes, is a known feature (e.g. Kourzeneva 2014). It is related to the difficulty of forecasting the mixed layer thermodynamics under strong solar heating. Cold and subzero temperatures occurred in spring and autumn. In a few large lakes like Saimaa, Haukivesi, Pielinen, LSWT tended to be slightly underestimated in autumn both according to the FLake and the analysis (not shown). The cold left-hand side columns in the frequency distributions (Figures 3a and 4a) are mainly related to spring, when HIRLAM tended to melt the lakes significantly too early (Sections 4.2 and 4.3).

There are problems, especially in the analysed LSWT, over (small) lakes of irregular form that fit poorly the HIRLAM grid and where the measurements may represent more the local than the mean or typical conditions over the lake. These are the only ones where an underestimation of summer LSWT was seen. Cases occurred where FLake results differ so much from the



(a) with all temperatures (also frozen conditions) included



(b) only open water temperatures included

Figure 4. As for Figure 3 but for observed and analysed (an) LSWT.

Table 2 confirms the warm bias by FLake in the unfrozen conditions. Similar results were obtained for all stations together and also for our example lakes Lappajärvi and Kilpisjärvi, to be discussed in detail in Section 4.3. There were three lakes with negative LSWT bias according to FLake forecast, namely the large lakes Saimaa and Päijänne and the smaller Ala-Rieveli. After the correction by objective analysis, a small positive bias converted to negative over 6 additional lakes, among them the large lakes Lappajärvi in the west and Inari in the north. The mean absolute error decreased from forecast to analysis on every lake.

In the frequency distributions, the warm temperatures were evidently related to summer. For FLake, the overestimation of maximum temperatures, especially in shallow lakes, is a known feature (e.g. Kourzeneva 2014). It is related to the difficulty of forecasting the mixed layer thermodynamics under strong solar heating. Cold and subzero temperatures occurred in spring and autumn. In a few large lakes like Saimaa, Haukivesi, Pielinen, LSWT tended to be slightly underestimated in autumn both according to the FLake and the analysis (not shown). The cold left-hand side columns in the frequency distributions (Figures 3a

Table 2. Statistical scores for LSWT at all stations and at two selected stations

station	fc or an	mean ob	bias	mae	stde	N
unit		K	K	K	K	
ALL	fc	286.3	0.91	1.94	2.34	30877
	an	286.3	0.35	1.32	1.72	30861
Lappajärvi	fc	286.9	0.33	1.23	1.62	1243
	an	286.9	-0.65	1.06	1.10	1243
Kilpisjärvi	fc	281.7	1.82	2.13	2.15	780
	an	281.7	1.10	1.42	1.51	780

Statistics over days when both forecast/analysis and observation indicate unfrozen conditions. bias = systematic difference fc/an - ob, mae = mean absolute error, stde = standard deviation of the error, N = number of days (06 UTC comparison, no ice).

observations that the HIRLAM quality control against background values rejected the observations, forcing also the analysis to follow the incorrect forecast (not shown).

4.2 Freeze-up and break-up dates

In this section the freeze-up and break-up dates from HIRLAM are verified against corresponding observed dates over 45 lakes (Appendix Table A2). In the following, 'LSWT an' refers to the LID estimated from analysed LSWT and 'IceD fc' to those estimated from the forecast ice thickness by FLake. The time period contains six freezing periods (from autumn 2012 to autumn 2017) and seven melting periods (from spring 2012 to spring 2018). Due to some missing data the number of freeze-up cases was 233 and break-up cases 258. The 'IceD fc' data for the first melting period in spring 2012 was missing. The overall statistics of the error in freeze-up and break-up dates are shown in Table 3. In most cases the difference in error between the dates based on forecast and analysis was small. This is natural as the first guess of the LSWT analysis is the forecast LSWT by FLake. We will discuss next the freeze-up, then the break-up dates.

The bias in the error of freeze-up dates was small according to both 'IceD fc' and 'LSWT an', -0.3 and -3.5 days, respectively. The minimum and maximum errors were large in both cases: the maximum freeze-up date occurred about two months too late, the minimum about one and a half months too early. However, as will be shown later, the largest errors mostly occurred on a few problematic lakes while in most cases the errors were reasonable.

Figure 5a) shows the frequency distribution of the error of freeze-up dates. Forecast freeze-up dates occurred slightly more often in the unbiased class (error between -5 - +5 days), compared to the estimated dates from the analysis. Of all cases 48 %/ 40 % (percentages here and in the following are given as 'IceD fc' / 'LSWT an') fell into this class. In 20% / 26% of cases the freeze-up occurred more than five days too late and only in 11% / 9% cases more than two weeks too late. In case of 'IceD fc', the class of freeze-up more than 15 days too late comprised 25 cases distributed over 15 lakes, thus mostly one or

Table 2. Statistical scores for LSWT at all stations and at two selected stations

station	fc or an	mean ob	bias	mae	stde	N
unit		K	K	K	K	
ALL	fc	286.3	0.91	1.94	2.34	30877
	an	286.3	0.35	1.32	1.72	30861
Lappajärvi	fc	286.9	0.33	1.23	1.62	1243
	an	286.9	-0.65	1.06	1.10	1243
Kilpisjärvi	fc	281.7	1.82	2.13	2.15	780
	an	281.7	1.10	1.42	1.51	780

Statistics over days when both forecast/analysis and observation indicate unfrozen conditions. bias = systematic difference fc/an - ob, mae = mean absolute error, stde = standard deviation of the error, N = number of days (06 UTC comparison, no ice).

and 4a) are mainly related to spring, when HIRLAM/FLake tended to melt the lakes significantly too early (Sections 4.2 and 4.3).

There are problems, especially in the analysed LSWT, over (small) lakes of irregular form that fit poorly the HIRLAM grid and where the measurements may represent more the local than the mean or typical conditions over the lake. These are the only ones where an underestimation of summer LSWT was seen. Cases occurred where FLake results differ so much from the observations that the HIRLAM quality control against background values rejected the observations, forcing also the analysis to follow the incorrect forecast (not shown).

4.2 Freezing and melting dates

In this section the freezing and melting dates from HIRLAM are verified against corresponding observed dates over 45 lakes (Appendix Table A2). In the following, 'LSWT an' refers to the LID estimated from analysed LSWT and 'IceD fc' to those estimated from the forecast ice thickness by FLake. The time period contains six freezing periods (from autumn 2012 to autumn 2017) and seven melting periods (from spring 2012 to spring 2018). Due to some missing data the number of freezing cases was 233 and melting cases 258. The 'IceD fc' data for the first melting period in spring 2012 was missing. The overall statistics of the error in freezing and melting dates are shown in Table 3. In most cases the difference in error between the dates based on forecast and analysis was small. This is natural as the first guess of the LSWT analysis is the forecast LSWT by FLake. We will discuss next the freezing, then the melting dates.

The bias in the error of freezing dates was small according to both 'IceD fc' and 'LSWT an', -0.3 and -3.5 days, respectively. The minimum and maximum errors were large in both cases: the maximum freezing day occurred about two months too late the minimum about one and a half months too early. However, as will be shown later, the largest errors mostly occurred on a few problematic lakes while in most cases the errors were reasonable.

Table 3. Statistical measures of the error of freeze-up and break-up date

unit		bias days	stde days	max days	min days	N
Freeze-up	LSWT an	-3.5	17.9	64	-52	233
	IceD fc	-0.3	17.8	67	-41	233
Break-up	LSWT an	-15.2	8.5	2	-54	288
	IceD fc	-20.5	9.2	-1	-56	258

Denotation: LSWT an - LID estimated from analysed LSWT, IceD fc - LID estimated from forecast ice thickness.

two events per lake. This suggests that the error was related more to individual years than to systematically problematic lakes. It is worth noting, that of the eight cases where the error was over 45 days, six cases were due to a single lake, Lake Kevojärvi. This lake is situated in the very north of Finland. It is very small and narrow, with an area of 1 km^2, and located in a steep canyon. Therefore it is poorly represented by the HIRLAM grid and the results seem unreliable.

Concerning too early freezing, in 33% / 44% of the cases freeze-up occurred more than five days too early and in 15% / 19% more than two weeks too early. According to the forecast, these 15% (34 cases) were distributed over 19 lakes. Each of the five large lakes Pielinen, Kallavesi, Haukivesi, Päijänne and Inari occurred in this category three times while all other lakes together shared the remaining 19 cases during the six winters.

The break-up dates (Table 3) show a large negative bias, about two ('LSWT an') or three weeks ('IceD fc'), indicating that lake ice break-up was systematically forecast to occur too early. However, the standard deviation of the error was only about half of that of the error of freeze-up dates and there were no long tails in the distribution (Figure 5b). Hence the distribution is strongly skewed towards too early break-up, but much narrower than that of freeze-up (Figure 5a). The large bias was most probably due missing snow over lake ice in this HIRLAM version (see Section 5). The maximum frequency (47 %) was in the class -24 - -15 days for 'IceD fc', while in case of 'LSWT an', the maximum frequency (52 %) occurred in the class -14 - -5 days. FLake forecast 'IceD fc' suggested only three cases in the unbiased class -4 - +5 while according to 'LSWT an' there were 12 cases in this class. Hence, the break-up dates derived from analysed LSWT corresponded the observations better than those derived from FLake ice thickness forecast.

Note that this kind of method of verifying LID compares two different types of data. The observations by SYKE are visual observations from the shore of the lake (see Section 3.2.2), while the freeze-up and break-up dates from HIRLAM are based on single-gridpoint values of LSWT or ice thickness (see Section 3.3.2). In addition, the resulting freeze-up and break-up dates from HIRLAM are somewhat sensitive to definition of the freezing and melting tresholds. Here we used 1 mm for the forecast ice thickness and the freezing point for the LSWT analysis as the critical values.

In conclusion, the validation statistics show that HIRLAM succeeded rather well in predicting freezing of Finnish lakes. Almost in half of the cases the error was less than ± 5 days. Some bias towards too early freeze-up can be seen both in forecast

Table 3. Statistical measures of the error of freezing and melting date

unit		bias days	stde days	max days	min days	N
Freezing	LSWT an	-3.5	17.9	64	-52	233
	IceD fc	-0.3	17.8	67	-41	233
Melting	LSWT an	-15.2	8.5	2	-54	288
	IceD fc	-20.5	9.2	-1	-56	258

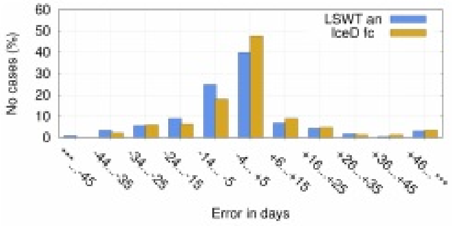
Denotation: LSWT an - LID estimated from analysed LSWT, IceD fc - LID estimated from forecast ice thickness.

Figure 5a) shows the frequency distribution of the error of freezing dates. Forecast freezing dates occurred slightly more often in the unbiased class (error between -5 - +5 days), compared to the estimated dates from the analysis. Of all cases 48 %/40 % (percentages here and in the following are given as 'IceD fc'/'LSWT an') fell into this class. In 20% / 26% of cases the freezing occurred more than five days too late and only in 11% / 9% cases more than two weeks too late. In case of 'IceD fc', the class of freezing more than 15 days too late comprised 25 cases distributed over 15 lakes, thus mostly one or two events per lake. This suggests that the error was related more to individual years than to systematically problematic lakes. It is worth noting, that of the eight cases where the error was over 45 days, six cases were due to a single lake, Lake Kevojärvi. This lake is situated in the very north of Finland. It is very small and narrow, with an area of 1 km^2, and located in a steep canyon. Therefore it is poorly represented by the HIRLAM grid and the results seem unreliable.

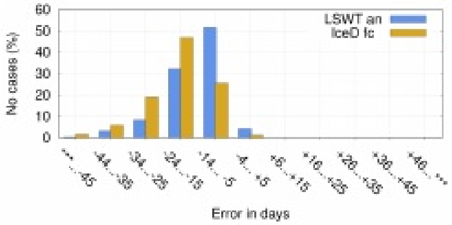
Concerning too early freezing, in 33% / 44% of the cases freezing occurred more than five days too early and in 15% / 19% more than two weeks too early. According to the forecast, these 15% (34 cases) were distributed over 19 lakes. Each of the five large lakes Pielinen, Kallavesi, Haukivesi, Päijänne and Inari occurred in this category three times while all other lakes together shared the remaining 19 cases during the six winters.

The melting dates (Table 3) show a large negative bias, about two ('LSWT an') or three weeks ('IceD fc'), indicating that lake ice melting was systematically forecast to occur too early. However, the standard deviation of the error was only about half of that of the error of freezing dates and there were no long tails in the distribution (Figure 5b). The distribution is strongly skewed towards too early melting, but much narrower than that of freezing (Figure 5a). The large bias was most probably due to the bug of this HIRLAM version that prevented the accumulation of snow over lake ice (see also Section 4.3). The maximum frequency (47 %) was in the class -24 - -15 days for 'IceD fc', while in case of 'LSWT an', the maximum frequency (52 %) occurred in the class -14 - -5 days. FLake forecast 'IceD fc' suggested only three cases in the unbiased class -4 - +5 while according to 'LSWT an' there were 12 cases in this class. Hence, the melting dates derived from analysed LSWT corresponded the observations better than those derived from FLake ice thickness forecast.

Note that this kind of method of verifying LID compares two different types of data. The observations by SYKE are visual observations from the shore of the lake (see Section 3.2.2), while the freezing and melting dates from HIRLAM are based on



(a) error of freeze-up dates



(b) error of break-up dates

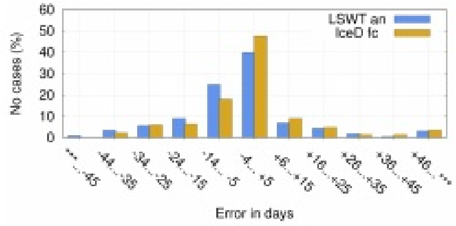
Figure 5. Frequency distribution of the difference between analysed/forecast and observed freeze-up and break-up dates over all lakes 2012-2018. Variables used in diagnosis of ice existence: analysed LSWT crossing the freezing point (blue) and forecast ice thickness > 1 mm (magenta). Observed variable: freeze-up date by SYKE. x-axis: difference (fc-ob), unit day, y-axis: percentage of all cases.

and in the analysis. Melting was more difficult. FLake predicted lake ice break-up always too early, with a mean error of over two weeks, and the analysis mostly followed it.

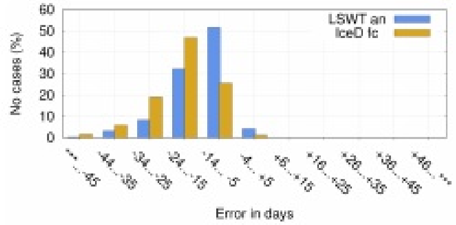
4.3 Comparisons on three lakes

In this section we present LSWT and LID time-series for two representative lakes, Kilpisjärvi in the north and Lappajärvi in the west (see the map in Figure 2). Observed and forecast ice and snow thickness are discussed, using also additional data from Lake Simpelejärvi in southeastern Finland.

Lake Kilpisjärvi is an Arctic lake at the elevation of 473 m, surrounded by fells. The lake occupies 40 % of the area of HIRLAM gridsquare covering it (the mean elevation of the gridsquare is 614 m). The average/maximum depths of the lake are 19.5/57 m and the surface area is 37.3 km^2. The heat balance as well as the ice and snow conditions on Lake Kilpisjärvi have been subject to several studies (Leppäranta et al., 2012; Lei et al., 2012; Yang et al., 2013). Typically, the ice season



(a) error in freezing days



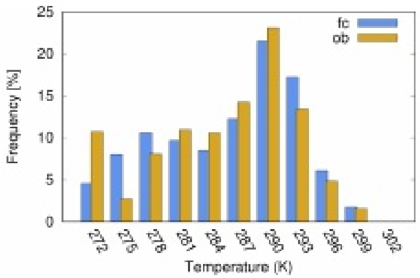
(b) error in melting days

Figure 5. Frequency distribution of the difference between analysed/forecast and observed freezing and melting days over all lakes 2012-2018. Variables used in diagnosis of ice existence: analysed LSWT crossing the freezing point (blue) and forecast ice thickness > 1 mm (magenta). Observed variable: freezing date by SYKE. x-axis: difference (fc-ob), unit day, y-axis: percentage of all cases.

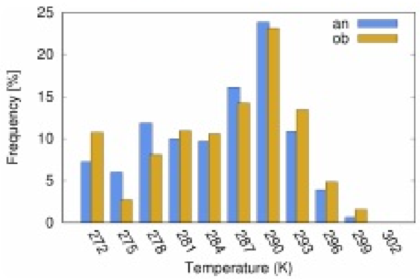
single-gridpoint values of LSWT or ice thickness (see Section 3.3.2). In addition, the resulting freezing and melting dates from HIRLAM are somewhat sensitive to definition of the freezing and melting tresholds. Here we used 1 mm for the forecast ice thickness and the freezing point for the LSWT analysis as the critical values.

In conclusion, the validation statistics show that HIRLAM/FLake succeeded rather well in predicting freezing of Finnish lakes. Almost in half of the cases the error was less than ± 5 days. Some bias towards too early freezing can be seen both in forecast and in the analysis. Melting was more difficult. FLake predicted ice melting always too early, with a mean error of over two weeks, and the analysis mostly followed it. These results are rather obvious because of the missing snow on ice.

lasts there seven months from November to May. Lake Lappajärvi is formed from a 23 km wide meteorite impact crater, which is estimated to be 76 million years old. It is Europe's largest crater lake with a surface area of 145.5 km^2 and an average/maximum depth of 6.9/36 m. Here the climatological ice season is shorter, typically about five months from December to April. The average/maximum depth of Lake Simpelejärvi is 8.7/34.4 m and the surface area 88.2 km^2. This lake is located at the border between Finland and Russia and belongs to the catchment area of Europe's largest lake, Lake Ladoga in Russia.



(a) forecast v.s. observation



(b) analysis v.s. observation

Figure 6. Frequency of observed (yellow) and forecast or analysed (blue) LSWT over Lake Lappajärvi 2012-2018, all temperatures included. x-axis: LSWT, unit K, y-axis: frequency, unit %.

Figures 6 and 7 show the frequency distributions of LSWT according to forecast v.s. observation and analysis v.s. observation for Lappajärvi and Kilpisjärvi. Features similar to the results averaged over all lakes (Section 4.1, Figures 3 and 4) are seen, i.e. underestimation of the amount of cold temperature cases and overestimation of the warmer temperatures by the forecast and analysis. On Lake Lappajärvi, only the amount of below-freezing temperatures was clearly underestimated, otherwise the distributions look quite balanced. According to the observations, on Lake Kilpisjärvi ice-covered days dominated during the

4.3 Comparisons on three lakes

In this section we present LSWT and LID time-series for two representative lakes, Kilpisjärvi in the north and Lappajärvi in the west (see the map in Figure 2). Observed and forecast ice and snow thickness are discussed, using also additional data from Lake Simpelejärvi in southeastern Finland.

Lake Kilpisjärvi is an Arctic lake at the elevation of 473 m, surrounded by fells. The lake occupies 40 % of the area of HIRLAM gridsquare covering it (the mean elevation of the gridsquare is 614 m). The average/maximum depths of the lake are 19.5/57 m and the surface area is 37.3 km^2. The heat balance as well as the ice and snow conditions on Lake Kilpisjärvi have been subject to several studies (Leppäranta et al., 2012; Lei et al., 2012; Yang et al., 2013). Typically, the ice season lasts there seven months from November to May. Lake Lappajärvi is formed from a 23 km wide meteorite impact crater, which is estimated to be 76 million years old. It is Europe's largest crater lake with a surface area of 145.5 km^2 and an average/maximum depth of 6.9/36 m. Here the climatological ice season is shorter, typically about five months from December to April. The average/maximum depth of Lake Simpelejärvi is 8.7/34.4 m and the surface area 88.2 km^2. This lake is located at the border between Finland and Russia and belongs to the catchment area of Europe's largest lake, Lake Ladoga in Russia.

Figures 6 and 7 show the frequency distributions of LSWT according to forecast v.s. observation and analysis v.s. observation for Lappajärvi and Kilpisjärvi. Features similar to the results averaged over all lakes (Section 4.1, Figures 3 and 4) are seen, i.e. underestimation of the amount of cold temperature cases and overestimation of the warmer temperatures by the forecast and analysis. On Lake Lappajärvi, only the amount of below-freezing temperatures was clearly underestimated, otherwise the distributions look quite balanced. According to the observations, on Lake Kilpisjärvi ice-covered days dominated during the periods from April to November. According to both FLake forecast and HIRLAM LSWT analysis the amount of these days was clearly smaller.

Yearly time series of the observed, forecast and analysed LSWT, with the observed LID marked, are shown in Figures 8 and 9. In the absence of observations, the HIRLAM analysis followed the forecast. Missing data in the time series close to freezing and melting are due to missing observations, hence missing information in the feedback files (see Section 2.2). Differences between the years due to the different prevailing weather conditions can be seen in the temperature variations.

Generally, FLake tended to melt the lakes too early in spring, as already indicated by the LID statistics (Section 4.2). The too early melting and too warm LSWT in summer show up clearly in Kilpisjärvi (Figure 9). In Lappajärvi, the model and analysis were able to follow even quite large and quick variations of LSWT in summer, but tended to somewhat overestimate the maximum temperatures. Overestimation of the maximum temperatures by FLake was still more prominent in shallow lakes (not shown). In autumn over Lakes Lappajärvi and Kilpisjärvi, the forecasts and analyses followed closely the LSWT observations and reproduced the freezing dates within a few days, which was also typical to the majority of lakes.

Figure 10 shows a comparison of forecast and observed evolution of ice thickness and snow depth on Lappajärvi, Kilpisjärvi and Simpelejärvi in winter 2012-2013, typical also for the other lakes and years studied. In all three lakes, the ice thickness started to grow after freezing both according to the forecast and the observations. In the beginning HIRLAM/FLake ice grew faster than observed. However, according to the forecast ice thickness started to decrease in March of every year but according periods from November to May. According to both LSWT analysis and forecast the amount of these days was clearly smaller in HIRLAM.

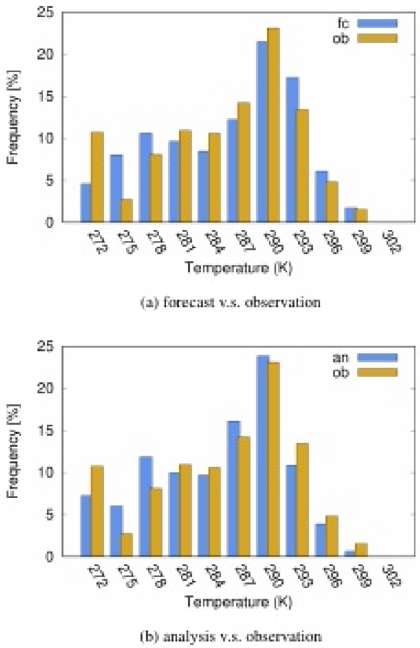


Figure 7. As for Figure 6 but for Lake Kilpisjärvi.

Yearly time series of the observed, forecast and analysed LSWT, with the observed LID marked, are shown in Figures 8 and 9. In the absence of observations, the HIRLAM analysis followed the forecast. Missing data in the time series close to freeze-up and break-up are due to missing observations, hence missing information in the feedback files (see Section 2.2). Differences between the years due to the different prevailing weather conditions are seen in the temperature variations.

Generally, FLake tended to melt the lakes too early in spring, as already indicated by the LID statistics (Section 4.2). The too early break-up and too warm LSWT in summer show up clearly in Kilpisjärvi (Figure 9). In Lappajärvi, the model and analysis were able to follow even quite large and quick variations of LSWT in summer, but tended to somewhat overestimate the maximum temperatures. Overestimation of the maximum temperatures by FLake was still more prominent in shallow

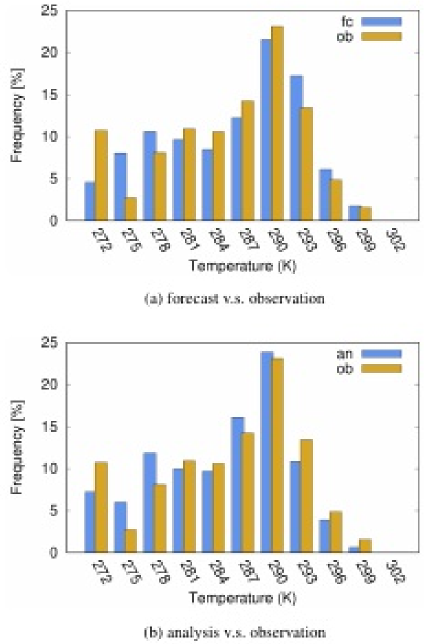


Figure 6. Frequency of observed (yellow) and forecast or analysed (blue) LSWT over Lake Lappajärvi 2012-2018, all temperatures included. x-axis: LSWT, unit K, y-axis: frequency, unit %.

to the observations only a month or two later. The most remarkable feature is that there was no snow in the FLake forecast. It was found that this was due to a coding error in the HIRLAM reference version 7.4 which is applied operationally in FMI.

The too early melting of ice in the absence of snow could be explained by the wrong absorption of the solar energy in the model. In reality, the main factor of snow and ice melt in spring is the increase of daily solar radiation. In HIRLAM, the downwelling short-wave irradiance at the surface is known to be reasonable, with some overestimation of the largest clear-sky fluxes and all cloudy fluxes (Rontu et al., 2017). Over lakes, HIRLAM/FLake uses constant values for the snow and ice shortwave reflection, with albedo values of 0.75 and 0.5, correspondingly. When there was no snow, the lake surface was thus assumed too dark. 25 % more absorption of an assumed maximum solar irradiance of 500 Wm^{-2} (valid for the latitude of Lappajärvi in the end of March) would mean availability of extra 125 Wm^{-2} for melting of the ice, which corresponds the magnitude of increase of available maximum solar energy within a month at the same latitude.

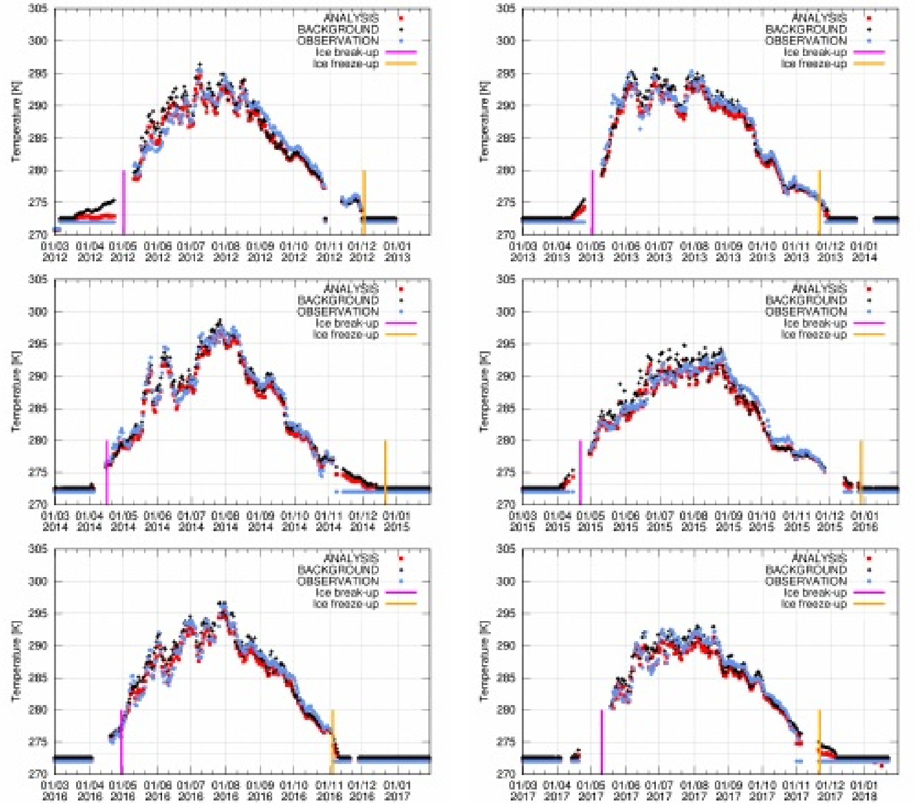
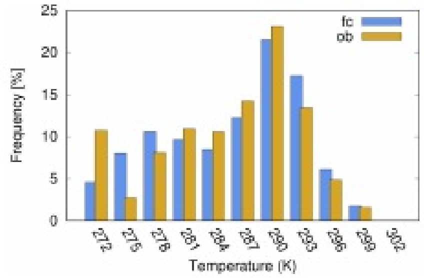
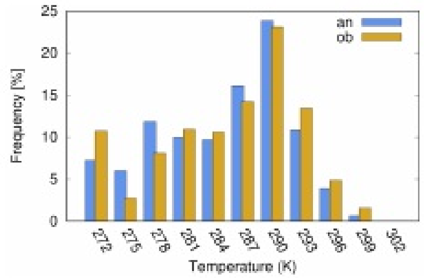


Figure 8. Time-series of the observed, analysed and forecast LSWT at the Lappajärvi observation location 23.67 E, 63.15 N for the years 2012-2018 based on 06 UTC data. Markers are shown in the inserted legend. Observed freeze-up date (blue) and break-up date (red) are marked with vertical lines.

lakes (not shown). In autumn over Lakes Lappajärvi and Kilpisjärvi, the forecasts and analyses followed closely the LSWT observations and reproduced the freeze-up dates within a few days, which was also typical to the majority of lakes.



(a) forecast v.s. observation



(b) analysis v.s. observation

Figure 7. As for Figure 6 but for Lake Kilpisjärvi.

The forecast of too thick ice can also be explained by the absence of snow in the model. When there is no insulation by the snow layer, the longwave cooling of the ice surface in clear-sky conditions is more intensive and leads to faster growth of ice compared to the situation of snow-covered ice. In nature, ice growth can also be due to the snow transformation, a process whose parametrization in the models is demanding (Yang et al., 2013; Cheng et al., 2014).

Also the downwelling longwave radiation plays a role in the surface energy balance. We may expect values from 150 Wm^{-2} to 400 Wm^{-2} in the Nordic spring conditions, with the largest values related to cloudy and the smallest to clear-sky situations. The standard deviation of the predicted by HIRLAM downwelling longwave radiation fluxes has been shown to be of the order of 20 Wm^{-2}, with a positive systematic error of a few Wm^{-2} (Rontu et al., 2017). Compared to the systematic effects related to absorption of the solar radiation, the impact of the longwave radiation variations on lake ice evolution is presumably small.

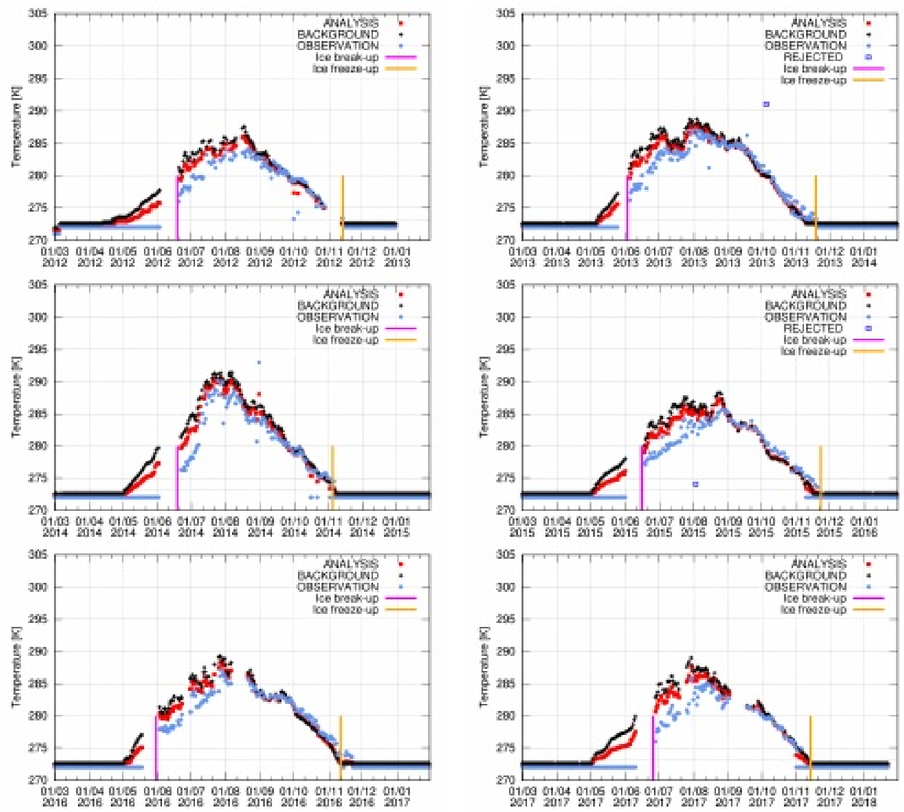


Figure 9. As for Figure 8 but for lake Kilpisjärvi, 20.82 E, 69.01 N.

Figure 10 shows a comparison of forecast and observed evolution of ice thickness and snow depth on Lappajärvi, Kilpisjärvi and Simpelejärvi in winter 2012-2013, typical also for the other lakes and years studied. The most striking feature is that there was no snow in the HIRLAM forecast.

On all three lakes, the ice thickness started to grow after freeze-up both according to the forecast and the observations. In the beginning HIRLAM ice grew faster than observed. However, according to the forecast ice thickness started to decrease in March of every year but according to the observations only a month or two later.

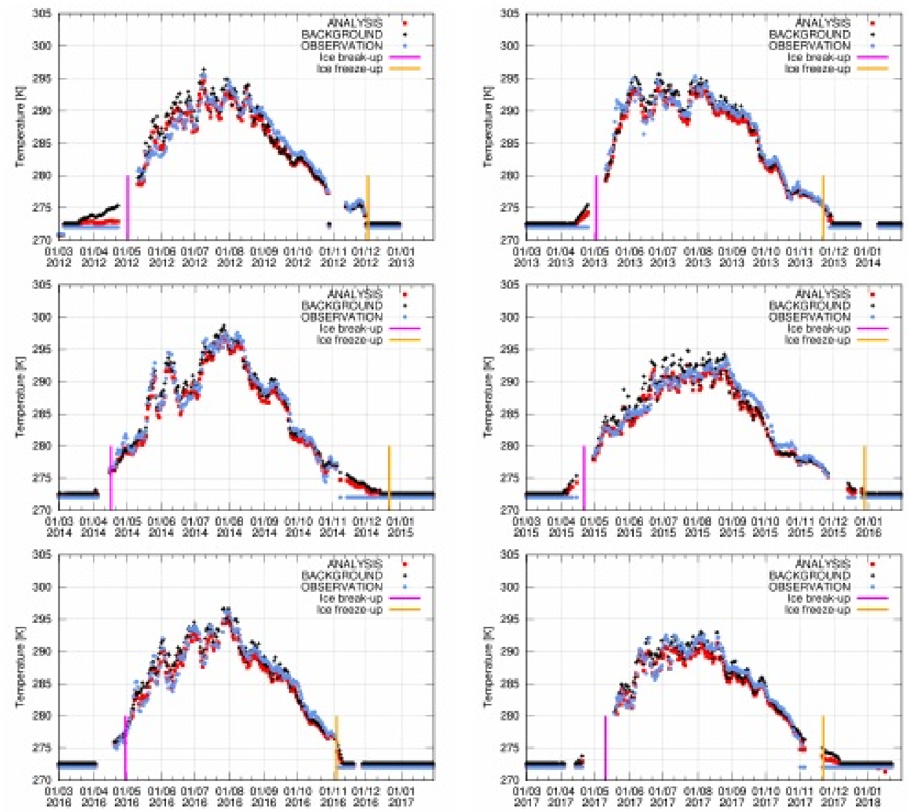


Figure 8. Time-series of the observed, analysed and forecast LSWT at the Lappajärvi observation location 23.67 E, 63.15 N for the years 2012-2018 based on 06 UTC data. Markers are shown in the inserted legend. Observed freezing date (blue) and melting date (red) are marked with vertical lines.

5 Conclusions and outlook

In this study, *in-situ* lake observations from the Finnish Environment Institute were used for validation of the HIRLAM NWP model, which is applied operationally in the Finnish Meteorological Institute. HIRLAM contains Freshwater Lake prognostic

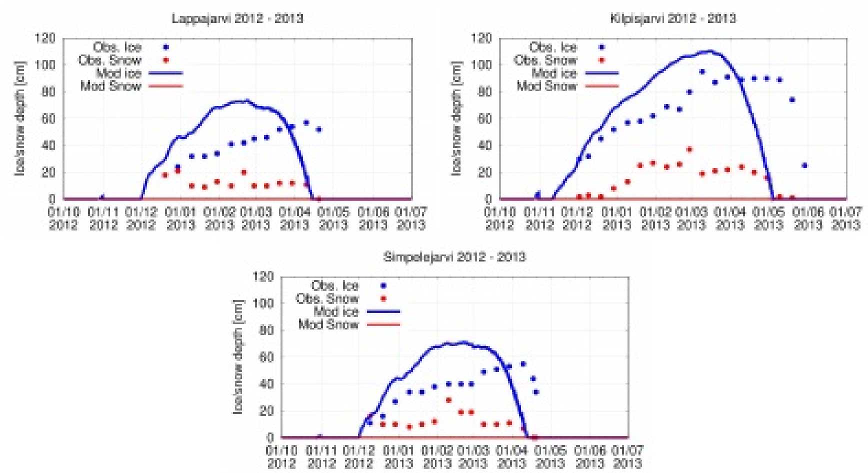


Figure 10. Evolution of ice (blue) and snow (red) thickness at Lakes Lappajärvi, Kilpisjärvi and Simpelejärvi during winter 2012-2013.

The too early break-up of lake ice in the absence of snow could be explained by the wrong absorption of the solar energy in the model. In reality, the main factor of snow and ice melt in spring is the increase of daily solar radiation. In HIRLAM, the downwelling short-wave irradiance at the surface is known to be reasonable, with some overestimation of the largest clear-sky fluxes and all cloudy fluxes (Rontu et al., 2017). Over lakes, HIRLAM uses constant values for the snow and ice shortwave reflection, with albedo values of 0.75 and 0.5, correspondingly. When there was no snow, the lake surface was thus assumed too dark. 25 % more absorption of an assumed maximum solar irradiance of 500 Wm^{-2} (valid for the latitude of Lappajärvi in the end of March) would mean availability of extra 125 Wm^{-2} for melting of the ice, which corresponds the magnitude of increase of available maximum solar energy within a month at the same latitude.

The forecast of too thick ice can also be explained by the absence of snow in the model. When there is no insulation by the snow layer, the longwave cooling of the ice surface in clear-sky conditions is more intensive and leads to faster growth of ice compared to the situation of snow-covered ice. In nature, ice growth can also be due to the snow transformation, a process whose parametrization in the models is demanding (Yang et al., 2013; Cheng et al., 2014).

Also the downwelling longwave radiation plays a role in the surface energy balance. We may expect values from 150 Wm^{-2} to 400 Wm^{-2} in the Nordic spring conditions, with the largest values related to cloudy and the smallest to clear-sky situations. The standard deviation of the predicted by HIRLAM downwelling longwave radiation fluxes has been shown to be of the order

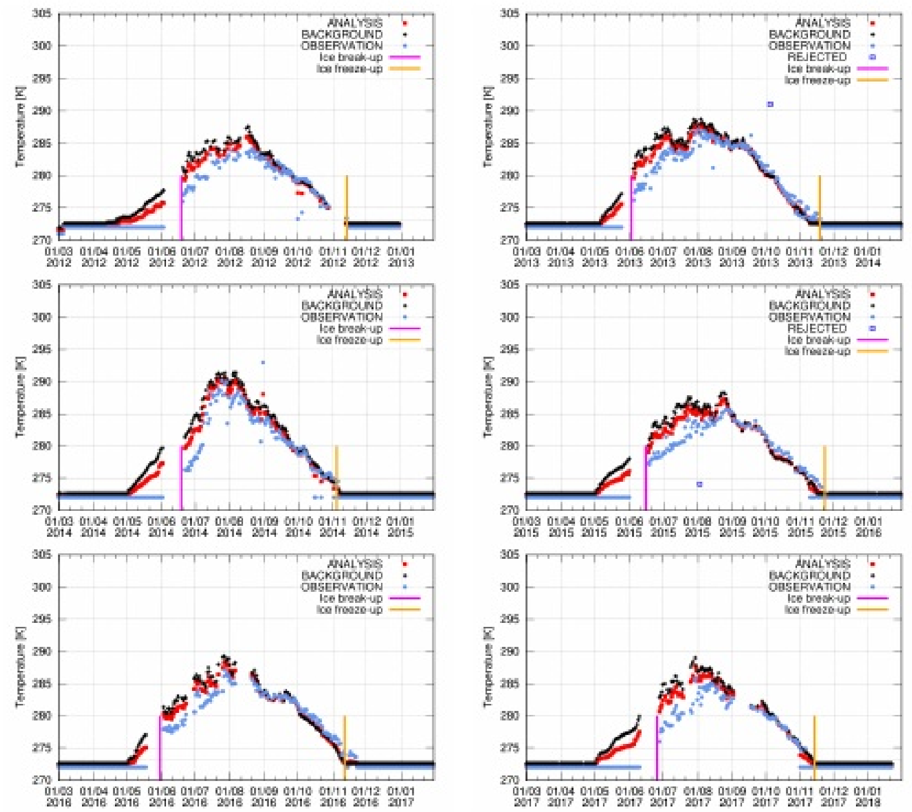


Figure 9. As for Figure 8 but for lake Kilpisjärvi, 20.82 E, 69.01 N.

parametrizations and an independent objective analysis of lake surface state. We focused on comparison of observed and forecast lake surface water temperature, ice thickness and snow depth in the years 2012 - 2018. Because the HIRLAM/FLake system was unmodified during this period, a long uniform dataset was available for evaluation of the performance of FLake integrated into an operational NWP model. On the other hand, no conclusions about the impact of the lake surface state on the operational forecast of the near-surface temperatures, cloudiness or precipitation can be drawn because of the lack of alternative (without FLake) forecasts for comparison.

of 20 Wm^{-2}, with a positive systematic error of a few Wm^{-2} (Rontu et al., 2017). Compared to the systematic effects related to absorption of the solar radiation, the impact of the longwave radiation variations on lake ice evolution is presumably small.

5 Discussion: snow on lake ice

The most striking result reported in Section 4 was the too early melting of the lake ice predicted by FLake in HIRLAM as compared to observations. We suggested that the early break-up is related to the missing snow on lake ice in HIRLAM. It was detected that a too large critical value to diagnose snow existence prevented practically all accumulation of the forecast snowfall on lake ice in the reference HIRLAM v.7.4, used operationally at FMI.

In general, handling of the snow cover on lake and sea ice is a demanding task for the NWP models. In HIRLAM, snow depth observations are included into the objective analysis over the land areas, but not over ice where no observations are widely available in real time. Snow depth and temperature over land are treated prognostically using dedicated parametrizations (in HIRLAM, similar to Samuelsson et al., 2006, 2011, see also Boone et al., 2017). Over the sea, a simple prognostic parametrization of sea ice temperature is applied in HIRLAM but neither the thickness of ice nor the depth or temperature of snow on ice are included (Samuelsson et al., 2006). Batrak et al. (2018) provide a useful review and references concerning prognostic sea ice schemes and their snow treatment in NWP models. An essential difference between the simple sea ice scheme and the lake ice scheme applied in HIRLAM is that the former relies on external data on the existence of sea ice cover, provided by the objective analysis, while the latter includes prognostic treatment of the lake water body also. This means that the lake ice freezes and melts in the model depending on the thermal conditions of lake water, evolving throughout the seasons.

The ice thickness, snow depth and ice and snow temperatures are prognostic variables of FLake. When the FLake parametrizations were introduced into HIRLAM (Kourzeneva et al., 2008; Eerola et al., 2010), parametrization of the snow thickness and snow temperature was first excluded. In the COSMO NWP model, snow is implicitly accounted for by modifying ice albedo using empirical data on its temperature dependence (Mironov et al., 2010). This way was applied also e.g. in a recent study over the Great Lakes (Baijnath-Rodino and Duguay, 2019).

Semmler et al. (2012) performed a detailed winter-time comparison between FLake and a more complex snow and ice thermodynamic model (HIGHTSI) on a small lake in Alaska. FLake includes only one ice and one soil layer, while HIGHTSI represents a more advanced multilayer scheme. Atmospheric forcing for the stand-alone experiments was provided by HIRLAM. Based on their sensitivity studies, Semmler et al. (2012) suggested three simplifications to the original, time-dependent snow-on-ice parametrizations of FLake: use a prescribed constant snow density, modify the value of the prescribed molecular heat conductivity and use prescribed constant albedos of dry snow and ice. Later, a similar comparison was performed over Lake Kilpisjärvi (Yang et al., 2013), confirming the improvements due to the updated snow parametrizations in FLake. Implementation of these modifications allowed to include the parametrization of snow on lake ice also into HIRLAM (Section 2.1).

In FLake, snow on lake ice is accumulated from the predicted snowfall. Snow melt on lake ice is related to snow and ice temperatures. In case of FLake integrated into HIRLAM, accumulation and melt are updated at every time step of the advancing forecast. Very small amounts of snow are considered to fall beyond the accuracy of parametrizations and removed.

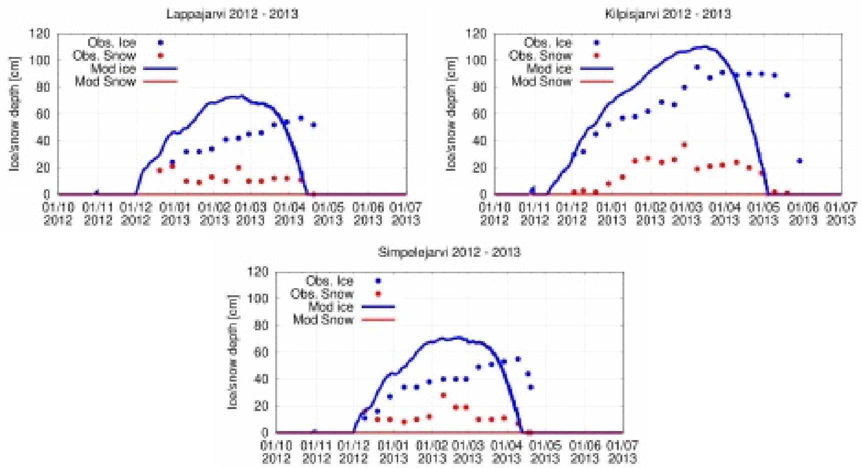


Figure 10. Evolution of ice (blue) and snow (red) thickness at Lakes Lappajärvi, Kilpisjärvi and Simpelejärvi during winter 2012-2013.

On average, the forecast and analysed LSWT were warmer than observed with systematic errors of 0.91 K and 0.35 K, correspondingly. The mean absolute errors were 1.94 and 1.32 K. Thus, the independent observation-based analysis of *in-situ* LSWT observations was able to improve the FLake +6 h forecast used as the first guess. However, the resulting analysis is by definition not used for correction of the FLake forecast but remains an independent by-product of HIRLAM. An overestimation of the FLake LSWT summer maxima was found, especially for the shallow lakes. This behaviour of FLake is well known, documented earlier e.g. by Kourzeneva, 2014. It arises due to the difficulty to handle correctly the mixing in the near-surface water layer that is intensively heated by the sun.

Forecast freezing dates were found to correspond the observations well, typically within a week. The forecast ice thickness tended to be overestimated, still the melting dates over most of the lakes occured systematically several weeks too early. Practically no forecast snow was on found on the lake ice, although the snow parametrization by FLake was included in HIRLAM. The reason for the incorrect behaviour was evidently related to a coding error in HIRLAM that prevented snow accumulation on lake ice. The too early melting and overestimated ice thickness differ from the results by Pietikäinen et al., 2018; Yang et al., 2013; Kourzeneva, 2014, who reported somewhat too late melting of the Finnish lakes when FLake with realistic snow parametrizations was applied within a climate model or stand-alone driven by NWP data. It can be concluded that a realistc parametrization of snow on lake ice is important in order to describe correctly the lake surface state in spring.

19

This is controlled by a critical limit, which was set too large (one millimeter instead of ten micrometers) in HIRLAM v.7.4. Due to the incorrect critical value, practically no snow accumulated on lake ice in the FMI operational HIRLAM, validated in this study. In a HIRLAM test experiment, where the original smaller value was used, up to 17 cm of snow accumulated on lake ice within a month (Janurary 2012, not shown).

6 Conclusions and outlook

In this study, *in-situ* lake observations from the Finnish Environment Institute were used for validation of the HIRLAM NWP model, which is applied operationally in the Finnish Meteorological Institute. HIRLAM contains Freshwater Lake prognostic parametrizations and an independent objective analysis of lake surface state. We focused on comparison of observed and forecast lake surface water temperature, ice thickness and snow depth in the years 2012 - 2018. Because the HIRLAM system was unmodified during this period, a long uniform dataset was available for evaluation of the performance of FLake integrated into an operational NWP model. On the other hand, no conclusions about the impact of the lake surface state on the operational forecast of the near-surface temperatures, cloudiness or precipitation can be drawn because of the lack of alternative (without FLake) forecasts for comparison.

On average, the forecast and analysed LSWT were warmer than observed with systematic errors of 0.91 K and 0.35 K, correspondingly. The mean absolute errors were 1.94 and 1.32 K. Thus, the independent observation-based analysis of *in-situ* LSWT observations was able to improve the FLake +6 h forecast used as the first guess. However, the resulting analysis is by definition not used for correction of the FLake forecast but remains an independent by-product of HIRLAM. An overestimation of the FLake LSWT summer maxima was found, especially for the shallow lakes. This behaviour of FLake is well known, documented earlier e.g. by Kourzeneva (2014). It arises due to the difficulty to handle correctly the mixing in the near-surface water layer that is intensively heated by the sun.

Forecast freeze-up dates were found to correspond the observations well, typically within a week. The forecast ice thickness tended to be overestimated, still the break-up dates over most of the lakes occured systematically several weeks too early. Practically no forecast snow was on found on the lake ice, although the snow parametrization by FLake was included in HIRLAM. The reason for the incorrect behaviour was evidently related to a too large critical value to diagnose snow existence that prevented the accumulation of snow on lake ice. The too early melting and overestimated ice thickness differ from the results by Pietikäinen et al. (2018); Yang et al. (2013); Kourzeneva (2014), who reported somewhat too late melting of the Finnish lakes when FLake with realistic snow parametrizations was applied within a climate model or stand-alone driven by NWP data. It can be concluded that a realistic parametrization of snow on lake ice is important in order to describe correctly the lake surface state in spring.

Small lakes and those of complicated geometry cause problems for the relatively coarse HIRLAM grid of 7 - kilometre resolution. The problems are related to the observation usage, forecast and validation, especially when interpolation and selection of point values are applied. The observations and model represent different spatial scales. For example, the comparison of the freeze-up and break-up dates was based on diagnostics of single-gridpoint values that were compared to observations which

Small lakes and those of complicated geometry cause problems for the relatively coarse HIRLAM grid of 7 - kilometre resolution. The problems are related to the observation usage, forecast and validation, especially when interpolation and selection of point values are applied. The observations and model represent different spatial scales. For example, the comparison of the freezing and melting dates was based on diagnostics of single-gridpoint values that were compared to observations representing entire lakes as overseen from the observation sites. Also the results of LID diagnostics were sensitive to the criteria for definition of the ice existence in HIRLAM/FLake. All this adds unavoidable inaccuracy into the model-observation intercomparison but does not change the main conclusions of the present study.

SYKE LSWT observations used for the real-time analysis are regular and reliable but do not always cover the days immediately after melting or close to freezing, partly because the quality control of HIRLAM LSWT analysis utilizes the SYKE statistical lake water temperature model results in a too strict way. Although the 27 observations are located all over the country, they cover a very small part of the lakes and their availability is limited to Finland. SYKE observations of the ice and snow depth as well as the freezing and melting dates provide valuable data for the validation purposes.

A need for minor technical corrections in the FMI HIRLAM/FLake system was revealed. The snow accumulation bug was corrected in October 2018, based on our findings. Further developments and modifications are not foreseen because the HIRLAM NWP systems, applied in the European weather services, are being replaced by kilometre-scale ALADIN-HIRLAM forecasting systems (Termonia et al., 2018; Bengtsson et al., 2017), where the prognostic FLake parametrizations are also available. HARMONIE/FLake uses the newest version of the global lake database (GLDB v.3) and contains updated snow and ice properties that were suggested by Yang et al., 2013. The objective analysis of lake surface state is yet to be implemented, taking into account the HIRLAM experience summarized in this study and earlier by Kheyrollah Pour et al., 2017. In the future, an important source of wider observational information on lake surface state are the satellite measurements, whose operational application in NWP models still requires further work.

Code and data availability. Observational data was obtained from SYKE open data archive SYKE, 2018 as follows: 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/. A supplementary file containing the freezing and melting dates as picked and prepared for the lakes studied here is attached. Data picked from HIRLAM archive are attached as supplementary files: data from the objective analysis feedback files (observed, analysed, forecast LSWT interpolated to the 27 active station locations) and from the gridded output of the HIRLAM analysis (analysed LSWT, forecast ice and snow thickness from the nearest gridpoint of all locations used in the present study).

In this study, FMI operational weather forecasts resulting from use of HIRLAM v.7.4 (rc1, with local updates) were validated against lake observations. The HIRLAM reference code is not open software but the property of the international HIRLAM-C programme. For research purposes, the codes can be requested from the programme (hirlam.org). The source codes of the version operational at FMI, relevant for the present study, are available from the authors upon request.

represent entire lakes as overseen from the observation sites. Also the results of LID diagnostics were sensitive to the criteria for definition of the ice existence in HIRLAM. All this adds unavoidable inaccuracy into the model-observation intercomparison but does not change the main conclusions of the present study.

SYKE LSWT observations used for the real-time analysis are regular and reliable but do not always cover the days immediately after break-up or close to freeze-up, partly because the quality control of HIRLAM LSWT analysis utilizes the SYKE statistical lake water temperature model results in a strict way. Although the 27 observations are located all over the country, they cover a very small part of the lakes and their availability is limited to Finland. SYKE observations of the ice and snow depth as well as the freeze-up and break-up dates provide valuable data for the validation purposes.

A need for minor technical corrections in the FMI HIRLAM system was revealed. The coefficient influencing snow accumulation on lake ice was corrected based on our findings. Further developments and modifications are not foreseen because the HIRLAM NWP systems, applied in the European weather services, are being replaced by kilometre-scale ALADIN-HIRLAM forecasting systems (Termonia et al., 2018; Bengtsson et al., 2017), where the prognostic FLake parametrizations are also available. HARMONIE/FLake uses the newest version of the global lake database (GLDB v.3) and contains updated snow and ice properties. The objective analysis of lake surface state is yet to be implemented, taking into account the HIRLAM experience summarized in this study and earlier by Kheyrollah Pour et al. (2017). In the future, an important source of wider observational information on lake surface state are the satellite measurements, whose operational application in NWP models still requires further work.

Code and data availability. Observational data was obtained from SYKE open data archive SYKE, 2018 as follows: 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/. A supplementary file containing the freeze-up and break-up dates as picked and prepared for the lakes studied here is attached. Data picked from HIRLAM archive are attached as supplementary files: data from the objective analysis feedback files (observed, analysed, forecast LSWT interpolated to the 27 active station locations) and from the gridded output of the HIRLAM analysis (analysed LSWT, forecast ice and snow thickness from the nearest gridpoint of all locations used in the present study).

In this study, FMI operational weather forecasts resulting from use of HIRLAM v.7.4 (rc1, with local updates) were validated against lake observations. The HIRLAM reference code is not open software but the property of the international HIRLAM-C programme. For research purposes, the codes can be requested from the programme (hirlam.org). The source codes of the version operational at FMI, relevant for the present study, are available from the authors upon request.

Author contributions. Laura Rontu computed the LSWT statistics based on HIRLAM feedback files. Kalle Eerola performed the freeze-up and break-up date, snow and ice thickness comparisons based on data picked from HIRLAM grib files. Matti Horttanainen prepared observation data obtained via SYKE open data interface and lake depths from GLDB v.3. Laura Rontu composed the manuscript text based on input from all authors.

Author contributions. Laura Rontu computed the LSWT statistics based on HIRLAM feedback files. Kalle Eerola performed the freezing and melting date, snow and ice thickness comparisons based on data picked from HIRLAM grib files. Matti Horttanainen prepared observation data obtained via SYKE open data interface and lake depths from GLDB v.3. Laura Rontu composed the manuscript text based on input from all authors.

Competing interests. No competing interests are present.

Acknowledgements. Our thanks are due to Joni-Pekka Pietikäinen and Ekaterina Kourzeneva for discussions and information, to Margarita Choulga and Olga Toptunova for the support with the updated GLDB v.3 data, and to Emily Gleeson for advice with English language.

Competing interests. No competing interests are present.

Acknowledgements. Our thanks are due to Joni-Pekka Pietikäinen and Ekaterina Kourzeneva for discussions and information, to Margarita Choulga and Olga Toptunova for the support with the updated GLDB v.3 data, and to Emily Gleeson for advice with English language. The comments of three anonymous reviewers and the editor significantly helped to improve the presentation of our results.

References

Balsamo, G., Salgado, R., Dutra, E., Boussetta, S., Stockdale, T. and Potes, M: On the contribution of lakes in predicting near-surface temperature in a global weather forecasting model. Tellus A, 64, 15829, https://doi.org/10.3402/tellusa.v64i0.15829, 2012.

Bengtsson, L., Andrae, U., Aspelien, T., Batrak, Y., Calvo, J., de Rooy, W., Gleeson, E., Hansen-Sass, B., Homleid, M., Hortal, M., Ivarsson, K., Lenderink, G., Niemelä, S., Pagh Nielsen, K., Onvlee, J., Rontu, L., Samuelsson, P., Santos Muñoz, D., Subias, A., Tijm, S., Toll, V., Yang, X. and Ødegaard Køltzow, M.: The HARMONIE-AROME model configuration in the ALADIN-HIRLAM NWP system. Monthly Wea. Rev., 145, 1919–1935, https://doi.org/10.1175/MWR-D-16-0417.1, 2018.

Cheng, B, Vihma. T, Rontu, L., Kontu A., Kheyrollah Pour, H., Duguay, C. and Pulliainen, J.: Evolution of snow and ice temperature, thickness and energy balance in Lake Orajärvi, northern Finland. Tellus A: Dynamic Meteorology and Oceanography ,66, 21564, https://doi.org/10.3402/tellusa.v66.21564, 2014.

Choulga, M., Kourzeneva, E., Zakharova, E. and Doganovsky, A.: Estimation of the mean depth of boreal lakes for use in numerical weather prediction and climate modelling. Tellus A, 66, 21295, https://doi.org/10.3402/tellusa.v66.21295, 2014.

Cordeira, J. M. and Laird, N. F. 2008. The influence of ice cover on two lake-effect snow events over lake Erie. *Mon. Wea. Rev.* **136**, 2747–2763. https://doi.org/DOI: 10.1175/2007MWR2310.1.

Elo, A-R.: Effects of climate and morphology on temperature conditions of lakes. University of Helsinki, Division of Geophysics, Report series in Geophysics. Available at http://urn.fi/URN:ISBN:978-952-10-3745-0, 2007.

Eerola, K., Rontu, L., Kourzeneva, E. and Shcherbak, E.: A study on effects of lake temperature and ice cover in HIRLAM. Boreal Env. Res., 15, 130–142, 2010.

Eerola, K.: Twenty-one years of verification from the HIRLAM NWP system. Wea. Forecasting, 28, 270–285. https://doi.org/10.1175/WAF-D-12-00068.1, 2013.

Eerola, K., Rontu, L., Kourzeneva, E., Kheyrollah Pour, H. and Duguay, C.: Impact of partly ice-free Lake Ladoga on temperature and cloudiness in an anticyclonic winter situation-a case study using a limited area model. Tellus A, 66, 23929. https://doi.org/10.3402/tellusa.v66.23929, 2014.

Gandin, L.: Objective analysis of meteorological fields. Gidrometizdat, Leningrad. Translated from Russian, Jerusalem, Israel Program for Scientific Translations, 1965.

Kheyrollah Pour, H., Rontu, L., Duguay, C. R., Eerola, K. and Kourzeneva, E.: Impact of satellite-based lake surface observations on the initial state of HIRLAM. Part II: Analysis of lake surface temperature and ice cover. Tellus A., 66, 21395. https://doi.org/10.3402/tellusa.v66.21395, 2014.

Kheyrollah Pour, H., Choulga, M., Eerola, K., Kourzeneva, E., Rontu, L., Pan F. and Duguay C. R.: Towards improved objective analysis of lake surface water temperature in a NWP model: preliminary assessment of statistical properties. Tellus A., 66, 21534. https://doi.org/10.1080/16000870.2017.1313025, 2017.

Korhonen J.: Suomen vesistöjen jääolot. Suomen ympäristö, (751), 145 pp (in Finnish). Available at http://hdl.handle.net/10138/40687, 2005.

Korhonen J.: Suomen vesistöjen lämpötilaolot 1900-luvulla. Suomen ympäristö, (566), 116 pp (in Finnish). Available at http://hdl.handle.net/10138/40478, 2002.

Kourzeneva, E., Samuelsson, P., Ganbat, G. and Mironov, D.: Implementation of lake model Flake into HIRLAM. HIRLAM Newsletter, (54), 54–64. Available at http://hirlam.org/, 2008.

References

Baijnath-Rodino, J. and Duguay, C.: Assessment of coupled CRCM5–FLake on the reproduction of wintertime lake-induced precipitation in the Great Lakes Basin. Theoretical and Applied Climatology https://doi.org/10.1007/s00704-019-02799-8, 2019.

Balsamo, G., Salgado, R., Dutra, E., Boussetta, S., Stockdale, T. and Potes, M: On the contribution of lakes in predicting near-surface temperature in a global weather forecasting model. Tellus A, 64, 15829, https://doi.org/10.3402/tellusa.v64i0.15829, 2012.

Batrak, Y., Kourzeneva, E. and Homleid, M.: Implementation of a simple thermodynamic sea ice scheme, SICE version 1.0-38h1, within the ALADIN–HIRLAM numerical weather prediction system version 38h1. Geosci. Model Dev., 11, 3347–3368, https://doi.org/10.5194/gmd-11-3347-2018, 2018.

Bengtsson, L., Andrae, U., Aspelien, T., Batrak, Y., Calvo, J., de Rooy, W., Gleeson, E., Hansen-Sass, B., Homleid, M., Hortal, M., Ivarsson, K., Lenderink, G., Niemelä, S., Pagh Nielsen, K., Onvlee, J., Rontu, L., Samuelsson, P., Santos Muñoz, D., Subias, A., Tijm, S., Toll, V., Yang, X. and Ødegaard Køltzow, M.: The HARMONIE-AROME model configuration in the ALADIN-HIRLAM NWP system. Monthly Wea. Rev., 145, 1919–1935, https://doi.org/10.1175/MWR-D-16-0417.1, 2018.

: Boone, A., Samuelsson, P., Gollvik, S., Napoly, A., Jarlan, L., Brun, E. and Decharme, B.: The interactions between soil–biosphere–atmosphere land surface model with a multi-energy balance (ISBA-MEB) option in SURFEXv8 – Part 1: Model description. Geosci. Model Dev., 10, 843–872. https://doi.org/10.5194/gmd-10-843-2017, 2017.

Cheng, B, Vihma. T, Rontu, L., Kontu A., Kheyrollah Pour, H., Duguay, C. and Pulliainen, J.: Evolution of snow and ice temperature, thickness and energy balance in Lake Orajärvi, northern Finland. Tellus A: Dynamic Meteorology and Oceanography ,66, 21564, https://doi.org/10.3402/tellusa.v66.21564, 2014.

Choulga, M., Kourzeneva, E., Zakharova, E. and Doganovsky, A.: Estimation of the mean depth of boreal lakes for use in numerical weather prediction and climate modelling. Tellus A, 66, 21295, https://doi.org/10.3402/tellusa.v66.21295, 2014.

Cordeira, J. M. and Laird, N. F. 2008. The influence of ice cover on two lake-effect snow events over lake Erie. *Mon. Wea. Rev.* **136**, 2747–2763. https://doi.org/DOI: 10.1175/2007MWR2310.1.

Elo, A-R.: Effects of climate and morphology on temperature conditions of lakes. University of Helsinki, Division of Geophysics, Report series in Geophysics. Available at http://urn.fi/URN:ISBN:978-952-10-3745-0, 2007.

Eerola, K., Rontu, L., Kourzeneva, E. and Shcherbak, E.: A study on effects of lake temperature and ice cover in HIRLAM. Boreal Env. Res., 15, 130–142, 2010.

Eerola, K.: Twenty-one years of verification from the HIRLAM NWP system. Wea. Forecasting, 28, 270–285. https://doi.org/10.1175/WAF-D-12-00068.1, 2013.

Eerola, K., Rontu, L., Kourzeneva, E., Kheyrollah Pour, H. and Duguay, C.: Impact of partly ice-free Lake Ladoga on temperature and cloudiness in an anticyclonic winter situation-a case study using a limited area model. Tellus A, 66, 23929. https://doi.org/10.3402/tellusa.v66.23929, 2014.

Gandin, L.: Objective analysis of meteorological fields. Gidrometizdat, Leningrad. Translated from Russian, Jerusalem, Israel Program for Scientific Translations, 1965.

Kheyrollah Pour, H., Rontu, L., Duguay, C. R., Eerola, K. and Kourzeneva, E.: Impact of satellite-based lake surface observations on the initial state of HIRLAM. Part II: Analysis of lake surface temperature and ice cover. Tellus A., 66, 21395. https://doi.org/10.3402/tellusa.v66.21395, 2014.

Kourzeneva, E., Asensio, H., Martin, E. and Faroux, S.: Global gridded dataset of lake coverage and lake depth for use in numerical weather prediction and climate modelling. Tellus A., 64,15640. https://doi.org/10.3402/tellusa.v64i0.15640, 2012a.

Kourzeneva, E., Martin, E., Batrak, Y. and Moigne, P. L.:. Climate data for parameterisation of lakes in numerical weather prediction models. Tellus A., 64, 17226. https://doi.org/10.3402/tellusa.v64i0.17226, 2012b.

Kourzeneva, E.: Assimilation of lake water surface temperature observations with Extended Kalman filter. Tellus A., 66, 21510. https://doi.org/10.3402/tellusa.v66.21510, 2014.

Laird, N. F., Kristovich, D. A. R. and Walsh, J. E.: Idealized model simulations examining the mesoscale structure of winter lake-effect circulations. *Mon. Wea. Rev.* **131**, 206–221. https://doi.org/DOI: 10.1175/1520-0493(2003)131<0206:IMSETM>2.0.CO;2, 2003.

Leppäranta, M., Lindgren, E. and Shirasawa, K.: The heat budget of Lake Kilpisjarvi in the Arctic tundra. Hydrology research, 48, 969-980. https://doi.org/10.2166/nh.2016.171, 2012.

Lei, R., Leppäranta, M., Cheng, B., Heil P. and Li, Z.: Changes in ice-season characteristics of a European Arctic lake from 1964 to 2008. Climatic Change, 115, 725–739. https://doi.org/10.1007/s10584-012-0489-2, 2012.

Lindstedt, D., P. Lind, E. Kjellström and C. Jones.: A new regional climate model operating at the meso-gamma scale: performance over Europe. Tellus A, 67, 24138. https://doi.org/10.3402/tellusa.v67.24138, 2015.

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, J. and co authors: Development of a global land cover characteristics database and IGBP DISCover from 1-km AVHRR data. Int. J. Rem. Sens., 21, 1303–1130. Data and documentation available online at: http://edc2.usgs.gov/glcc/glcc.php, 2000.

Masson, V., P. Le Moigne, E. Martin, S. Faroux, S., Alias, A., Alkama, R., Belamari, S., Barbu, A., Boone, A., Bouyssel, F., Brousseau, P., Brun, E., Calvet, J.-C., Carrer, D., Decharme, B., Delire, C., Donier, S., Essaouini, K., Gibelin, A.-L., Giordani, H., Habets, F., Jidane, M., Kerdraon, G., Kourzeneva, E., Lafaysse, M., Lafont, S., Lebeaupin Brossier, C., Lemonsu, A., Mahfouf, J.-F., Marguinaud, P., Mokhtari, M., Morin, S., Pigeon, G., Salgado, R., Seity, Y., Taillefer, F., Tanguy, G., Tulet, P., Vincendon, B., Vionnet, V., and Voldoire, A.: The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes, Geosci. Model Dev., 6, 929–960. https://doi.org/:10.5194/gmd-6-929-2013, 2013.

Mironov, D., Heise, E., Kourzeneva, E., Ritter, B., Schneider, N. and Terzhevik, A.: Implementation of the lake parameterisation scheme FLake into the numerical weather prediction model COSMO. Boreal Env. Res., 15, 218–230, 2010.

Pietikäinen, J.-P., Markkanen, T., Sieck, K., Jacob, D., Korhonen, J., Räisänen, P., Gao, Y., Ahola, J., Korhonen, H., Laaksonen, A., and Kaurola, J.: The regional climate model REMO (v2015) coupled with the 1-D freshwater lake model FLake (v1): Fenno-Scandinavian climate and lakes. Geosci. Model Dev., 11, 1321–1342, https://doi.org/10.5194/gmd-11-1321-2018, 2018.

Rontu, L., Eerola, K., Kourzeneva, E. and Vehviläinen, B.: Data assimilation and parametrisation of lakes in HIRLAM. Tellus A., 64, 17611. https://doi.org/10.3402/tellusa.v64i0.17611, 2012.

Rontu L., Gleeson, E., Räisänen, P., Nielsen, K.P., Savijärvi, H. and Sass, B.H.: The HIRLAM fast radiation scheme for mesoscale numerical weather prediction models. Adv.Sci.Res., 14, 195–215, https://doi.org/10.5194/asr-14-195-2017, 2017.

Rooney, G. G. and Bornemann, F. J.: The performance of FLake in the Met Office Unified Model. Tellus A, 65, 21363, https://doi.org/10.3402/tellusa.v65i0.21363, 2013.

Samuelsson, P., Kourzeneva, E. and Mironov, D.: The impact of lakes on the European climate as stimulated by a regional climate model. Boreal Env. Res., 15, 113–129, 2010.

SYKE. ISO 19115 Geographic Information Metadata on Hydrological data. Available at http://metatieto.ymparisto.fi:8080/geoportal/catalog/search/resource/details.page?uuid=%7B86FC3188-6796-4C79-AC58-8DBC7B568827%7D, 2018.

Kheyrollah Pour, H., Choulga, M., Eerola, K., Kourzeneva, E., Rontu, L., Pan F. and Duguay C. R.: Towards improved objective analysis of lake surface water temperature in a NWP model: preliminary assessment of statistical properties. Tellus A., 66, 21534. https://doi.org/10.1080/16000870.2017.1313025, 2017.

Korhonen J.: Long-term changes and variability of the winter and spring season hydrological regime in Finland. Report series in Geophysics, No. 79, 83 pp. University of Helsinki, Faculty of science, Institute for atmospheric and earth system research. Available at: https://helda.helsinki.fi/bitstream/handle/10138/298308/longterm.pdf, 2019.

Kourzeneva, E., Samuelsson, P., Ganbat, G. and Mironov, D.: Implementation of lake model Flake into HIRLAM. HIRLAM Newsletter, (54), 54–64. Available at http://hirlam.org/, 2008.

Kourzeneva, E., Asensio, H., Martin, E. and Faroux, S.: Global gridded dataset of lake coverage and lake depth for use in numerical weather prediction and climate modelling. Tellus A., 64,15640. https://doi.org/10.3402/tellusa.v64i0.15640, 2012a.

Kourzeneva, E., Martin, E., Batrak, Y. and Moigne, P. L.:. Climate data for parameterisation of lakes in numerical weather prediction models. Tellus A., 64, 17226. https://doi.org/10.3402/tellusa.v64i0.17226, 2012b.

Kourzeneva, E.: Assimilation of lake water surface temperature observations with Extended Kalman filter. Tellus A., 66, 21510. https://doi.org/10.3402/tellusa.v66.21510, 2014.

Laird, N. F., Kristovich, D. A. R. and Walsh, J. E.: Idealized model simulations examining the mesoscale structure of winter lake-effect circulations. *Mon. Wea. Rev.* **131**, 206–221. https://doi.org/DOI: 10.1175/1520-0493(2003)131<0206:IMSETM>2.0.CO;2, 2003.

Leppäranta, M., Lindgren, E. and Shirasawa, K.: The heat budget of Lake Kilpisjarvi in the Arctic tundra. Hydrology research, 48, 969-980. https://doi.org/10.2166/nh.2016.171, 2012.

Lei, R., Leppäranta, M., Cheng, B., Heil P. and Li, Z.: Changes in ice-season characteristics of a European Arctic lake from 1964 to 2008. Climatic Change, 115, 725–739. https://doi.org/10.1007/s10584-012-0489-2, 2012.

Lindstedt, D., P. Lind, E. Kjellström and C. Jones.: A new regional climate model operating at the meso-gamma scale: performance over Europe. Tellus A, 67, 24138. https://doi.org/10.3402/tellusa.v67.24138, 2015.

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, J. and co authors: Development of a global land cover characteristics database and IGBP DISCover from 1-km AVHRR data. Int. J. Rem. Sens., 21, 1303–1130. Data and documentation available online at: http://edc2.usgs.gov/glcc/glcc.php, 2000.

Masson, V., P. Le Moigne, E. Martin, S. Faroux, S., Alias, A., Alkama, R., Belamari, S., Barbu, A., Boone, A., Bouyssel, F., Brousseau, P., Brun, E., Calvet, J.-C., Carrer, D., Decharme, B., Delire, C., Donier, S., Essaouini, K., Gibelin, A.-L., Giordani, H., Habets, F., Jidane, M., Kerdraon, G., Kourzeneva, E., Lafaysse, M., Lafont, S., Lebeaupin Brossier, C., Lemonsu, A., Mahfouf, J.-F., Marguinaud, P., Mokhtari, M., Morin, S., Pigeon, G., Salgado, R., Seity, Y., Taillefer, F., Tanguy, G., Tulet, P., Vincendon, B., Vionnet, V., and Voldoire, A.: The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes, Geosci. Model Dev., 6, 929–960. https://doi.org/:10.5194/gmd-6-929-2013, 2013.

Mironov, D., Heise, E., Kourzeneva, E., Ritter, B., Schneider, N. and Terzhevik, A.: Implementation of the lake parameterisation scheme FLake into the numerical weather prediction model COSMO. Boreal Env. Res., 15, 218–230, 2010.

Pietikäinen, J.-P., Markkanen, T., Sieck, K., Jacob, D., Korhonen, J., Räisänen, P., Gao, Y., Ahola, J., Korhonen, H., Laaksonen, A., and Kaurola, J.: The regional climate model REMO (v2015) coupled with the 1-D freshwater lake model FLake (v1): Fenno-Scandinavian climate and lakes. Geosci. Model Dev., 11, 1321–1342, https://doi.org/10.5194/gmd-11-1321-2018, 2018.

Rontu, L., Eerola, K., Kourzeneva, E. and Vehviläinen, B.: Data assimilation and parametrisation of lakes in HIRLAM. Tellus A., 64, 17611. https://doi.org/10.3402/tellusa.v64i0.17611, 2012.

Termonia, P., Fischer, C., Bazile, E., Bouyssel, F., Brožková, R., Bénard, P., Bochenek, B., Degrauwe, D., Derkova, M., El Khatib, R., Hamdi, R., Mašek, J., Pottier, P., Pristov, N., Seity, Y., Smolíková, P., Spaniel, O., Tudor, M., Wang, Y., Wittmann, C., and Joly, A.: The ALADIN System and its canonical model configurations AROME CY41T1 and ALARO CY40T1, Geosci. Model Dev., 11, 257-281, https://doi.org/10.5194/gmd-11-257-2018, 2018.

Undén, P., Rontu, L., Järvinen, H., Lynch, P., Calvo, J., Cats, G., Cuxart, J., Eerola, K., Fortelius, C., Garcia-Moya, J. A. et al.: HIRLAM-5 scientific documentation. Available at http://www.hirlam.org, 2002.

Yang, Y., Cheng, B., Kourzeneva, E., Semmler, T., Rontu, L., Leppäranta, M., Shirasawa, K. and Li, Z. J.: Modelling experiments on air–snow–ice interactions over Kilpisjärvi, a lake in northern Finland. Boreal Env. Res. 18: 341–358, 2013.

24

Rontu L., Gleeson, E., Räisänen, P., Nielsen, K.P., Savijärvi, H. and Sass, B.H.: The HIRLAM fast radiation scheme for mesoscale numerical weather prediction models. Adv.Sci.Res., 14, 195–215, https://doi.org/10.5194/asr-14-195-2017, 2017.

Rooney, G. G. and Bornemann, F. J.: The performance of FLake in the Met Office Unified Model. Tellus A, 65, 21363, https://doi.org/10.3402/tellusa.v65i0.21363, 2013.

. Samuelsson P., Gollvik, S., and Ullerstig, A.: The land-surface scheme of the Rossby Centre regional atmospheric climate model (RCA3). Report in Meteorology 122, SMHI, SE-60176 Norrköping, Sweden, 2006.

Samuelsson, P., Kourzeneva, E. and Mironov, D.: The impact of lakes on the European climate as stimulated by a regional climate model. Boreal Env. Res., 15, 113–129, 2010.

Samuelsson, P., Jones, C., Willén, U., Ullerstig, A., Gollvik, S., Hansson, U., Jansson, C., Kjellström, E., Nikulin, G., and Wyser, K.: The Rossby Centre Regional Climate Model RCA3: Model description and performance, Tellus A, 63, 1–3, https://doi.org/10.1111/j.1600-0870.2010.00478.x, 2011.

Semmler, T., B. Cheng, Y. Yang, and L. Rontu. Snow and ice on Bear Lake (Alaska) – sensitivity experiments with two lake ice models. Tellus A, 64, 17339, https://doi.org/10.3402/tellusa.v64i0.17339, 2012.

SYKE. ISO 19115 Geographic Information Metadata on Hydrological data. Available at http://metatieto.ymparisto.fi:8080/geoportal/catalog/search/resource/details.page?uuid=%7B86FC3188-6796-4C79-AC58-8DBC7B568827%7D, 2018.

Termonia, P., Fischer, C., Bazile, E., Bouyssel, F., Brožková, R., Bénard, P., Bochenek, B., Degrauwe, D., Derkova, M., El Khatib, R., Hamdi, R., Mašek, J., Pottier, P., Pristov, N., Seity, Y., Smolíková, P., Spaniel, O., Tudor, M., Wang, Y., Wittmann, C., and Joly, A.: The ALADIN System and its canonical model configurations AROME CY41T1 and ALARO CY40T1, Geosci. Model Dev., 11, 257-281, https://doi.org/10.5194/gmd-11-257-2018, 2018.

Undén, P., Rontu, L., Järvinen, H., Lynch, P., Calvo, J., Cats, G., Cuxart, J., Eerola, K., Fortelius, C., Garcia-Moya, J. A. et al.: HIRLAM-5 scientific documentation. Available at http://www.hirlam.org, 2002.

Yang, Y., Cheng, B., Kourzeneva, E., Semmler, T., Rontu, L., Leppäranta, M., Shirasawa, K. and Li, Z. J.: Modelling experiments on air–snow–ice interactions over Kilpisjärvi, a lake in northern Finland. Boreal Env. Res. 18: 341–358, 2013.

Table A1. Prognostic and diagnostic lake variables within HIRLAM

variable	unit	type
temperature of snow on lake ice	K	prog by FLake
temperature of lake ice	K	prog by FLake
mean water temperature	K	prog by FLake
mixed layer temperature	K	prog by FLake
bottom temperature	K	prog by FLake
temperature of upper layer sediments	K	prog by FLake
mixed layer depth	m	prog by FLake
thickness of upper layer sediments	m	prog by FLake
thermocline shape factor	-	prog by FLake
lake ice thickness	m	prog by FLake
snow depth on lake ice	m	prog by FLake
LSWT	K	diag by FLake
		= mixed layer temperature if no ice
lake surface temperature	K	diag by FLake
		uppermost temperature: LSWT or ice or snow
LSWT	K	anal by HIRLAM
		flag value 272 K when there is ice
fraction of lake ice	[0-1]	diag fraction in HIRLAM grid
lake surface roughness	m	diag by HIRLAM
screen level temperature over lake	K	diag by HIRLAM
screen level abs.humidity over lake	$kgkg^{-1}$	diag by HIRLAM
anemometer level u-component over lake	ms^{-1}	diag by HIRLAM
anemometer level v-component over lake	ms^{-1}	diag by HIRLAM
latent heat flux over lake	Wm^{-2}	diag by HIRLAM
sensible heat flux over lake	Wm^{-2}	diag by HIRLAM
scalar momentum flux over lake	Pa	diag by HIRLAM
SW net radiation over lake	Wm^{-2}	diag by HIRLAM
LW net radiation over lake	Wm^{-2}	diag by HIRLAM
depth of lake	m	pres in HIRLAM grid
fraction of lake	[0-1]	pres in HIRLAM grid

Denotation: prog = prognostic, diag = diagnostic, pres = prescribed, anal = result of OI

Table A1. Prognostic and diagnostic lake variables within HIRLAM

variable	unit	type
temperature of snow on lake ice	K	prog by FLake
temperature of lake ice	K	prog by FLake
mean water temperature	K	prog by FLake
mixed layer temperature	K	prog by FLake
bottom temperature	K	prog by FLake
temperature of upper layer sediments	K	prog by FLake
mixed layer depth	m	prog by FLake
thickness of upper layer sediments	m	prog by FLake
thermocline shape factor	-	prog by FLake
lake ice thickness	m	prog by FLake
snow depth on lake ice	m	prog by FLake
LSWT	K	diag by FLake
		= mixed layer temperature if no ice
lake surface temperature	K	diag by FLake
		uppermost temperature: LSWT or ice or snow
LSWT	K	anal by HIRLAM
		flag value 272 K when there is ice
fraction of lake ice	[0-1]	diag fraction in HIRLAM grid
lake surface roughness	m	diag by HIRLAM
screen level temperature over lake	K	diag by HIRLAM
screen level abs.humidity over lake	$kgkg^{-1}$	diag by HIRLAM
anemometer level u-component over lake	ms^{-1}	diag by HIRLAM
anemometer level v-component over lake	ms^{-1}	diag by HIRLAM
latent heat flux over lake	Wm^{-2}	diag by HIRLAM
sensible heat flux over lake	Wm^{-2}	diag by HIRLAM
scalar momentum flux over lake	Pa	diag by HIRLAM
SW net radiation over lake	Wm^{-2}	diag by HIRLAM
LW net radiation over lake	Wm^{-2}	diag by HIRLAM
depth of lake	m	pres in HIRLAM grid
fraction of lake	[0-1]	pres in HIRLAM grid

Denotation: prog = prognostic, diag = diagnostic, pres = prescribed, anal = result of OI

Table A2. Lakes with SYKE observations used in this study.

NAME	LON	LAT	MEAND (m)	MAXD (m)	AREA (kgm^{-2})	HIRD (m)	HIRFR	HIRID
Pielinen	29.607	63.271	10.1	61.0	894.2	10.0	0.916	4001
Kallavesi	27.783	62.762	9.7	75.0	316.1	10.0	0.814	4002
Haukivesi	28.389	62.108	9.1	55.0	560.4	10.0	0.725	4003
Saimaa	28.116	61.338	10.8	85.8	1,377.0	10.0	0.950	4004
Pääjärvi1	24.789	62.864	3.8	14.9	29.5	3.0	0.430	4005
Nilakka	26.527	63.115	4.9	21.7	169.0	10.0	0.866	4006
Konnevesi	26.605	62.633	10.6	57.1	189.2	10.0	0.937	4007
Jääsjärvi	26.135	61.631	4.6	28.2	81.1	10.0	0.750	4008
Päijänne	25.482	61.614	14.1	86.0	864.9	10.0	0.983	4009
Ala-Rieveli	26.172	61.303	11.3	46.9	13.0	10.0	0.549	4010
Kyyvesi	27.080	61.999	4.4	35.3	130.0	10.0	0.810	4011
Tuusulanjärvi	25.054	60.441	3.2	9.8	5.9	3.0	0.174	4012
Pyhäjärvi	22.291	61.001	5.5	26.2	155.2	5.0	0.922	4013
Längelmävesi	24.370	61.535	6.8	59.3	133.0	10.0	0.875	4014
Pääjärvi2	25.132	61.064	14.8	85.0	13.4	14.0	0.350	4015
Vaskivesi	23.764	62.142	7.0	62.0	46.1	10.0	0.349	4016
Kuivajärvi	23.860	60.786	2.2	9.9	8.2	10.0	0.419	4017
Näsijärvi	23.750	61.632	14.7	65.6	210.6	10.0	0.850	4018
Lappajärvi	23.671	63.148	6.9	36.0	145.5	10.0	1.000	4019
Pesiöjärvi	28.650	64.945	3.9	15.8	12.7	7.0	0.290	4020
Rehja-Nuasjärvi	28.016	64.184	8.5	42.0	96.4	10.0	0.534	4021
Oulujärvi	26.965	64.451	6.9	35.0	887.1	10.0	1.000	4022
Ounasjärvi	23.602	68.377	6.6	31.0	6.9	10.0	0.166	4023
Unari	25.711	67.172	5.0	24.8	29.1	10.0	0.491	4024
Kilpisjärvi	20.816	69.007	19.5	57.0	37.3	22.0	0.399	4025
Kevojärvi	27.011	69.754	11.1	35.0	1.0	10.0	0.016	4026
Inarijärvi	27.924	69.082	14.3	92.0	1,039.4	14.0	0.979	4027

Denotation: LON and LAT are the longitude E and latitude N in degrees, MEAND and MAXD are the mean and maximum depths and AREA is the water surface area from the updated lake list of GLDB v.3 (Margarita Choulga, personal communication), HIRD and HIRFR are the mean lake depth and fraction of lakes [0…1] interpolated to the selected HIRLAM gridpoint, taken from the operational HIRLAM that uses GLDB v.2 as the source for lake depths. HIRID is the lake index used by HIRLAM and in this study. Above the middle line are the 27 lakes with both LSWT and LID observations, below the 18 lakes where only LID was available.

Table A2. Lakes with SYKE observations used in this study.

NAME	LON	LAT	MEAND (m)	MAXD (m)	AREA (kgm^{-2})	HIRD (m)	HIRFR	HIRID
Pielinen	29.607	63.271	10.1	61.0	894.2	10.0	0.916	4001
Kallavesi	27.783	62.762	9.7	75.0	316.1	10.0	0.814	4002
Haukivesi	28.389	62.108	9.1	55.0	560.4	10.0	0.725	4003
Saimaa	28.116	61.338	10.8	85.8	1,377.0	10.0	0.950	4004
Pääjärvi1	24.789	62.864	3.8	14.9	29.5	3.0	0.430	4005
Nilakka	26.527	63.115	4.9	21.7	169.0	10.0	0.866	4006
Konnevesi	26.605	62.633	10.6	57.1	189.2	10.0	0.937	4007
Jääsjärvi	26.135	61.631	4.6	28.2	81.1	10.0	0.750	4008
Päijänne	25.482	61.614	14.1	86.0	864.9	10.0	0.983	4009
Ala-Rieveli	26.172	61.303	11.3	46.9	13.0	10.0	0.549	4010
Kyyvesi	27.080	61.999	4.4	35.3	130.0	10.0	0.810	4011
Tuusulanjärvi	25.054	60.441	3.2	9.8	5.9	3.0	0.174	4012
Pyhäjärvi	22.291	61.001	5.5	26.2	155.2	5.0	0.922	4013
Längelmävesi	24.370	61.535	6.8	59.3	133.0	10.0	0.875	4014
Pääjärvi2	25.132	61.064	14.8	85.0	13.4	14.0	0.350	4015
Vaskivesi	23.764	62.142	7.0	62.0	46.1	10.0	0.349	4016
Kuivajärvi	23.860	60.786	2.2	9.9	8.2	10.0	0.419	4017
Näsijärvi	23.750	61.632	14.7	65.6	210.6	10.0	0.850	4018
Lappajärvi	23.671	63.148	6.9	36.0	145.5	10.0	1.000	4019
Pesiöjärvi	28.650	64.945	3.9	15.8	12.7	7.0	0.290	4020
Rehja-Nuasjärvi	28.016	64.184	8.5	42.0	96.4	10.0	0.534	4021
Oulujärvi	26.965	64.451	6.9	35.0	887.1	10.0	1.000	4022
Ounasjärvi	23.602	68.377	6.6	31.0	6.9	10.0	0.166	4023
Unari	25.711	67.172	5.0	24.8	29.1	10.0	0.491	4024
Kilpisjärvi	20.816	69.007	19.5	57.0	37.3	22.0	0.399	4025
Kevojärvi	27.011	69.754	11.1	35.0	1.0	10.0	0.016	4026
Inarijärvi	27.924	69.082	14.3	92.0	1,039.4	14.0	0.979	4027

Denotation: LON and LAT are the longitude E and latitude N in degrees, MEAND and MAXD are the mean and maximum depths and AREA is the water surface area from the updated lake list of GLDB v.3 (Margarita Choulga, personal communication), HIRD and HIRFR are the mean lake depth and fraction of lakes [0…1] interpolated to the selected HIRLAM gridpoint, taken from the operational HIRLAM that uses GLDB v.2 as the source for lake depths. HIRID is the lake index used by HIRLAM and in this study. Above the middle line are the 27 lakes with both LSWT and LID observations, below the 18 lakes where only LID was available.

Table A3. Lakes with SYKE observations used in this study. Part 2

NAME	LON	LAT	MEAND (m)	MAXD (m)	AREA (kgm^{-2})	HIRD (m)	HIRFR	HIRID
Simpelejärvi	29.482	61.601	9.3	34.4	88.2	10.0	0.548	40241
Pökkäänlahti	27.264	61.501	8.0	84.3	58.0	10.0	0.299	40261
Muurasjärvi	25.353	63.478	9.0	35.7	21.1	10.0	0.060	40263
Kalmarinselkä	25.001	62.786	5.7	21.9	7.1	5.0	0.330	40271
Summasjärvi	25.344	62.677	6.7	40.5	21.9	10.0	0.555	40272
Iisvesi	27.021	62.679	17.2	34.5	164.9	18.0	0.456	40277
Hankavesi	26.826	62.614	7.0	49.0	18.2	18.0	0.100	40278
Petajävesi	25.173	62.255	4.2	26.6	8.8	3.0	0.245	40282
Kukkia	24.618	61.329	5.2	35.6	43.9	10.0	0.299	40308
Ähtärinjärvi	24.045	62.755	5.2	27.0	39.9	10.0	0.266	40313
Kuortaneenjärvi	23.407	62.863	3.3	16.2	14.9	10.0	0.277	40328
Lestijärvi	24.716	63.584	3.6	6.9	64.7	10.0	0.513	40330
Pyhäjärvi	25.995	63.682	6.3	27.0	121.8	10.0	0.266	40331
Lentua	29.690	64.204	7.4	52.0	77.8	7.0	0.600	40335
Lammasjärvi	29.551	64.131	4.3	21.0	46.8	3.0	0.200	40336
Naamankajärvi	28.246	65.104	2.9	14.0	8.5	7.0	0.299	40342
Korvuanjärvi	28.663	65.348	6.0	37.0	15.4	10.0	0.342	40343
Oijärvi	25.930	65.621	1.1	2.4	21.0	10.0	0.333	40345

Denotation: LON and LAT are the longitude E and latitude N in degrees, MEAND and MAXD are the mean and maximum depths and AREA is the water surface area from the updated lake list of GLDB v.3 (Margarita Choulga, personal communication), HIRD and HIRFR are the mean lake depth and fraction of lakes [0…1] interpolated to the selected HIRLAM gridpoint, taken from the operational HIRLAM that uses GLDB v.2 as the source for lake depths. HIRID is the lake index used by HIRLAM and in this study. Above the middle line are the 27 lakes with both LSWT and LID observations, below the 18 lakes where only LID was available.

Second reply to the three reviewers

This manuscript has evolved from the original version of the 29th October 2018, which is available at the manuscript page, to the first revision of the 20th February 2019 and to the second revision of the 21st of March 2019.

Reviewers comments to the original revision were replied in detail on 20th February, available at the manuscript page. The revised manuscript is not available there.

Editor's comments on 21st February to the first revision were replied on 21 March 2019. The editor's comments, our reply and the editors reply are available at the manuscript page. A second revision was written in responce to the editors comments. It is not available at the manuscript page.

The essential differences between the first and second revisions were listed in the reply to the editor:

"1) Through the manuscript, we corrected our unfortunate formulations related to "bug" and "technical error" in order to avoid creating misunderstandings. 2) We wrote a short discussion section about snow on lake ice, with proper references with respect to the current results. We coordinated the conclusion section and the new discussion section to avoid overlap. 3) We checked the whole manuscript in order to make it crystal clear for the editors, reviewers and readers that we are validating operational model results, aiming at detecting problems and suggesting improvements for further developments. At the same time, we took the opportunity to modify and add a few references, improve the terminology concerning the lake ice melting and freezing and make a few minor text corrections."

In our first reply, we did not react to the second reviewer's opinion that the results and discussions in this paper are of very limited interest due to a bug in treatment snow accumulation on ice. We presented our viewpoint on this in the reply to the editor on the 21st of March.

The terminology on lake ice melting and freezing dates was modified as originally suggested by the first reviewer, to melt-up and freeze-up, with a reference to Korhonen, 2019. In our first reply, we left this question open.

We introduced remarks related to the snow bug into several places of the first revision, as was also suggested by the reviewers. In the second revision, these are now mostly placed in a new discussion section about snow on ice instead of the scattered remarks.

About the technical side of FMI operational HIRLAM. Our suggested snow-on-ice correction was implemented into operations only 4.3.2019, not in October 2018 as we assumed. This means that it did not yet have time to influence properly in the results of lake ice melting this spring. At FMI, it has been decided that HIRLAM will be gradually decommissioned from operational usage during the next two years. This leaves the next winter 2019-2020 for the final proof of the influence of this modification in HIRLAM. However, we are convinced that the results of our unpublished experiment of January 2014, now mentioned in the new discussion section on snow on lake ice, did show that the modification works as expected.

To show the differences between the original, first and second revisions in a proper way, three difference pdf files were now generated by using latexdiff. They are: between the original and the first revision, between the original and the second revision and between the first and the second revisions.

Please find below the original replies to three reviewers, with a few sentences in this color and highlighting to show the differences made between the first and second revisions in the context of earlier replies.

We hope these remarks and additional comparison documents make the situation clear. We apologize for not preparing and providing all needed files directly to the reviewers via the manuscript pages.

The 20th of May 2019

Laura Rontu, Kalle Eerola, Matti Horttanainen

Original replies
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 documented 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. Especially 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. The operational correction was made on the 4th of March, 2019. thus this is not valid.

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 provide 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 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. This is not valid, we now use freeze-up and break-up dates, following Korhonen, 2019.

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.
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 operational correction was made on the 4th of March, 2019. thus this is not valid. We now discuss more in depth in the reply to the editor's comments why we disagree with your statement that "the corresponding results and discussions are of very limited interest". A new discussion section about snow on ice was added in the manuscript.

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 downwelling 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, T_{2m} 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 -0.5°C, none when it is 0°C 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. We have added a discussion section on snow on lake ice to explain this issue systematically.

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 temperature 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 ASSIMILATION 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 kmˆ-2" should be "with an area of 1 kmˆ2"
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.
Corrected

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. The operational correction was made on the 4th of March, 2019. thus this is not valid.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.

Validation of lake surface state in the HIRLAM v.7.4 NWP model against *in-situ* measurements in Finland

Laura Rontu[1], Kalle Eerola[1], and Matti Horttanainen[1]

[1]Finnish Meteorological Institute, P.O. Box 503, 00101 Helsinki, Finland

Correspondence: laura.rontu@fmi.fi

Abstract. High Resolution Limited Area Model (HIRLAM), used for the operational numerical weather prediction in the Finnish Meteorological Institute (FMI), includes prognostic treatment of lake surface state since 2012. Forecast is based on the Freshwater Lake (FLake) model integrated to into HIRLAM. Additionally, an independent objective analysis of lake surface water temperature (LSWT) combines the short forecast of FLake to observations from the Finnish Environment Institute (SYKE). The resulting description of lake surface state - forecast FLake variables and analysed LSWT - was compared to SYKE observations of lake water temperature, freezing and melting dates as well as the ice and snow thickness thickness and snow depth for 2012-2018 over 45 lakes in Finland. During the ice-free period, the predicted LSWT corresponded to the observations with a slight overestimation, with a systematic error of + 0.91 K. The colder temperatures were underrepresented and the maximum temperatures were too high. The objective analysis of LSWT was able to reduce the bias to + 0.35 K. The predicted freezing dates corresponded well the observed dates, mostly within the accuracy of a week. The forecast melting dates were far too early, typically several weeks ahead of the observed dates. The growth of ice thickness after freezing was generally overestimated. However, practically no predicted snow appeared on lake ice. The absence of snow, found to be due to a technical error in HIRLAM, is suggested to be also the reason of the inaccurate simulation of the lake ice melt in spring.

1 Introduction

Lakes influence the energy exchange between the surface and the atmosphere, the dynamics of the atmospheric boundary layer and the near-surface weather. This is important for weather forecasting over the areas where lakes, especially those with a large yearly variation of the water temperature, freezing in autumn and melting in spring, cover a significant area of the surface (Kheyrollah Pour et al., 2017; Laird et al., 2003 and references therein). Description of the lake surface state influences the numerical weather prediction (NWP) results, in particular in the models whose resolution is high enough to account for even the smaller lakes (Eerola et al., 2014 and references therein). Especially, the existence of ice can be important for the numerical forecast (Eerola et al., 2014; Cordeira and Laird, 2008).

~~The~~ In the Finnish Meteorological Institute (FMI), the High Resolution Limited Area Model HIRLAM (Undén et al., 2002; Eerola, 2013) has been applied since 1990 for the numerical short-range weather forecast~~over the Northern Europe~~. In the beginning, the monthly climatological water surface temperature for both sea (sea surface temperature SST) and lakes (Lake Surface Water Temperature LSWT) was used. Since 2012, HIRLAM includes a prognostic lake temperature parameterization based on the Freshwater Lake Model (FLake, Mironov et al., 2010). An independent objective analysis of observed LSWT (~~Kheyrollah Pour et al. (2017)~~ Kheyrollah Pour et al., 2017 and references therein) was implemented in 2011. ~~Fractional~~ The fractional ice cover (lake ice concentration in each ~~grid-square~~ gridsquare of the model) is ~~estimated separately based on~~ diagnosed from the analysed LSWT~~and the ice thickness predicted by FLake~~.

FLake was designed to be used as a parametrization scheme for the forecast of the lake surface state in NWP and climate models. It allows to predict the lake surface state in interaction with the atmospheric processes treated by the NWP model. The radiative and turbulent fluxes from the atmospheric model are combined with FLake processes at each time-step of the model integration ~~(with a typical interval of one or several minutes)~~ in the model grid, where the fraction and depth of lakes are prescribed.

FLake has been implemented into the main European NWP and regional climate models, first into COSMO (Mironov et al., 2010) then into ECMWF (Balsamo et al., 2012), Unified Model (Rooney and Bornemann, 2013), SURFEX surface modelling framework (Masson et al., 2016), regional climate models RCA (Samuelsson et al., 2010), HCLIM (Lindstedt et al., 2015) and REMO (Pietikäinen et al., 2018), among others. Description of lake surface state and its influence in the numerical weather and climate prediction has been validated in various ways. Results of case studies, e.g. Eerola et al. (2014) and shorter-period NWP experiments, e.g. Eerola et al. (2010); Rontu et al. (2012); Kheyrollah Pour et al. (2014); Kheyrollah Pour et al. (2017) as well as climate model results, e.g. ~~Samuelsson et al. (2010); Pietikäinen et al. (2018)~~ Samuelsson et al. (2010); Pietikäinen et al. (2018), have been compared with remote-sensing satellite data and ~~*in situ*~~*in-situ* lake temperature and ice measurements as well as validated against the standard weather observations. In general, improvement of the scores has been seen over regions where lakes occupy a significant area. However, specific features of each of the host models influence the results of the coupled atmosphere-lake system as FLake appears to be quite sensitive to the forcing by the atmospheric model.

The aim of the present study is to use ~~*in situ*~~*in-situ* LSWT measurements, lake ice freezing and melting dates and measurements of ice and snow thickness by the Finnish Environment Institute (Suomen Ympäristökeskus = SYKE) for validation of the lake surface state forecast by the operational HIRLAM NWP model. For this purpose, HIRLAM analyses and forecasts archived by the Finnish Meteorological Institute (FMI) were compared with the observations by SYKE over the lakes of Finland from spring 2012 to summer 2018. To our knowledge, this is the longest available detailed dataset that allows to evaluate how well the lake surface state is simulated by an operational NWP model that applies FLake parametrizations.

2 Lake surface state in HIRLAM

FLake was implemented in the HIRLAM forecasting system in 2012 (Kourzeneva et al., 2008; Eerola et al., 2010). The model utilizes external datasets on the lake depth (Kourzeneva et al., 2012a; Choulga et al., 2014) and the lake climatology (Kourzeneva et al., 2012b). The latter is only needed in order to provide initial values of FLake prognostic variables in the very first forecast (so-called cold start). Real-time *in-situ* LSWT observations by SYKE for 27 Finnish lakes ~~became available~~ were obtained in 2011 to be used for the operational ~~HIRLAM analysis in 2011~~ LSWT analysis in HIRLAM (Eerola et al., 2010; Rontu et al., 2012). In the current operational HIRLAM at FMI FLake provides the background for the optimal interpolation analysis (OI, based on ~~Gandin 1965) analysis~~ Gandin, 1965) of LSWT. However, the prognostic FLake variables are not corrected using the analysed LSWT~~, which would require~~. This would require more advanced data assimilation methods based on e.g. the extended Kalman filter (Kourzeneva, 2014). ~~The relations between the OI analysis and the prognostic FLake in HIRLAM are schematically illustrated in Figure 1.~~

~~Coexistence of the independent objective analysis of the observed LSWT and prognostic FLake parametrizations in HIRLAM.~~

2.1 Freshwater lake model in HIRLAM

FLake is a bulk model capable of predicting the vertical temperature structure and mixing conditions in lakes of various depths on time-scales from hours to years (Mironov et al., 2010). The model is based on two-layer parametric representation of the evolving temperature profile in the water and on the integral budgets of energy for the layers in question. Bottom sediments and the thermodynamics of the ice and snow on ice layers are treated separately. FLake depends on prescribed lake depth information. The prognostic and diagnostic variables of HIRLAM ~~/FLake plus~~ FLake together with the analysed lake surface variables in HIRLAM are listed in the Appendix (Table A1).

At each time step of the HIRLAM forecast, FLake is driven by the atmospheric radiative and turbulent fluxes provided by the physical parameterisations in HIRLAM. This couples the atmospheric variables over lakes with the lake surface properties as provided by FLake. 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.

Implementation of FLake model as a ~~parametrizations~~ parametrization scheme in HIRLAM was based on the experiments described by ~~Rontu et al. (2012)~~Rontu et al., 2012. Compared to the reference version of FLake (Mironov et al., 2010), minor modifications were introduced, namely, use of constant snow density = 300 $\mathrm{kgm^{-3}}$, molecular heat conductivity = 1 $\mathrm{Jm^{-1}s^{-1}K^{-1}}$, constant albedos of dry snow = 0.75 and ice = 0.5. Bottom sediment calculations were excluded. Global lake depth database (GLDB v.2, ~~Choulga et al. (2014)) is~~ Choulga et al., 2014) was used for derivation of mean lake depth in each gridsquare. Fraction of lake ~~is~~ was taken from HIRLAM physiography database, where it originates from GLCC (Loveland et al., 2000).

~~In this study, lake surface temperature and thickness of ice predicted by FLake were used for the model-observation comparison.~~ Lake surface temperature is diagnosed from the mixed layer temperature for the unfrozen lake gridpoints and from the ice or snow-on-ice temperature for the frozen points. In FLake, ice starts to grow from an assumed value of one millimeter when temperature reaches the freezing point. The whole lake tile in a gridsquare is considered by FLake either frozen or unfrozen. Snow on ice is accumulated from the model's snowfall at each time step during the numerical integration.

2.2 Objective analysis of LSWT observations

A comprehensive description of the optimal interpolation (OI) of the LSWT observations in HIRLAM is given by Kheyrollah Pour et al. (20 (Kheyrollah Pour et al., 2017). Shortly, LSWT analysis is obtained by correcting the FLake forecast at each gridpoint by using the weighted average of the deviations of observations from their background values. Prescribed statistical information about the observation and background error variance as well as the distance-dependent autocorrelation between the locations (observations and gridpoints) are applied. The real-time observations entering the HIRLAM surface analysis system are subject to quality control in two phases. First, the observations are compared to the background, provided by the FLake short forecast. Second, optimal interpolation is done at each observation location, using the neighbouring observations only (excluding the current observation) and comparing the result to the observed value at the station.

A specific feature of the lake surface temperature OI is that the interpolation is performed not only within the (large) lakes but also across the lakes: within a statistically pre-defined radius, the observations affect all gridpoints containing a fraction of lake. This ensures that the analysed LSWT on lakes without own observations may also be influenced by observations from neighbouring lakes, not only by the first guess provided by FLake forecast.

The relations between the OI analysis and the prognostic FLake in HIRLAM are schematically illustrated in Figure 1. Within the present HIRLAM setup, the background for the analysis is provided by the short (6-hour) FLake forecast . However, but the next forecast is not initialised initialized from the analysis, see Figure 1. Instead, FLake continues running from the previous forecast, driven by the atmospheric state given by HIRLAM at each time step. This means that FLake does not benefit from the result of OI analysis does not benefit FLake but the analysis remains to some extent as an extra diagnostic field, to some extent independent of the LSWT forecast. Note that However, FLake background has a large influence in the analysis, especially over distant lakes where neighbouring observations are not available. The diagnostic LSWT analysis, available at every gridpoint of HIRLAM, might be useful e.g. for hydrological, agricultural or road weather applications.

Missing LSWT observations in spring and early winter are interpreted to represent presence of ice and given a flag value of -1.2°C. If, however, the results of the statistical moving-average-type LSWT model (Elo, 2007), provided by SYKE along with the real-time observations, indicate unfrozen conditions, the observations are considered missing. This prevents appearance of ice in summer when observations are missing but leads to a misinterpretation of data in spring if the SYKE model indicates too early melting. In the analysis, fraction of ice is diagnosed from the LSWT field in a simple way. The lake surface within a gridsquare is assumed fully ice-covered when LSWT falls below -0.5°C and fully ice-free when LSWT is above 0°C. Between these temperature thresholds, the fraction of ice changes linearly (Kheyrollah Pour et al., 2014).

The HIRLAM surface data assimilation system produces comprehensive feedback information from every analysis-forecast cycle. The feedback consists of the observed value and its deviations from the background and from the final analysis at the observation point. Bilinear interpolation of the analysed and forecast values is done to the observation location from the

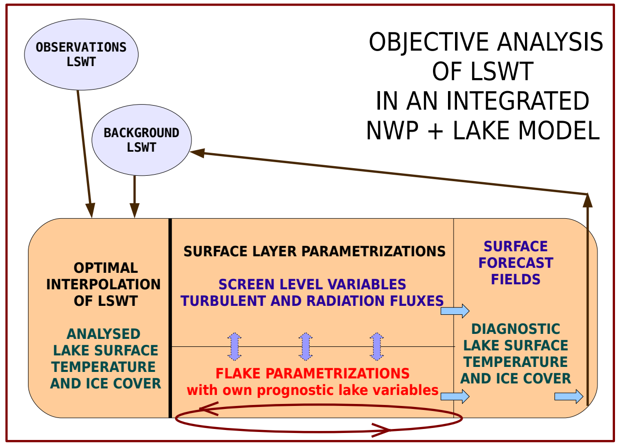


Figure 1. Coexistence of the independent objective analysis of the observed LSWT and prognostic FLake parametrizations in HIRLAM. The thin arrows are related to data flow between HIRLAM analysis-forecast cycles while the thick arrows describe processes within each cycle.

nearest gridpoints that contain a fraction of lake. In addition, information about the quality check and usage of observations is provided. Fractions of land and lake in the model grid as well as the weights, which were used to interpolate gridpoint values to the observation location, are given. ~~We use this information as basic material~~ This information is the basis of the present study (see sections 3.3 and 4).

3 Model-observation intercomparison 2012-2018

In this intercomparison we validated HIRLAM/FLake results against observations about the lake surface state. The impact of FLake parametrizations to the weather forecast by HIRLAM ~~is~~ was not considered. This is because ~~the archived observations and the operational HIRLAM results were used during the period from spring 2012 to summer 2018 when FLake was always an integral part of HIRLAM. This means that there are no~~ no non-FLake ~~weather forecasts to compare with~~weather forecasts exist for comparison with the operational forecasts during the validation period.

Throughout the following text, the analysed LSWT refers to the result of OI analysis, where FLake forecast has been used as background (Section 2.2) while the forecast LSWT refers to the value diagnosed from the mixed layer water temperature predicted by FLake (Section 2.1). Observed LSWT refers to the measured by SYKE lake water temperature (Section 3.2).

3.1 FMI operational HIRLAM

FMI operational HIRLAM is based on the last reference version (v.7.4), implemented in spring 2012. (~~Eerola (2013)~~ Eerola, 2013 and references therein). FLake was introduced into this version. After that the development of HIRLAM was frozen. Thus, during the years of the present comparison, the FMI operational HIRLAM system remains unmodified, which offers a clean time

Table 1. FMI operational HIRLAM

Domain	From Atlantic to Ural, from North Africa beyond North Pole
Model horizontal / vertical resolution	7 km / 65 levels
HIRLAM version	7.4
Model dynamics	Hydrostatic, semi-Lagrangian, grid-point
Atmospheric physical parametrizations	Savijärvi radiation, CBR turbulence,
	Rasch-Kristiansson cloud microphysics + Kain-Fritsch convection
Surface physical parametrizations	ISBA-newsnow for surface, FLake for lakes
Data assimilation	Default atmospheric (4DVAR) and surface (OI) analysis
Lateral boundaries	ECMWF forecast
Forecast	Up to +54 h initiated every 6h (00, 06, 12, 18 UTC)

series of data for the model-observation intercomparison. The general properties of the system are summarised in Table 1. In the present study, a coding error in FLake implementation was revealed in the reference HIRLAM v.7.4. A too large critical value to diagnose snow existence prevented practically all accumulation of the forecast snowfall on lake ice in the FMI HIRLAM-FLake operational system.

3.2 SYKE lake observations

In this study we used three different types of SYKE lake observations: LSWT, lake ice dates (LID) and ice thickness and snow depth on lake ice. In total, observations on 45 lakes listed in Appendix (Table A2) were included as detailed in the following. The lake depths and surface areas given in Table A2 are based on he updated lake list of GLDB v.3 (Margarita Choulga, personal communication).

3.2.1 Lake temperature measurements

Regular *in-situ* lake water temperature ~~(LWT)~~ measurements are performed by SYKE. ~~SYKE operates 32~~ Currently SYKE operates 34 regular lake and river water temperature measurement sites in Finland. The temperature of the lake water is measured every morning at 8.00 AM local time, close to shore, at 20 cm below the water surface. The measurements are recorded either automatically or manually and are performed only during the ice-free season (Korhonen, 2002; Rontu et al., 2012). Further, we will for simplicity denote also these data as LSWT observations although they do not represent exactly the same surface water temperature (skin temperature, radiative temperature) that could be estimated by satellite measurements. These data are available in the SYKE open data archive (SYKE, 2018). Measurements from 27 of these 34 lakes (Figure 2, white dots) ~~used by~~ were selected for use in the FMI operational HIRLAM ~~, were included~~ in 2011, and the list has been kept unmodified since that. The set of 27 daily observations, quality-controlled by HIRLAM, were obtained from the analysis

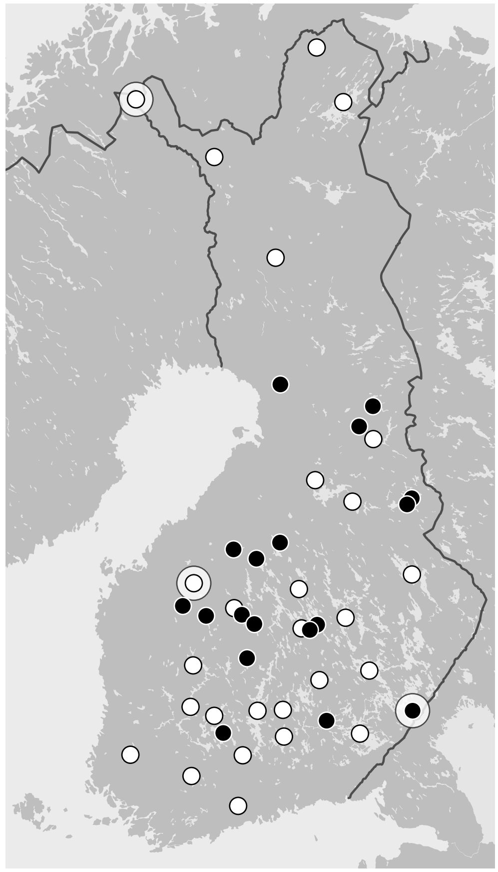


Figure 2. Map of SYKE observation points used in this study: lakes with both lake surface water temperature (LSWT) and lake ice date (LID) observations (white), lakes where only LID is available (black). On Lakes Lappajärvi, Kilpisjärvi and Simpelejärvi also ice thickness and snow ~~thickness~~ depth measurements were used (Section 4.3), they are surrounded with a large white circle. List of the lakes with coordinates is given in Appendix A2.

feedback files and used in all comparisons reported in this study. ~~These data are also available in the SYKE open data archive (SYKE, 2018).~~

3.2.2 Freezing and melting dates

Regular visual observations of freezing and melting of lakes have been recorded in Finland for centuries, the longest time series starting in the middle of the 19th century (Korhonen, 2005). Presently, dates of freezing and melting are available from SYKE (2018) on 123 lakes, but the time series for many lakes are discontinuous. Further, we will denote the melting and freezing dates together by “lake ice dates” (LID). For both freezing and melting the dates are available in two categories: for freezing “freezing of the visible area” (code 29 by SYKE) and “permanent freezing of the visible area” (code 30). For melting the dates are defined as “no ice visible from the observation site” (code 28) and “no ice on the outer open water areas” (code 27). LID observations aim at representing conditions on entire lakes. LID observations by SYKE are made independently of their LSWT measurements and possibly from different locations on the same lakes. The LSWT measurements may be started later than the date of reported lake ice melting or end earlier than the reported freezing date.

LID from the 27 lakes whose LWT LSWT measurements are used in HIRLAM were available and selected for this study. In addition, 18 lakes with only LID available (Figure 2, black dots) were chosen for comparison with HIRLAM/FLake LID.

3.2.3 Ice thickness and snow thickness depth on lakes

SYKE records the lake ice and snow thickness thickness and snow depth on around 50 locations in Finland, archived . Archived data are available in total from 160 measurement sites. The manual measurements are done three times a month during the ice season. Thickness of ice and the snow snow depth on ice are measured by drilling holes through snow and ice layers along chosen tracks, normally at least 50 m from the coast (Korhonen, 2005). The locations may differ from those of the LSWT measurement or LID observation over the same lakes. In this study, measurements from lakes Lappajärvi, Kilpisjärvi and Simplejärvi were utilised as additional data for validation in Section 4.3. These lakes, sufficiently large in order to fit well the HIRLAM grid, represent the western, northern and south-eastern Finland.

3.3 Lake Validation of HIRLAM/FLake lake surface statederived from HIRLAM output

3.3.1 Lake surface water temperature

Diagnosed LSWT from LSWT by HIRLAM/FLakeanalysis and forecastcycles , resulting from the objective analysis or diagnosed from the forecast, was compared with the observed LWT LSWT by SYKE using data extracted from the analysis feedback files (Section 2.2) at the observation locations on 06 UTC every day. , excluding the winter periods 1 December - 31 March. The observations (ob) at 27 SYKE stations were assumed to represent the true value, while the analysis (an) is the result of OI that combines the background forecast (fc) with the observations. Time-series, maps and statistical scores, to be presented in Section 4.1, were derived from these.

3.3.2 Freezing and melting dates

3.3.2 Lake ice conditions

Both the analysed LSWT and the lake ice thickness forecast by FLake were separately used to define LID. The values For this study, the observed LID, ice and snow thickness observations were obtained from SYKE open data base, relying on their quality control. The HIRLAM/FLake analysed LSWT as well as the predicted ice thickness and snow depth were picked afterwards from the HIRLAM archive for a single gridpoint nearest to each of the 45 observation locations (not interpolated as in the analysis feedback file that was used for the LSWT comparison). For the definition of LID, it It was assumed that the gridpoint value nearest to the location of the LSWT observation represents the ice conditions over the chosen lake.

LID were defined in two independent ways: from the analysed LSWT and from the forecast lake ice thickness. Note that the ice thickness and snow depth on ice are not analysed variables in HIRLAM. In autumn a lake can freeze and melt several times before final freezing. The last date when the forecast ice thickness crossed a critical value of 1 mm or the analysed LSWT fell below freezing point was selected as the date of freezing. To decrease the effect of oscillation of the gridpoint values between the HIRLAM forecast-analysis cycles, the mean of the four daily ice thickness forecasts or analysed LSWT values was used. In the same way, the last date when the forecast ice thickness fell below the critical value of 1 mm or the analysed LSWT value crossed the freezing point was selected as melting day.

3.4 Validation methods

For LSWT statistics we used data collected during the HIRLAM surface analysis at each active observation location (Section 2.2), excluding the winter periods 1 December - 31 March. The observations (ob) at 27 SYKE stations were assumed to represent the true value, while the analysis (an) is the result of OI that combines the background forecast (fc) with the observations. Time-series, maps and statistical scores, to be presented in Section 4.1, were derived from theseTo decrease the effect of oscillation of the gridpoint values between the HIRLAM forecast-analysis cycles, the mean of the four daily ice thickness forecasts or analysed LSWT values was used.

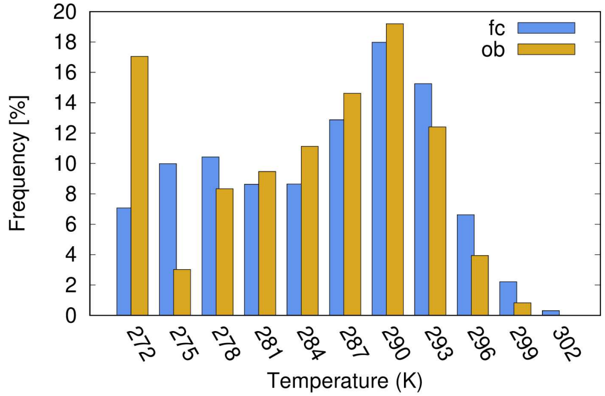
LID by HIRLAM/FLake were compared to the observed dates during 2012-2018, including in the comparison data over all months. In this comparison we included data also during the winter period. The category 29 observations (“freezing of the visible area”, see Section 3.2.2) were used. In this category the time series were the most complete at the selected stations. For the same reason, the melting observations of category 28 (“no ice visible from the observation site”) were used for comparison. Furthermore, using a single gridpoint value for the calculation of LID also seems to correspond best the observation definition based on what is visible from the observation site. The statistics were calculated as obfc - fc and ob ob and an - anob. Hence, positive values mean that melting or freezing takes place too late in the model as compared to the observations.

In this study, lake ice thickness and snow depth measurements from lakes Lappajärvi, Kilpisjärvi and Simpelejärvi were utilised as additional data for validation of predicted by HIRLAM/FLake ice thickness and snow depth (Section 4.3). These lakes, representing the western, northern and south-eastern Finland, were selected for illustration based on the best data availability during the study years. They are also sufficiently large in order to fit well the HIRLAM grid.

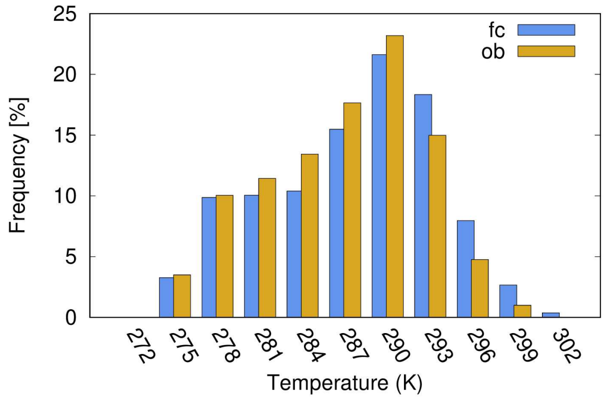
4 Results

4.1 Analysed and forecast LSWT at observation points

Figure 3 shows the frequency distribution of LSWT according to FLake forecast and SYKE observations. It is evident that the amount of data in the class of temperatures which represents frozen conditions (LSWT flag value 272 K) is was underestimated by the forecast (Figure 3a). When subzero temperatures are were excluded from the comparison (Figure 3b), underestimation in the colder temperature classes and overestimation in the warmer classes still remains.



(a) with all temperatures (also frozen conditions) included



(b) only open water temperatures included

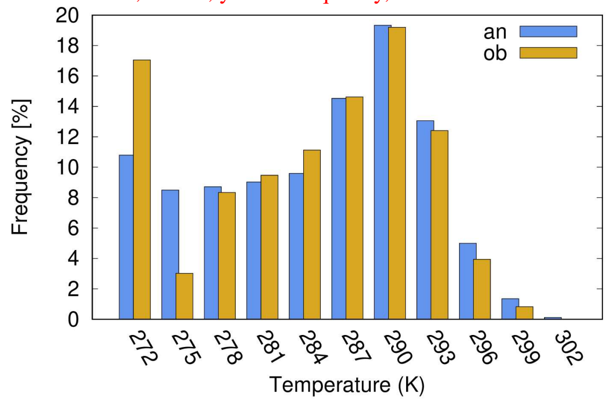
Figure 3. Frequency of observed (ob, yellow) and forecast (fc, blue) LSWT over all 27 SYKE lakes 2012-2018. x-axis: LSWT, unit K, y-axis: frequency, unit %.

LSWT analysis (Figure 4) ~~improves~~ improved the situation somewhat but the basic features remain. This is due to the dominance of FLake forecast via the background of the analysis. In Section 4.3, we will show time-series illustrating the physics behind these LSWT statistics.

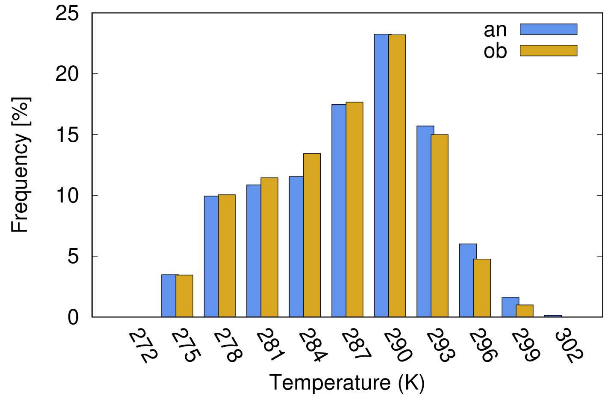
Table 2 confirms the warm bias by FLake in the unfrozen conditions. Similar results were obtained for all stations together and also for our example lakes Lappajärvi and Kilpisjärvi, to be discussed in detail in Section 4.3. There were three lakes with negative LSWT bias according to FLake forecast, namely the large lakes Saimaa and Päijänne ~~,~~ and the smaller Ala-Rieveli. After the correction by objective analysis, a small positive bias converted to negative over 6 additional lakes, among them the large lakes Lappajärvi in the west and Inari in the north. The mean absolute error decreased from forecast to analysis ~~in~~ on every lake.

~~with all temperatures (also frozen conditions) included~~

~~only open water temperatures included Frequency of observed (yellow) and forecast (blue) LSWT over all 27 SYKE lakes 2012-2018. x-axis: LSWT, unit K, y-axis: frequency, unit %.~~



(a) with all temperatures (also frozen conditions) included



(b) only open water temperatures included

Figure 4. As for Figure 3 but for observed and analysed (an) LSWT.

In the frequency distributions, the warm temperatures ~~are~~ were evidently related to summer. For FLake, the overestimation of maximum temperatures, especially in shallow lakes, is a ~~knowm~~ known feature (e.g. Kourzeneva 2014). It is related to the difficulty of forecasting the mixed layer thermodynamics under strong solar heating. Cold and subzero temperatures ~~occur~~ occurred in spring and autumn. In a few large lakes like Saimaa, Haukivesi, Pielinen, LSWT ~~tends~~ tended to be slightly underestimated in autumn both according to the FLake and the analysis (not shown). ~~However, as will be shown in Sections 4.2 and 4.3, the~~ The cold left-hand side columns in the frequency distributions (Figures 3a and 4a) are mainly related to spring, when HIRLAM/FLake ~~tends~~ tended to melt the lakes significantly too early (Sections 4.2 and 4.3).

Table 2. Statistical scores for LSWT at all stations and at two selected stations

station	fc or an	mean ob	bias	mae	stde	N
unit		K	K	K	K	
ALL	fc	286.3	0.91	1.94	2.34	30877
	an	286.3	0.35	1.32	1.72	30861
Lappajärvi	fc	286.9	0.33	1.23	1.62	1243
	an	286.9	-0.65	1.06	1.10	1243
Kilpisjärvi	fc	281.7	1.82	2.13	2.15	780
	an	281.7	1.10	1.42	1.51	780

Statistics over days when both forecast/analysis and observation indicate unfrozen conditions. bias = systematic difference fc/an - ob, mae = mean absolute error, stde = standard deviation of the error, N = number of days (06 UTC comparison, no ice).

There are problems, especially in the analysed LSWT, over (small) lakes of irregular form that fit poorly the HIRLAM grid and where the measurements may represent more the local than the mean or typical conditions over the lake. These are the only ones where an underestimation of summer LSWT can be was seen. Cases occur occurred where FLake results differ so much from the observations that the quality control of the HIRLAM surface data assimilation rejects HIRLAM quality control against background values rejected the observations, forcing also the analysis to follow the incorrect forecast (not shown).

4.2 Freezing and melting dates

In this section the freezing and melting dates from HIRLAM are verified against corresponding observed dates over 45 lakes (Appendix Table A2). In the following, 'LSWT an' refers to the LID estimated from analysed LSWT and 'IceD fc' to those estimated from the forecast ice thickness by FLake. The time period contains six freezing periods (from autumn 2012 to autumn 2017) and seven melting periods (from spring 2012 to spring 2018). Due to some missing data the number of freezing cases was 233 and melting cases 258. The 'IceD fc' data for the first melting period in spring 2012 was missing. The overall statistics of the error in freezing and melting dates are shown in Table 3. In most cases the difference in error between the dates based on forecast and analysis was small. This is natural as the first guess of the LSWT analysis is the forecast LSWT by FLake. We will discuss next the freezing, then the melting dates.

Statistics of the error in melting and freezing dates are shown in Table 3. 'LSWT an' refers to the melting/freezing dates computed from analysed lake surface temperature and The bias in the error of freezing dates was small according to both 'IceD fc' to those estimated from the forecast ice thickness. Over the 45 lakes included in this comparison, the number of cases of melting was 288 as estimated from the analyzed LSWT and 258 as estimated from the forecast ice thickness. The difference is due to starting time of our data. When the data started at the 1st of April 2012, at several stations the lake was already open according to FLake forecast while the analysed LSWT still indicated frozen conditions. For freezing, the number of cases

Table 3. Statistical measures of the error of freezing and melting date

		bias	~~sde~~ stde	max	min	N
unit		days	days	days	days	
Freezing	LSWT an	-3.5	17.9	64	-52	233
	IceD fc	-0.3	17.8	67	-41	233
Melting	LSWT an	-15.2	8.5	2	-54	288
	IceD fc	-20.5	9.2	-1	-56	258

Denotation: LSWT an - LID estimated from analysed LSWT, IceD fc - LID estimated from forecast ice thickness.

~~was 233 according to both estimates. As the data contains the time period from the 1st April 2012 to the 30th June 2018, the maximum number of freezing events on an individual lake is six and that of melting events seven. In practice, the number may be less for some lakes because of missing observations~~and 'LSWT an', -0.3 and -3.5 days, respectively. The minimum and maximum errors were large in both cases: the maximum freezing day occurred about two months too late the minimum about one and a half months too early. However, as will be shown later, the largest errors mostly occurred on a few problematic lakes while in most cases the errors were reasonable.

Figure 5a) shows the frequency distribution of the error of freezing dates. ~~Definition of the freezing date from the ice thickness by FLake gave slightly more occurrences~~ Forecast freezing dates occurred slightly more often in the unbiased class (error between -5 - +5 days), compared to the ~~estimate from the analysed LSWT~~estimated dates from the analysis. Of all cases 48 %~~and~~ /40 % ~~fell in this class according to ice thickness and LSWT, respectively. In 16~~(percentages here and in the following are given as 'IceD fc'/'LSWT an') fell into this class. In 20% / ~~20~~26% of cases the freezing ~~occurs~~ occurred more than five days too late and only in ~~9~~11% / ~~11~~9% cases more than two weeks too late. ~~This class of more than two weeks too late freezing consists of~~ In case of 'IceD fc', the class of freezing more than 15 days too late comprised 25 cases ~~which are~~ distributed over 15 lakes, thus ~~in most cases one event~~ mostly one or two events per lake. This suggests that the error ~~is~~ was related more to individual years than to systematically problematic lakes. It is worth noting, that of the eight cases where the error ~~is~~ was over 45 days, ~~are all but one due to one lake,~~ six cases were due to a single lake, Lake Kevojärvi~~which~~ . This lake is situated in the very north of Finland. ~~This lake is~~ It is very small and narrow, with an area of 1 km~~$^{-2}$, and situated~~ ^{2}, and located in a steep canyon. Therefore it is poorly represented by the HIRLAM grid and ~~both FLake and analysis~~ the results seem unreliable.

Concerning ~~the cases of~~ too early freezing, in ~~44~~33% / ~~32~~44% of the cases freezing ~~occurs~~ occurred more than five days too early and in ~~19~~15% / ~~15~~19% more than two weeks too early. ~~The last mentioned~~ According to the forecast, these 15% (34 cases) ~~are~~ were distributed over 19 lakes. Each of the five large lakes Pielinen, Kallavesi, Haukivesi, Päijänne and Inari ~~occur~~ occurred in this category three times while all other lakes together ~~share~~ shared the remaining 19 cases during the six winters.

~~Looking at the errors in~~ The melting dates (~~Figure 5b) , both estimates indicate too earlymelting and the distribution~~ Table 3) show a large negative bias, about two ('LSWT an') or three weeks ('IceD fc'), indicating that lake ice melting was systematically forecast to occur too early. However, the standard deviation of the error was only about half of that of the error of freezing dates and there were no long tails in the distribution (Figure 5b). The distribution is strongly skewed towards too early ~~dates. Based on the LSWT analysis, the~~ melting, but much narrower than that of freezing (Figure 5a). The large bias was most probably due to the bug of this HIRLAM version that prevented the accumulation of snow over lake ice (see also Section 4.3). The maximum frequency (~~52~~47 %) ~~occurs~~ was in the class ~~-14~~-24 - ~~-5 days while based on the ice thickness~~-15 days for 'IceD fc', while in case of 'LSWT an', the maximum frequency (~~47~~52 %) ~~is~~ occurred in the class ~~-24~~-14 - ~~-15 days. The mean values are -15.2 and -20.5 days and the standard deviations are 8.5 and 9.2 days, respectively. FLake suggests~~ -5 days. FLake forecast 'IceD fc' suggested only three cases in the unbiased class -4 - +5 while according to ~~the LSWT analysis there are~~ 'LSWT an' there were 12 cases in this class. Hence, the melting dates derived from analysed ~~LWST correspond~~ LSWT corresponded the observations better than those derived from FLake ice thickness forecast~~but both are strongly biased towards too early melting. In the class where the error is over 35 days too early there are 19 cases on 12 different lakes. Four cases of these occured in the largest error category (over 45 days)on Lake Kilpisjärvi.~~

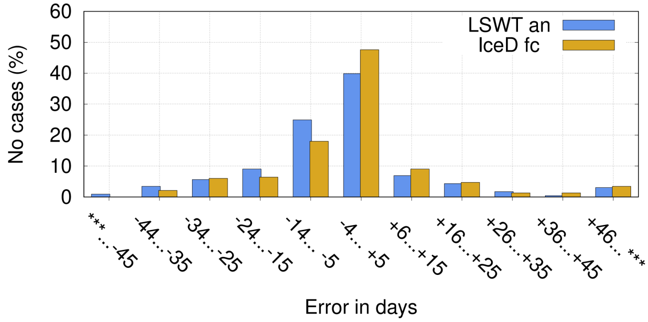
~~If we compare the error in freezing or melting dates based on analysed LSWT on those (27) lakes where SYKE temperature observations are available and used in the analysis to the rest (18) of lakes with no observations, it appears that the differences are small (not shown). Furthermore, similar differences appear also on the error estimates based on ice thickness from FLake. This suggests that the differences between these groups are related to the individual properties of the lakes: their depth, size, shape etc. rather than to the usage of LSWT observations in the analysis~~.

~~We can conclude~~ Note that this kind of method of verifying LID compares two different types of data. The observations by SYKE are visual observations from the shore of the lake (see Section 3.2.2), while the freezing and melting dates from HIRLAM are based on single-gridpoint values of LSWT or ice thickness (see Section 3.3.2). In addition, the resulting freezing and melting dates from HIRLAM are somewhat sensitive to definition of the freezing and melting tresholds. Here we used 1 mm for the forecast ice thickness and the freezing point for the LSWT analysis as the critical values.

In conclusion, the validation statistics show that HIRLAM/FLake ~~succeeds~~ succeeded rather well in predicting ~~the~~ freezing of Finnish lakes. Almost in half of the cases the error ~~is~~ was less than ± 5 days. Some bias towards too early freezing can be seen ~~. Melting is~~ both in forecast and in the analysis. Melting was more difficult. FLake ~~predicts~~ predicted ice melting always too early, with a mean error of over two weeks, and the ~~LSWT analysis mostly follows it. The statistics suggest that only on a few stations the freezing or melting dates were systematically wrong during most of the years. Instead, most of the large errors were distributed among many lakes. The result of the freezing or melting dates diagnostics is somewhat sensitive to how the tresholds for freezing and melting are set. Here we used 1 mm for icethickness and the freezing point for the LSWT analysis as the critical values~~analysis mostly followed it. These results are rather obvious because of the missing snow on ice.

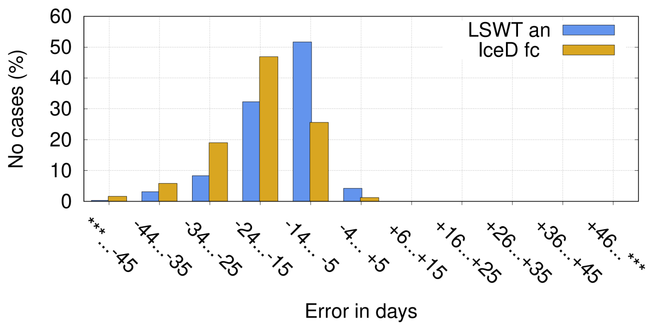
4.3 Comparisons on three lakes

In this section we ~~combine the analysis of LWST time-series and LID~~ present LSWT and LID time-series for two representative lakes, Kilpisjärvi in the north and Lappajärvi in the west (see the map in Figure 2). Observed and forecast ice and snow thickness are discussed, using also additional data from Lake Simpelejärvi in ~~the south-east of~~ southeastern Finland.



~~Frequency distribution of the difference between analysed/forecast and observed freezing days over all lakes 2012-2018. Variables used in diagnosis of ice existence: analysed LSWT crossing the freezing point (blue) and forecast ice thickness > 1 mm (orange). Observed variable: freezing date by SYKE. x-axis: difference (fc-ob), unit day, y-axis: percentage of all cases.~~

(a) error in freezing days



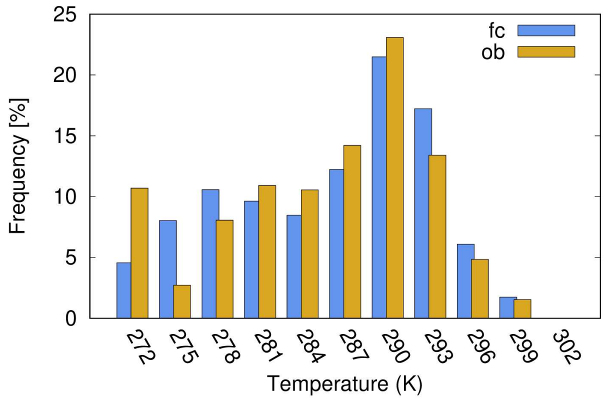
~~As for Figure 5a but for melting days.~~

(b) error in melting days

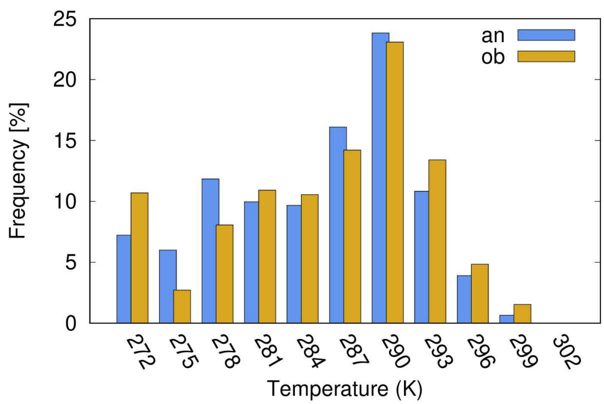
Figure 5. Frequency distribution of the difference between analysed/forecast and observed freezing and melting days over all lakes 2012-2018. Variables used in diagnosis of ice existence: analysed LSWT crossing the freezing point (blue) and forecast ice thickness > 1 mm (magenta). Observed variable: freezing date by SYKE. x-axis: difference (fc-ob), unit day, y-axis: percentage of all cases.

Lake Kilpisjärvi is an Arctic lake at the elevation of 473 m, surrounded by fells. ~~Its~~ The lake occupies 40 % of the area of HIRLAM gridsquare covering it (the mean elevation of the gridsquare is 614 m). The average/maximum ~~depth is 22.5~~depths of the lake are 19.5/57 m and the surface area ~~37.33~~is 37.3 km^2. The heat balance as well as the ice and snow conditions on Lake Kilpisjärvi have been ~~a subject of~~ subject to several studies (Leppäranta et al., 2012; Lei et al., 2012; Yang et al., 2013). Typically, the ice season lasts there seven months from November to May. Lake Lappajärvi is formed from a 23 km wide meteorite impact crater, which is estimated to be 76 million years old. It is Europe's largest crater lake with a surface area of 145.5 km^2 and an average/maximum depth of ~~12~~6.9/36 m. Here the climatological ice season is shorter, typically about five months from December to April. The average/maximum depth of Lake Simpelejärvi is ~~9.3~~8.7/34.4 m and the surface area 88.2 km^2. This lake is located at the border between Finland and Russia and belongs to the catchment area of Europe's largest lake, Lake Ladoga in Russia.

~~Figures 6a - 7b~~



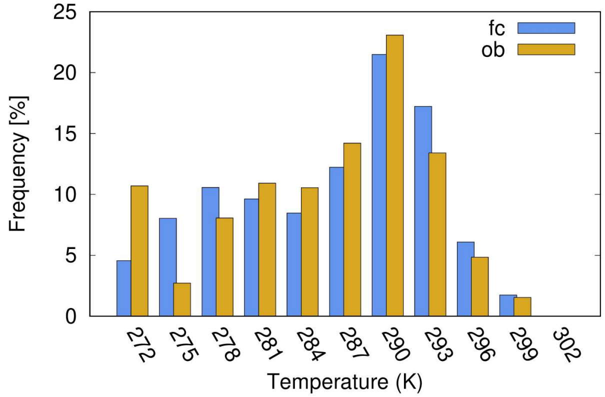
(a) forecast v.s. observation



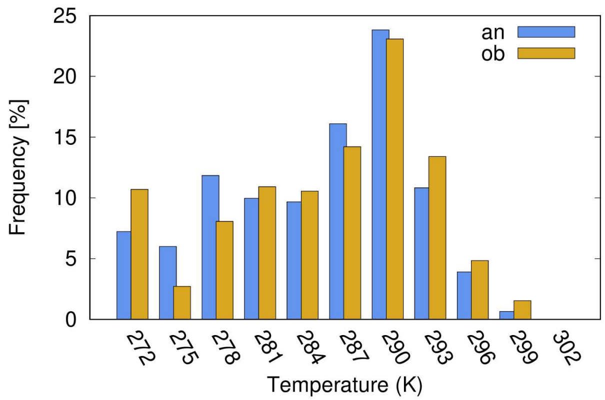
(b) analysis v.s. observation

Figure 6. Frequency of observed (yellow) and forecast or analysed (blue) LSWT over Lake Lappajärvi 2012-2018, all temperatures included. x-axis: LSWT, unit K, y-axis: frequency, unit %.

Figures 6 and 7 show the frequency distributions of LSWT according to ~~the observations~~ forecast v.s. ~~forecast and analysis for these lakes. Features common to the majority of~~ observation and analysis v.s. observation for Lappajärvi and Kilpisjärvi. Features similar to the results averaged over all lakes (Section 4.1, Figures 3 and 4) are seen, i.e. underestimation of the amount of cold temperature cases and overestimation of the warmer temperatures by the forecast and analysis. On Lake Lappajärvi, only the amount of below-freezing temperatures ~~is~~ was clearly underestimated, otherwise the distributions look quite balanced. According to the observations, on Lake Kilpisjärvi ~~the days with frozen surface dominate during the April–November periods~~ ice-covered days dominated during the periods from April to November. According to both FLake forecast and HIRLAM LSWT analysis the amount of these days ~~is~~ was clearly smaller.



(a) ~~As for Figure 3a) but for Lake Lappajärvi, with all temperatures included~~forecast v.s. observation
~~As for Figure 4a) but for Lake Lappajärvi~~
~~As for Figure 6a but for Lake Kilpisjärvi.~~



(b) ~~As for Figure 6b but for Lake Kilpisjärvi~~analysis v.s. observation

Figure 7. As for Figure 6 but for Lake Kilpisjärvi.

Yearly time series of the observed, forecast and analysed LSWT, with the observed LID marked, are shown in Figures 8 and 9. In the absence of observations, the HIRLAM analysis ~~follows~~ followed the forecast. Missing data in the time series close to freezing and melting are due to missing observations, hence missing information in the feedback files (see Section 2.2). Differences between the years due to the different prevailing weather conditions can be seen in the temperature variations.

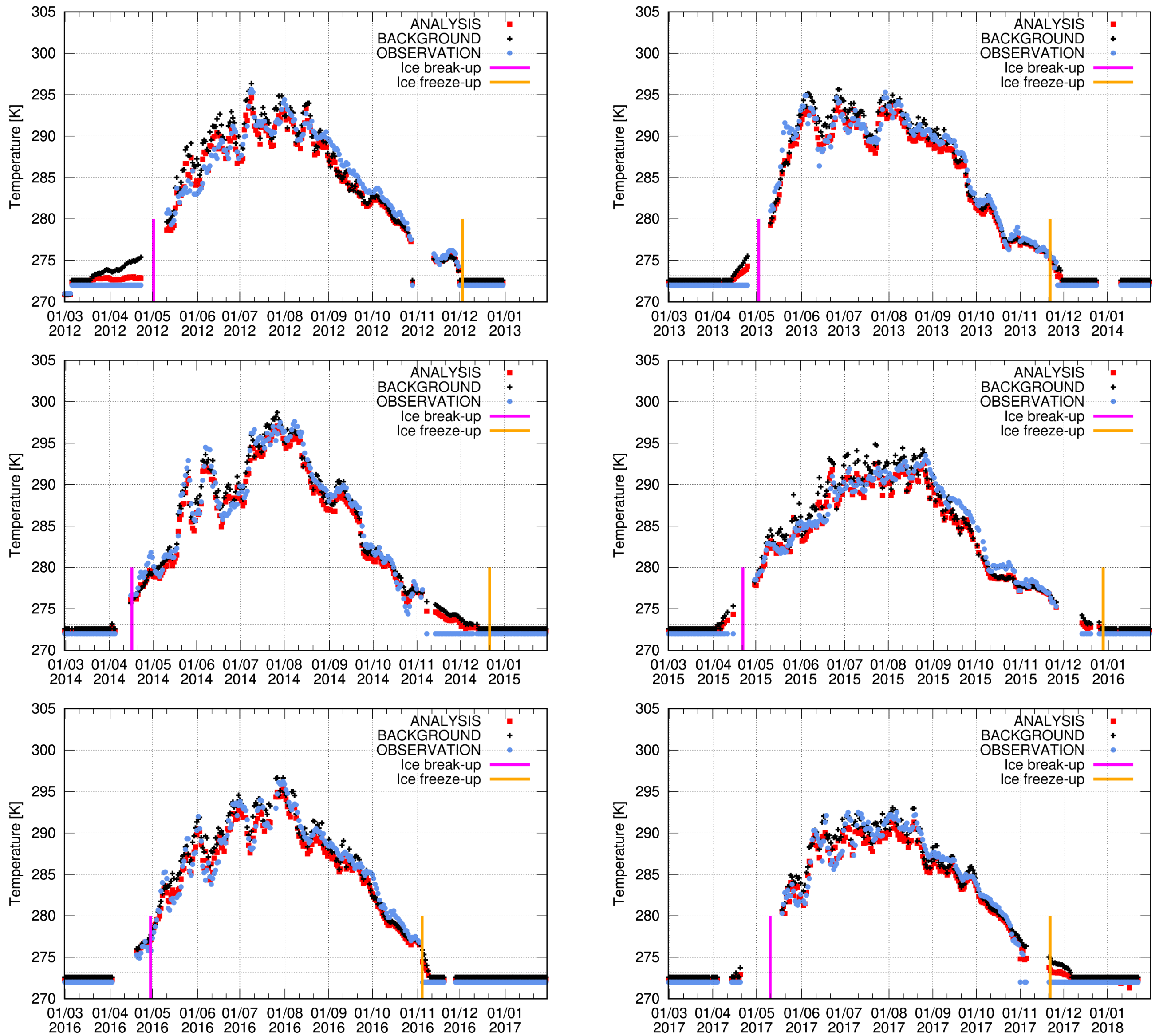


Figure 8. Time-series of the observed, analysed and forecast LSWT at the Lappajärvi observation location 23.67 E, 63.15 N for the years 2012-2018 based on 06 UTC data. Markers are shown in the inserted legend. Observed freezing date (blue) and melting date (red) are marked with vertical lines.

Generally, ~~in spring FLake tends~~ FLake tended to melt the lakes too early in spring, as already indicated by the LID statistics (Section 4.2). The too early melting and too warm LSWT in summer show up clearly in Kilpisjärvi (Figure 9). In Lappajärvi, the model and analysis ~~are~~ were able to follow even quite large and quick variations of LSWT in summer, but ~~tend~~ tended to somewhat overestimate the maximum temperatures. Overestimation of the maximum temperatures by FLake ~~is~~ was still more

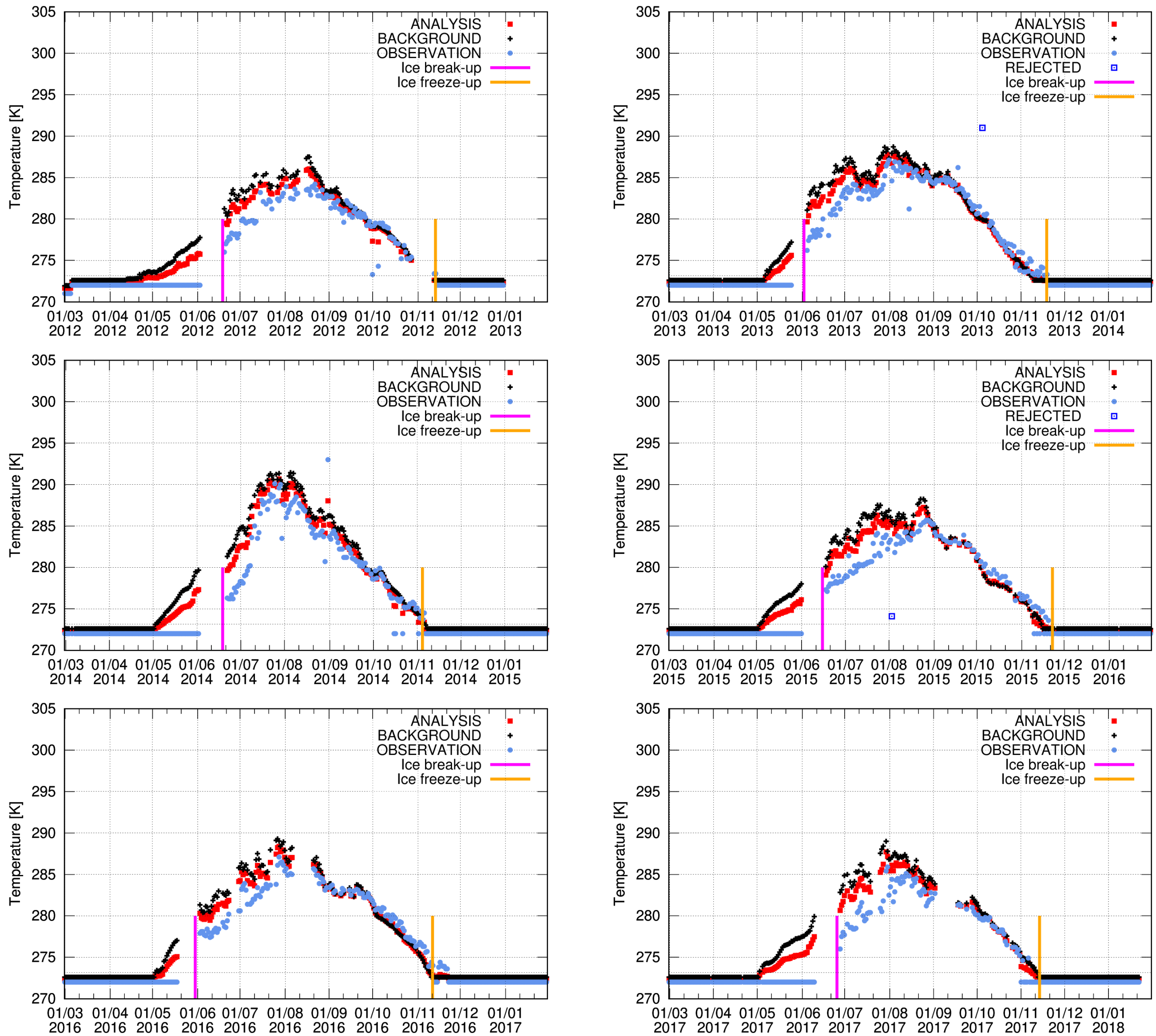


Figure 9. As for Figure 8 but for lake Kilpisjärvi, 20.82 E, 69.01 N.

prominent in shallow lakes (not shown). In autumn over Lakes Lappajärvi and Kilpisjärvi, the forecasts and analyses ~~follow~~ followed closely the LSWT observations and ~~reproduce the freezing date~~ reproduced the freezing dates within a few days, which ~~is~~ was also typical to the majority of lakes.

Figure 10 shows a comparison of forecast and observed evolution of ice ~~and snow thickness~~ thickness and snow depth on Lappajärvi, Kilpisjärvi and Simpelejärvi in winter 2012-2013, typical also for the other lakes and years studied. In all three lakes, the ice thickness ~~starts~~ started to grow after freezing both according to the forecast and the observations. In the beginning

HIRLAM/FLake ice grows grew faster than observed. However, according to the forecast ice thickness starts started to decrease in March of every year but according to the observations only a month or two later. The most remarkable feature is that there is was no snow in the FLake forecast. It was found that this was due to a coding error in the HIRLAM reference version 7.4 which is applied operationally in FMI.

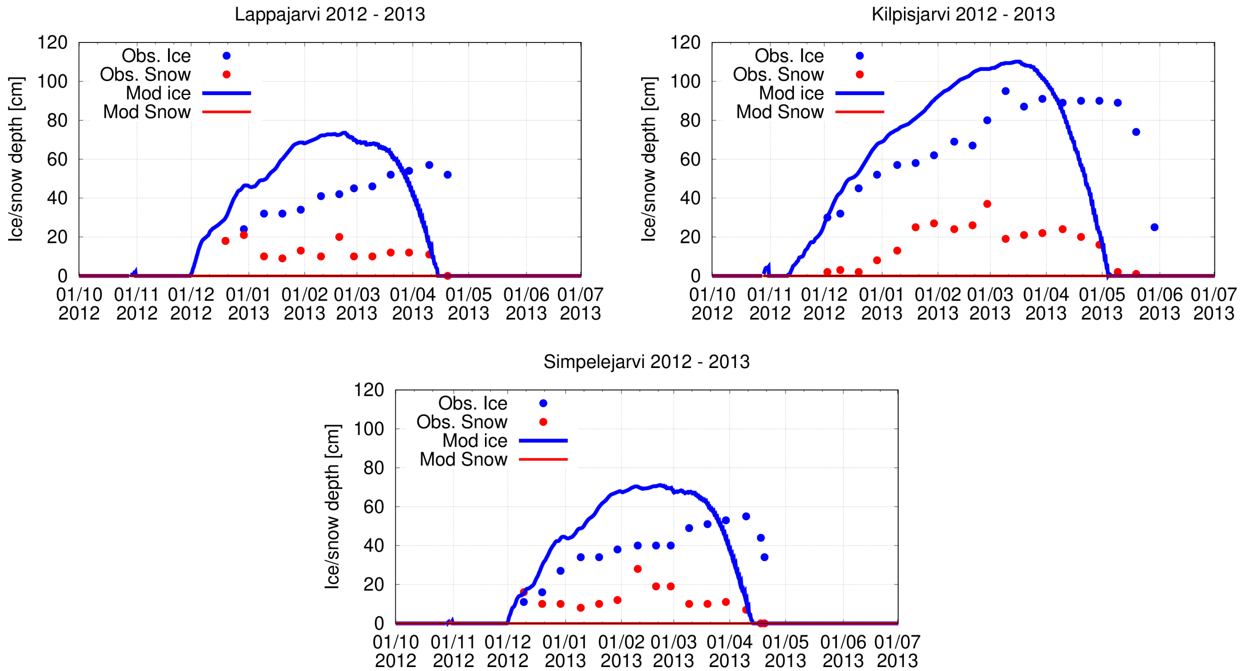


Figure 10. Evolution of ice (blue) and snow (red) thickness at Lakes Lappajärvi, Kilpisjärvi and Simpelejärvi during winter 2012-2013.

The too early melting of ice in the absence of snow could be explained by the wrong absorption of the solar energy in the model. In reality, the main factor of snow and ice melt in spring is the increase of daily solar radiation. In HIRLAM, the downwelling short-wave irradiance at the surface is known to be reasonable, with some overestimation of the largest clear-sky fluxes and all cloudy fluxes (Rontu et al., 2017). Over lakes, HIRLAM/FLake uses constant values for the snow and ice shortwave reflection, with albedo values of 0.75 and 0.5, correspondingly. When there was no snow, the lake surface was thus assumed too dark. 25 % more absorption of an assumed maximum solar irradiance of 500 Wm^{-2} (valid for the latitude of Lappajärvi in the end of March) would mean availability of extra 125 Wm−2 $^{-2}$ for melting of the ice, which corresponds the magnitude of increase of available maximum solar energy within a month at the same latitude.

The forecast of too thick ice can also be explained by the absence of snow in the model. When there is no insulation by the snow layer, the longwave cooling of the ice surface in clear-sky conditions is more intensive and leads to faster growth of ice compared to the situation of snow-covered ice. In nature, ice growth can also be due to the snow transformation, a process whose parametrization in the models is demanding (Yang et al., 2013; Cheng et al., 2014).

Also the downwelling longwave radiation plays a role in the surface energy balance. We may expect values from 150 Wm^{-2} to 400 Wm^{-2} in the Nordic spring conditions, with the largest values related to cloudy and the smallest to clear-sky situations. The standard deviation of the predicted by HIRLAM downwelling longwave radiation fluxes has been shown to be of the order of 20 Wm^{-2}, with a positive systematic error of a few Wm^{-2} (Rontu et al., 2017). Compared to the systematic effects related to absorption of the solar radiation, the impact of the longwave radiation variations on lake ice evolution is presumably small.

5 Conclusions and outlook

In this study, *in-situ* lake observations from the Finnish Environment Institute were used for validation of the HIRLAM NWP model, which is applied operationally in the Finnish Meteorological Institute. ~~It~~ HIRLAM contains Freshwater Lake prognostic parametrizations and an independent objective analysis of lake surface state. We focused on comparison of observed and forecast lake surface water temperature, ice ~~and snow thickness~~ thickness and snow depth in the years 2012 - 2018. Because the HIRLAM/FLake system was unmodified during this period, a long uniform dataset was available for evaluation of the performance of FLake integrated ~~in~~ into an operational NWP model. On the other hand, no conclusions about the impact of the lake surface state on the operational forecast of the near-surface temperatures, cloudiness or precipitation can be drawn because of the lack of alternative (without FLake) forecasts for comparison.

On average, the forecast and analysed LSWT were warmer than observed with systematic errors of 0.91 K and 0.35 K, correspondingly. The mean absolute errors were 1.94 and 1.32 K. Thus, the independent observation-based analysis of *in-situ* LSWT observations was able to improve the FLake +6 h forecast used as the first guess. However, the resulting analysis is by definition not used for correction of the FLake forecast but remains an independent by-product of HIRLAM.

An overestimation of the FLake LSWT summer maxima was found, especially for the shallow lakes. This behaviour of FLake is well known, documented earlier e.g. by ~~Kourzeneva (2014)~~Kourzeneva, 2014. It arises due to the difficulty to handle correctly the mixing in the near-surface water layer that is intensively heated by the sun.

Forecast freezing dates were found to correspond the observations well, typically within a week. The forecast ice thickness tended to be overestimated, still the melting dates over most of the lakes occured systematically several weeks too early. Practically no forecast snow was on found on the lake ice, although the snow parametrization by FLake was included in HIRLAM. The reason for the ~~wrong behaviour in HIRLAM~~ incorrect behaviour was evidently related to a coding error in HIRLAM that prevented snow accumulation on lake ice. The too early melting and overestimated ice thickness differ from the results by ~~Pietikäinen et al. (2018); Yang et al. (2013); Kourzeneva (2014)~~Pietikäinen et al., 2018; Yang et al., 2013; Kourzeneva, 2014, who reported somewhat too late melting of the Finnish lakes when FLake with realistic snow parametrizations was applied within a climate model or ~~independently,~~ stand-alone driven by NWP data. It can be concluded that a realistc parametrization of snow on lake ice is importart in order to describe correctly the lake surface state in spring.

Small lakes and those of complicated geometry cause problems for the relatively coarse HIRLAM grid of 7 - kilometre resolution. The problems are related to the observation usage, forecast and validation, especially when interpolation and selection of point values are applied. The observations and model represent different spatial scales. For example, the comparison of the freezing and melting dates was based on diagnostics of single-gridpoint values that were compared to observations representing entire lakes as seen overseen from the observation sites. Also the results of LID diagnostics were sensitive to the criteria for definition of the ice existence in HIRLAM/FLake. All this adds unavoidable inaccuracy into the model-observation intercomparison but does not change the main conclusions of the present study.

SYKE LSWT observations used for the real-time analysis are regular and reliable but did do not always cover the days immediately after melting or close to freezing, partly because the quality control of HIRLAM LSWT analysis utilizes the SYKE statistical lake water temperature model results in a too strict way. Although the 27 observations are located all over the country, they cover a very small part of the lakes and their availability is limited to Finland. SYKE observations of the ice and snow depth as well as the freezing and melting dates provide valuable data for the validation purposes.

A need for minor technical corrections in the FMI HIRLAM/FLake system was revealed. The snow accumulation bug was corrected in October 2018, based on our findings. Further developments and modifications are not foreseen because the HIRLAM NWP systems, applied in the European weather services, are being replaced by kilometre-scale HARMONIE-AROME-based operational systems (Bengtsson et al., 2017)ALADIN-HIRLAM forecasting systems (Termonia et al., 2018; Bengtsson et al., 2017) , where the prognostic FLake parametrizations are also available. HARMONIE/FLake uses the newest version of the global lake database (GLDB v.3) and contains updated snow and ice properties that were suggested by (Yang et al., 2013)Yang et al., 2013 . The objective analysis of lake surface state is yet to be implementedinto HARMONIE-AROME, taking into account the HIRLAM experience summarized in this study and earlier by Kheyrollah Pour et al. (2017). Kheyrollah Pour et al., 2017. In the future, an important source of wider observational information on lake surface state are the satellite measurements, whose operational application in NWP models still requires further work.

Prognostic and diagnostic lake variables within HIRLAM variable unit type temperature of snow on lake ice K prog by FLake temperature of lake ice K prog by FLakemean water temperature K prog by FLakemixed layer temperature K prog by FLakebottom temperature K prog by FLaketemperature of upper layer sediments K prog by FLakemixed layer depth m prog by FLakethickness of upper layer sediments m prog by FLakethermocline shape factor - prog by FLakelake ice thickness m prog by FLakesnow depth on lake ice m prog by FLakeLSWT K diag by FLake= mixed layer temperature if no icelake surface temperature K diag by FLake uppermost temperature: LSWT or ice or snowLSWT K anal by HIRLAMflag value 272 K when there is icelake surface roughness m diag by HIRLAMscreen level temperature over lake m diag by HIRLAMscreen level abs.humidity over lake m diag by HIRLAManemometer level u-component over lake m diag by HIRLAManemometer level v-component over lake m diag by HIRLAMlatent heat flux over lake Wm^{-2} diag by HIRLAMsensible heat flux over lake Wm^{-2} diag by HIRLAMscalar momentum flux over lake Wm^{-2} diag by HIRLAMSW net radiation over lake Wm^{-2} diag by HIRLAMLW net radiation over lake Wm^{-2} diag by HIRLAMdepth of lake m pres in HIRLAM gridfraction of lake 0-1pres in HIRLAM gridfraction of lake ice 0-1diag in HIRLAM grid

Lakes with SYKE observations used in this study NAME LON LAT MEAND HIRD HIRFR HIRID Pielinen 29.607 63.271 11.1 10.0 0.916 4001 Kallavesi 27.783 62.762 12.1 10.0 0.814 4002 Haukivesi 28.389 62.108 9.0 10.0 0.725 4003 Saimaa 28.116 61.338 17.0 10.0 0.950 4004 Pääjärvi1 24.789 62.864 3.9 3.0 0.430 4005 Nilakka 26.527 63.115 4.9 10.0 0.866 4006 Konnevesi 26.605 62.633 15.9 10.0 0.937 4007 Jääsjärvi 26.135 61.631 4.6 10.0 0.750 4008 Päijänne 25.482 61.614 14.1

~~10.0 0.983 4009 Ala-Rieveli 26.172 61.303 11.3 10.0 0.549 4010 Kyyvesi 27.080 61.999 4.4 10.0 0.810 4011 Tuusulanjärvi 25.054 60.441 3.2 3.0 0.174 4012 Pyhäjärvi 22.291 61.001 5.5 5.0 0.922 4013 Längelmävesi 24.370 61.535 6.8 10.0 0.875 4014 Pääjärvi2 25.132 61.064 14.8 14.0 0.350 4015 Vaskivesi 23.764 62.142 7.0 10.0 0.349 4016 Kuivajärvi 23.860 60.786 2.2 10.0 0.419 4017 Näsijärvi 23.750 61.632 14.1 10.0 0.850 4018 Lappajärvi 23.671 63.148 12.0 10.0 1.000 4019 Pesiöjärvi 28.650 64.945 7.3 7.0 0.290 4020 Rehja-Nuasjärvi 28.016 64.184 8.5 10.0 0.534 4021 Oulujärvi 26.965 64.451 7.6 10.0 1.000 4022 Ounasjärvi 23.602 68.377 6.6 10.0 0.166 4023 Unari 25.711 67.172 6.1 10.0 0.491 4024 Kilpisjärvi 20.816 69.007 22.5 22.0 0.399 4025 Kevojärvi 27.011 69.754 7.0 10.0 0.016 4026 Inarijärvi 27.924 69.082 14.4 14.0 0.979 4027 Simpelejärvi 29.482 61.601 9.3 10.0 0.548 40241 Pokkaanlahti 27.264 61.501 7.00 10.0 0.299 40261 Muurasjärvi 25.353 63.478 9.10 10.0 0.060 40263 Kalmarinjärvi 25.001 62.786 5.80 5.0 0.330 40271 Summasjärvi 25.344 62.677 6.70 10.0 0.555 40272 Iisvesi 27.021 62.679 17.2 18.0 0.456 40277 Hankavesi 26.826 62.614 7.00 18.0 0.100 40278 Petajävesi 25.173 62.255 2.70 3.0 0.245 40282 Kukkia 24.618 61.329 6.00 10.0 0.299 40308 Ähtärinjärvi 24.045 62.755 7.00 10.0 0.266 40313 Kuortaneenjärvi 23.407 62.863 7.00 10.0 0.277 40328 Lestijärvi 24.716 63.584 7.00 10.0 0.513 40330 Pyhäjärvi 25.995 63.682 7.00 10.0 0.266 40331 Lentua 29.690 64.204 7.60 7.0 0.600 40335 Lammasjärvi 29.551 64.131 4.40 3.0 0.200 40336 Naamankajärvi 28.246 65.104 7.00 7.0 0.299 40342 Korvuanjärvi 28.663 65.348 18.50 10.0 0.342 40343 Oijärvi 25.930 65.621 7.00 10.0 0.333 40345~~

Code and data availability. Observational data was obtained from SYKE open data archive SYKE, 2018 as follows: 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/. A supplementary file containing the freezing and melting dates as picked and prepared for the lakes studied here is attached. Data picked from HIRLAM archive are attached as supplementary files: data from the objective analysis feedback files (observed, analysed, forecast LSWT interpolated to the 27 active station locations) and from the gridded output of the HIRLAM analysis (analysed LSWT, forecast ice and snow thickness from the nearest gridpoint of all locations used in the present study).

In this study, FMI operational weather forecasts resulting from use of HIRLAM v.7.4 (rc1, with local updates) were validated against lake observations. The HIRLAM reference code is not open software but the property of the international HIRLAM-C programme. For research purposes, the codes can be requested from the programme (hirlam.org). The source codes of the version operational at FMI, relevant for the present study, are available from the authors upon request.

Author contributions. Laura Rontu computed the LSWT statistics based on HIRLAM feedback files. Kalle Eerola performed the freezing and melting date, snow and ice thickness comparisons based on data picked from HIRLAM grib files. Matti Horttanainen prepared observation data obtained via SYKE open data interface and lake depths from GLDB v.3. Laura Rontu composed the manuscript text based on input from all authors.

Competing interests. No competing interests are present.

Acknowledgements. Our thanks are due to Joni-Pekka Pietikäinen and Ekaterina Kourzeneva for discussions and information, to Margarita Choulga and Olga Toptunova for the support with the ~~GLDB lake depth data~~updated GLDB v.3 data, and to Emily Gleeson for advice with English language.

References

Balsamo, G., ~~R. Salgado~~Salgado, R., Dutra, E.~~Dutra~~, Boussetta, S.~~Boussetta~~, Stockdale, T. ~~Stockdale and M. Potes~~and Potes, M: On the contribution of lakes in predicting near-surface temperature in a global weather forecasting model. Tellus A, 64, 15829, https://doi.org/10.3402/tellusa.v64i0.15829, 2012.

Bengtsson, L., ~~U. Andrae, T.Aspelien~~Andrae, U., Aspelien, T., Batrak, Y.~~Batrak, ~~, Calvo, J.~~Calvo~~, ~~W.~~de Rooy, W., Gleeson, E.~~Gleeson~~, ~~B.~~Hansen-Sass, ~~M.~~B., Homleid, M.~~Hortal, ~~, Hortal, M., Ivarsson, K.~~Ivarsson, ~~, Lenderink, G.~~Lenderink~~, ~~S.Niemelä, K.~~Niemelä, S., Pagh Nielsen, K., Onvlee, J.~~Onvlee, ~~, Rontu, L.~~Rontu, ~~, Samuelsson, P.~~Samuelsson~~, ~~D.~~Santos Muñoz, D., Subias, A.~~Subias, ~~, Tijm, S.~~Tijm, ~~, Toll, V.~~Toll, ~~, Yang, X. ~~Yang, and M. ~~and Ødegaard Køltzow, M.: The HARMONIE-AROME model configuration in the ALADIN-HIRLAM NWP system. Monthly Wea. Rev., 145, 1919–1935, https://doi.org/10.1175/MWR-D-16-0417.1, 2018.

Cheng, B, Vihma. T, Rontu, L., Kontu A., Kheyrollah Pour, H., Duguay, C. and Pulliainen, J.: Evolution of snow and ice temperature, thickness and energy balance in Lake Orajärvi, northern Finland. Tellus A: Dynamic Meteorology and Oceanography ,66, 21564, https://doi.org/10.3402/tellusa.v66.21564, 2014.

Choulga, M., Kourzeneva, E., Zakharova, E. and Doganovsky, A.: Estimation of the mean depth of boreal lakes for use in numerical weather prediction and climate modelling. Tellus A, 66, 21295, https://doi.org/10.3402/tellusa.v66.21295, 2014.

Cordeira, J. M. and Laird, N. F. 2008. The influence of ice cover on two lake-effect snow events over lake Erie. *Mon. Wea. Rev.* **136**, 2747–2763. https://doi.org/DOI: 10.1175/2007MWR2310.1.

Elo, A-R.: Effects of climate and morphology on temperature conditions of lakes. University of Helsinki, Division of Geophysics, Report series in Geophysics. Available at http://urn.fi/URN:ISBN:978-952-10-3745-0, 2007.

Eerola, K., Rontu, L., Kourzeneva, E. and Shcherbak, E.: A study on effects of lake temperature and ice cover in HIRLAM. Boreal Env. Res., 15, 130–142, 2010.

Eerola, K.: Twenty-one years of verification from the HIRLAM NWP system. Wea. Forecasting, 28, 270–285. https://doi.org/10.1175/WAF-D-12-00068.1, 2013.

Eerola, K., Rontu, L., Kourzeneva, E., Kheyrollah Pour, H. and Duguay, C.: Impact of partly ice-free Lake Ladoga on temperature and cloudiness in an anticyclonic winter situation-a case study using a limited area model. Tellus A, 66, 23929. https://doi.org/10.3402/tellusa.v66.23929, 2014.

Gandin, L.: Objective analysis of meteorological fields. Gidrometizdat, Leningrad. Translated from Russian, Jerusalem, Israel Program for Scientific Translations, 1965.

Kheyrollah Pour, H., Rontu, L., Duguay, C. R., Eerola, K. and Kourzeneva, E.: Impact of satellite-based lake surface observations on the initial state of HIRLAM. Part II: Analysis of lake surface temperature and ice cover. Tellus A., 66, 21395. https://doi.org/10.3402/tellusa.v66.21395, 2014.

Kheyrollah Pour, H., Choulga, M., Eerola, K., Kourzeneva, E., Rontu, L., Pan F. and Duguay C. R.: Towards improved objective analysis of lake surface water temperature in a NWP model: preliminary assessment of statistical properties. Tellus A., 66, 21534. https://doi.org/10.1080/16000870.2017.1313025, 2017.

Korhonen J.: Suomen vesistöjen jääolot. Suomen ympäristö, (751), 145 pp (in Finnish). Available at http://hdl.handle.net/10138/40687, 2005.

Korhonen J.: Suomen vesistöjen lämpötilaolot 1900-luvulla. Suomen ympäristö, (566), 116 pp (in Finnish). Available at http://hdl.handle.net/10138/40478, 2002.

Kourzeneva, E., Samuelsson, P., Ganbat, G. and Mironov, D.: Implementation of lake model Flake into HIRLAM. HIRLAM Newsletter, (54), 54–64. Available at http://hirlam.org/, 2008.

Kourzeneva, E., Asensio, H., Martin, E. and Faroux, S.: Global gridded dataset of lake coverage and lake depth for use in numerical weather prediction and climate modelling. Tellus A., 64,15640. https://doi.org/10.3402/tellusa.v64i0.15640, 2012a.

Kourzeneva, E., Martin, E., Batrak, Y. and Moigne, P. L.:. Climate data for parameterisation of lakes in numerical weather prediction models. Tellus A., 64, 17226. https://doi.org/10.3402/tellusa.v64i0.17226, 2012b.

Kourzeneva, E.: Assimilation of lake water surface temperature observations with Extended Kalman filter. Tellus A., 66, 21510. https://doi.org/10.3402/tellusa.v66.21510, 2014.

Laird, N. F., Kristovich, D. A. R. and Walsh, J. E.: Idealized model simulations examining the mesoscale structure of winter lake-effect circulations. *Mon. Wea. Rev.* **131**, 206–221. https://doi.org/DOI: 10.1175/1520-0493(2003)131<0206:IMSETM>2.0.CO;2, 2003.

Leppäranta, M., Lindgren, E. and Shirasawa, K.: The heat budget of Lake Kilpisjarvi in the Arctic tundra. Hydrology research +research, 48, 969-980. https://doi.org/10.2166/nh.2016.171, 2012.

Lei, R., Leppäranta, M., Cheng, B., Heil P. and Li, Z.: Changes in ice-season characteristics of a European Arctic lake from 1964 to 2008. Climatic Change, 115, 725–739. https://doi.org/10.1007/s10584-012-0489-2, 2012.

Lindstedt, D., P. Lind, E. Kjellström and C. Jones.: A new regional climate model operating at the meso-gamma scale: performance over Europe. Tellus A, 67, 24138. https://doi.org/10.3402/tellusa.v67.24138, 2015.

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, J. and co authors: Development of a global land cover characteristics database and IGBP DISCover from 1-km AVHRR data. Int. J. Rem. Sens., 21, 1303–1130. Data and documentation available online at: http://edc2.usgs.gov/glcc/glcc.php, 2000.

Masson, V., P. Le Moigne, E. Martin, S. Faroux, S., and coauthorsAlias, A., Alkama, R., Belamari, S., Barbu, A., Boone, A., Bouyssel, F., Brousseau, P., Brun, E., Calvet, J.-C., Carrer, D., Decharme, B., Delire, C., Donier, S., Essaouini, K., Gibelin, A.-L., Giordani, H., Habets, F., Jidane, M., Kerdraon, G., Kourzeneva, E., Lafaysse, M., Lafont, S., Lebeaupin Brossier, C., Lemonsu, A., Mahfouf, J.-F., Marguinaud, P., Mokhtari, M., Morin, S., Pigeon, G., Salgado, R., Seity, Y., Taillefer, F., Tanguy, G., Tulet, P., Vincendon, B., Vionnet, V., and Voldoire, A.: The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes, Geosci. Model Dev., 6, 929–960. https://doi.org/:10.5194/gmd-6-929-2013, 2013.

Mironov, D., Heise, E., Kourzeneva, E., Ritter, B., Schneider, N. and Terzhevik, A.: Implementation of the lake parameterisation scheme FLake into the numerical weather prediction model COSMO. Boreal Env. Res., 15, 218–230, 2010.

Pietikäinen, J.-P., Markkanen, T., Sieck, K., Jacob, D., Korhonen, J., Räisänen, P., Gao, Y., Ahola, J., Korhonen, H., Laaksonen, A., and Kaurola, J.: The regional climate model REMO (v2015) coupled with the 1-D freshwater lake model FLake (v1): Fenno-Scandinavian climate and lakes. Geosci. Model Dev., 11, 1321–1342, https://doi.org/10.5194/gmd-11-1321-2018, 2018.

Rontu, L., Eerola, K., Kourzeneva, E. and Vehviläinen, B.: Data assimilation and parametrisation of lakes in HIRLAM. Tellus A., 64, 17611. https://doi.org/10.3402/tellusa.v64i0.17611, 2012.

Rontu L., Gleeson, E., Räisänen, P., Nielsen, K.P., Savijärvi, H. and Sass, B.H.: The HIRLAM fast radiation scheme for mesoscale numerical weather prediction models. Adv.Sci.Res., 14, 195–215, https://doi.org/10.5194/asr-14-195-2017, 2017.

Rooney, G. G. and Bornemann, F. J.: The performance of FLake in the Met Office Unified Model. Tellus A, 65, 21363, https://doi.org/10.3402/tellusa.v65i0.21363, 2013.

Samuelsson, P., Kourzeneva, E. and Mironov, D.: The impact of lakes on the European climate as stimulated by a regional climate model. Boreal Env. Res., 15, 113–129, 2010.

SYKE. ISO 19115 Geographic Information Metadata on Hydrological data. Available at http://metatieto.ymparisto.fi:8080/geoportal/catalog/search/resource/details.page?uuid=%7B86FC3188-6796-4C79-AC58-8DBC7B568827%7D, 2018.

Termonia, P., Fischer, C., Bazile, E., Bouyssel, F., Brožková, R., Bénard, P., Bochenek, B., Degrauwe, D., Derkova, M., El Khatib, R., Hamdi, R., Mašek, J., Pottier, P., Pristov, N., Seity, Y., Smolíková, P., Spaniel, O., Tudor, M., Wang, Y., Wittmann, C., and Joly, A.: The ALADIN System and its canonical model configurations AROME CY41T1 and ALARO CY40T1, Geosci. Model Dev., 11, 257-281, https://doi.org/10.5194/gmd-11-257-2018, 2018.

Undén, P., Rontu, L., Järvinen, H., Lynch, P., Calvo, J., Cats, G., Cuxart, J., Eerola, K., Fortelius, C., Garcia-Moya, J. A. et al.: HIRLAM-5 scientific documentation. Available at http://www.hirlam.org, 2002.

Yang, Y., Cheng, B., Kourzeneva, E., Semmler, T., RontuL. and coauthors: 2013. , L., Leppäranta, M., Shirasawa, K. and Li, Z. J.: Modelling experiments on air-snow-ice interactions over Kilpisjärviair–snow–ice interactions over Kilpisjärvi, a lake in northern Finland. Boreal EnvironEnv. Res. , 18, 341–358. : 341–358, 2013.

Table A1. Prognostic and diagnostic lake variables within HIRLAM

variable	unit	type
temperature of snow on lake ice	K	prog by FLake
temperature of lake ice	K	prog by FLake
mean water temperature	K	prog by FLake
mixed layer temperature	K	prog by FLake
bottom temperature	K	prog by FLake
temperature of upper layer sediments	K	prog by FLake
mixed layer depth	m	prog by FLake
thickness of upper layer sediments	m	prog by FLake
thermocline shape factor	-	prog by FLake
lake ice thickness	m	prog by FLake
snow depth on lake ice	m	prog by FLake
LSWT	K	diag by FLake
		= mixed layer temperature if no ice
lake surface temperature	K	diag by FLake
		uppermost temperature: LSWT or ice or snow
LSWT	K	anal by HIRLAM
		flag value 272 K when there is ice
fraction of lake ice	[0-1]	diag fraction in HIRLAM grid
lake surface roughness	m	diag by HIRLAM
screen level temperature over lake	K	diag by HIRLAM
screen level abs.humidity over lake	$kgkg^{-1}$	diag by HIRLAM
anemometer level u-component over lake	ms^{-1}	diag by HIRLAM
anemometer level v-component over lake	ms^{-1}	diag by HIRLAM
latent heat flux over lake	Wm^{-2}	diag by HIRLAM
sensible heat flux over lake	Wm^{-2}	diag by HIRLAM
scalar momentum flux over lake	Pa	diag by HIRLAM
SW net radiation over lake	Wm^{-2}	diag by HIRLAM
LW net radiation over lake	Wm^{-2}	diag by HIRLAM
depth of lake	m	pres in HIRLAM grid
fraction of lake	[0-1]	pres in HIRLAM grid

Denotation: prog = prognostic, diag = diagnostic, pres = prescribed, anal = result of OI

Table A2. Lakes with SYKE observations used in this study.

NAME	LON	LAT	MEAND (m)	MAXD (m)	AREA (kgm^{-2})	HIRD (m)	HIRFR	HIRID
Pielinen	29.607	63.271	10.1	61.0	894.2	10.0	0.916	4001
Kallavesi	27.783	62.762	9.7	75.0	316.1	10.0	0.814	4002
Haukivesi	28.389	62.108	9.1	55.0	560.4	10.0	0.725	4003
Saimaa	28.116	61.338	10.8	85.8	1,377.0	10.0	0.950	4004
Pääjärvi1	24.789	62.864	3.8	14.9	29.5	3.0	0.430	4005
Nilakka	26.527	63.115	4.9	21.7	169.0	10.0	0.866	4006
Konnevesi	26.605	62.633	10.6	57.1	189.2	10.0	0.937	4007
Jääsjärvi	26.135	61.631	4.6	28.2	81.1	10.0	0.750	4008
Päijänne	25.482	61.614	14.1	86.0	864.9	10.0	0.983	4009
Ala-Rieveli	26.172	61.303	11.3	46.9	13.0	10.0	0.549	4010
Kyyvesi	27.080	61.999	4.4	35.3	130.0	10.0	0.810	4011
Tuusulanjärvi	25.054	60.441	3.2	9.8	5.9	3.0	0.174	4012
Pyhäjärvi	22.291	61.001	5.5	26.2	155.2	5.0	0.922	4013
Längelmävesi	24.370	61.535	6.8	59.3	133.0	10.0	0.875	4014
Pääjärvi2	25.132	61.064	14.8	85.0	13.4	14.0	0.350	4015
Vaskivesi	23.764	62.142	7.0	62.0	46.1	10.0	0.349	4016
Kuivajärvi	23.860	60.786	2.2	9.9	8.2	10.0	0.419	4017
Näsijärvi	23.750	61.632	14.7	65.6	210.6	10.0	0.850	4018
Lappajärvi	23.671	63.148	6.9	36.0	145.5	10.0	1.000	4019
Pesiöjärvi	28.650	64.945	3.9	15.8	12.7	7.0	0.290	4020
Rehja-Nuasjärvi	28.016	64.184	8.5	42.0	96.4	10.0	0.534	4021
Oulujärvi	26.965	64.451	6.9	35.0	887.1	10.0	1.000	4022
Ounasjärvi	23.602	68.377	6.6	31.0	6.9	10.0	0.166	4023
Unari	25.711	67.172	5.0	24.8	29.1	10.0	0.491	4024
Kilpisjärvi	20.816	69.007	19.5	57.0	37.3	22.0	0.399	4025
Kevojärvi	27.011	69.754	11.1	35.0	1.0	10.0	0.016	4026
Inarijärvi	27.924	69.082	14.3	92.0	1,039.4	14.0	0.979	4027
Simpelejärvi	29.482	61.601	9.3	34.4	88.2	10.0	0.548	40241
Pökkäänlahti	27.264	61.501	8.0	84.3	58.0	10.0	0.299	40261
Muurasjärvi	25.353	63.478	9.0	35.7	21.1	10.0	0.060	40263
Kalmarinselkä	25.001	62.786	5.7	21.9	7.1	5.0	0.330	40271
Summasjärvi	25.344	62.677	6.7	40.5	21.9	10.0	0.555	40272
Iisvesi	27.021	62.679	17.2	34.5	164.9	18.0	0.456	40277
Hankavesi	26.826	62.614	7.0	49.0	18.2	18.0	0.100	40278
Petajävesi	25.173	62.255	4.2	26.6	8.8	3.0	0.245	40282
Kukkia	24.618	61.329	5.2	35.6	43.9	10.0	0.299	40308
Ähtärinjärvi	24.045	62.755	5.2	27.0	39.9	10.0	0.266	40313
Kuortaneenjärvi	23.407	62.863	3.3	16.2	14.9	10.0	0.277	40328
Lestijärvi	24.716	63.584	3.6	6.9	64.7	10.0	0.513	40330

Validation of lake surface state in the HIRLAM v.7.4 NWP model against *in-situ* measurements in Finland

Laura Rontu[1], Kalle Eerola[1], and Matti Horttanainen[1]

[1]Finnish Meteorological Institute, P.O. Box 503, 00101 Helsinki, Finland

Correspondence: laura.rontu@fmi.fi

Abstract. High Resolution Limited Area Model (HIRLAM), used for the operational numerical weather prediction in the Finnish Meteorological Institute (FMI), includes prognostic treatment of lake surface state since 2012. Forecast is based on the Freshwater Lake (FLake) model integrated ~~to~~ into HIRLAM. Additionally, an independent objective analysis of lake surface water temperature (LSWT) combines the short forecast of FLake to observations from the Finnish Environment Institute (SYKE). The resulting description of lake surface state - forecast FLake variables and analysed LSWT - was compared to SYKE observations of lake water temperature, ~~freezing and melting~~ freeze-up and break-up dates as well as the ice ~~and snow thickness~~ thickness and snow depth for 2012-2018 over 45 lakes in Finland. During the ice-free period, the predicted LSWT corresponded to the observations with a slight overestimation, with a systematic error of + 0.91 K. The colder temperatures were underrepresented and the maximum temperatures were too high. The objective analysis of LSWT was able to reduce the bias to + 0.35 K. The predicted ~~freezing~~ freeze-up dates corresponded well the observed dates, mostly within the accuracy of a week. The forecast ~~melting~~ break-up dates were far too early, typically several weeks ahead of the observed dates. The growth of ice thickness after ~~freezing~~ freeze-up was generally overestimated. However, practically no predicted snow appeared on lake ice. The absence of snow, ~~found to~~ presumably be due to ~~a technical error in HIRLAM~~an incorrect security coefficient value, is suggested to be also the main reason of the inaccurate simulation of the ~~ice melt~~ lake ice melting in spring.

1 Introduction

Lakes influence the energy exchange between the surface and the atmosphere, the dynamics of the atmospheric boundary layer and the near-surface weather. This is important for weather forecasting over the areas where lakes, especially those with a large yearly variation of the water temperature, freezing in autumn and melting in spring, cover a significant area of the surface (Kheyrollah Pour et al., 2017; Laird et al., 2003 and references therein). Description of the lake surface state influences the numerical weather prediction (NWP) results, in particular in the models whose resolution is high enough to account for even the smaller lakes (Eerola et al., 2014 and references therein). Especially, the existence of ice can be important for the numerical forecast (Eerola et al., 2014; Cordeira and Laird, 2008).

~~The~~ In the Finnish Meteorological Institute (FMI), the High Resolution Limited Area Model HIRLAM (Undén et al., 2002; Eerola, 2013) has been applied since 1990 for the numerical short-range weather forecast~~over the Northern Europe~~. In the beginning, the monthly climatological water surface temperature for both sea (sea surface temperature SST) and lakes (Lake Surface Water Temperature LSWT) was used. Since 2012, HIRLAM includes a prognostic lake temperature parameterization based on the Freshwater Lake Model (FLake, Mironov et al., 2010). An independent objective analysis of observed LSWT (~~Kheyrollah Pour et al. (2017)~~ Kheyrollah Pour et al., 2017 and references therein) was implemented in 2011. ~~Fractional~~ The fractional ice cover (lake ice concentration in each ~~grid-square~~ gridsquare of the model) is ~~estimated separately based on~~ diagnosed from the analysed LSWT~~and the ice thickness predicted by FLake~~.

FLake was designed to be used as a parametrization scheme for the forecast of the lake surface state in NWP and climate models. It allows to predict the lake surface state in interaction with the atmospheric processes treated by the NWP model. The radiative and turbulent fluxes as well as the predicted snow precipitation from the atmospheric model are combined with FLake processes at each time-step of the model integration ~~(with a typical interval of one or several minutes)~~ in the model grid, where the fraction and depth of lakes are prescribed.

FLake has been implemented into the other main European NWP and regional climate models, first into COSMO (Mironov et al., 2010) then into ECMWF (Balsamo et al., 2012), Unified Model (Rooney and Bornemann, 2013), SURFEX surface modelling framework (Masson et al., 2016), regional climate models RCA (Samuelsson et al., 2010), HCLIM (Lindstedt et al., 2015) and REMO (Pietikäinen et al., 2018), among others. Description of lake surface state and its influence in the numerical weather and climate prediction has been validated in various ways. Results of case studies, e.g. Eerola et al. (2014) and shorter-period NWP experiments, e.g. Eerola et al. (2010); Rontu et al. (2012); Kheyrollah Pour et al. (2014); Kheyrollah Pour et al. (2017) as well as climate model results, e.g. ~~Samuelsson et al. (2010); Pietikäinen et al. (2018)~~ Samuelsson et al. (2010); Pietikäinen et al. (2018), have been compared with remote-sensing satellite data and ~~*in situ*~~*in-situ* lake temperature and ice measurements as well as validated against the standard weather observations. In general, improvement of the scores has been seen over regions where lakes occupy a significant area. However, specific features of each of the host models influence the results of the coupled atmosphere-lake system as FLake appears to be quite sensitive to the forcing by the atmospheric model.

The aim of the present study is to ~~use~~ validate the lake surface state forecast by the operational HIRLAM NWP model using the ~~*in situ*~~*in-situ* LSWT measurements, lake ~~freezing and melting~~ ice freeze-up and break-up dates and measurements of ice and snow thickness by the Finnish Environment Institute (Suomen Ympäristökeskus = SYKE)~~for validation of the lake surface state forecast by the operational HIRLAM NWP model~~. For this purpose, HIRLAM analyses and forecasts archived by ~~the Finnish Meteorological Institute (FMI)~~ FMI were compared with the observations by SYKE over the lakes of Finland from spring 2012 to summer 2018. To our knowledge, this is the longest available detailed dataset that allows to evaluate how well the lake surface state is simulated by an operational NWP model that applies FLake parametrizations.

2 Lake surface state in HIRLAM

FLake was implemented in the HIRLAM forecasting system in 2012 (Kourzeneva et al., 2008; Eerola et al., 2010). The model utilizes external datasets on the lake depth (Kourzeneva et al., 2012a; Choulga et al., 2014) and the lake climatology (Kourzeneva et al., 2012b). The latter is only needed in order to provide initial values of FLake prognostic variables in the very first forecast ~~. Real-time~~ (so-called cold start). The use of real-time *in-situ* LSWT observations by SYKE for 27 Finnish lakes ~~became available for the operational HIRLAM analysis in~~ was introduced in 2011 into the operational LSWT analysis in HIRLAM (Eerola et al., 2010; Rontu et al., 2012). In the current operational HIRLAM ~~at FMI~~of FMI, FLake provides the background for the optimal interpolation analysis (OI, based on ~~Gandin 1965) analysis~~ Gandin, 1965) of LSWT. However, the prognostic FLake variables are not corrected using the analysed LSWT~~, which would require~~. This would require more advanced data assimilation methods based on e.g. the extended Kalman filter (Kourzeneva, 2014). ~~The relations between the OI analysis and the prognostic FLake in HIRLAM are schematically illustrated in Figure 1.~~

~~Coexistence of the independent objective analysis of the observed LSWT and prognostic FLake parametrizations in HIRLAM.~~

2.1 Freshwater lake model in HIRLAM

FLake is a bulk model capable of predicting the vertical temperature structure and mixing conditions in lakes of various depths on time-scales from hours to years (Mironov et al., 2010). The model is based on two-layer parametric representation of the evolving temperature profile in the water and on the integral budgets of energy for the layers in question. Bottom sediments and the thermodynamics of the ice and snow on ice layers are treated separately. FLake depends on prescribed lake depth information. The prognostic and diagnostic variables of HIRLAM ~~/FLake plus~~ FLake together with the analysed lake surface variables in HIRLAM are listed in the Appendix (Table A1).

At each time step ~~of~~ during the HIRLAM forecast, FLake is driven by the atmospheric radiative and turbulent fluxes as well as the predicted snowfall, provided by the physical parameterisations in HIRLAM. This couples the atmospheric variables over lakes with the lake surface properties as provided by FLake ~~.~~ parametrization. Most importantly, FLake provides HIRLAM with the evolving lake surface (water, ice, snow) temperature and radiative properties, that influence the HIRLAM forecast of the grid-average near-surface temperatures.

Implementation of FLake model as a ~~parametrizations~~ parametrization scheme in HIRLAM was based on the experiments described by Rontu et al. (2012). Compared to the reference version of FLake (Mironov et al., 2010), minor modifications were introduced, namely, use of constant snow density = 300 $\mathrm{kgm^{-3}}$, molecular heat conductivity = 1 $\mathrm{Jm^{-1}s^{-1}K^{-1}}$, constant albedos of dry snow = 0.75 and ice = 0.5. Bottom sediment calculations were excluded. Global lake depth database (GLDB v.2, ~~Choulga et al. (2014)) is~~ Choulga et al., 2014) was used for derivation of mean lake depth in each gridsquare. Fraction of lake ~~is~~ was taken from HIRLAM physiography database, where it originates from GLCC (Loveland et al., 2000).

~~In this study, lake surface temperature and thickness of ice predicted by FLake were used for the model-observation comparison.~~ Lake surface temperature is diagnosed from the mixed layer temperature for the unfrozen lake gridpoints and from the ice or snow-on-ice temperature for the frozen points. In FLake, ice starts to grow from an assumed value of one millimeter when temperature reaches the freezing point. The whole lake tile in a gridsquare is considered by FLake either frozen or unfrozen. Snow on ice is accumulated from the model's snowfall at each time step during the numerical integration.

2.2 Objective analysis of LSWT observations

A comprehensive description of the optimal interpolation (OI) of the LSWT observations in HIRLAM is given by Kheyrollah Pour et al. (20 (Kheyrollah Pour et al., 2017). Shortly, LSWT analysis is obtained by correcting the FLake forecast at each gridpoint by using the weighted average of the deviations of observations from their background values. Prescribed statistical information about the observation and background error variance as well as the distance-dependent autocorrelation between the locations (observations and gridpoints) are applied. The real-time observations entering the HIRLAM surface analysis system are subject to quality control in two phases. First, the observations are compared to the background, provided by the FLake short forecast. Second, optimal interpolation is done at each observation location, using the neighbouring observations only (excluding the current observation) and comparing the result to the observed value at the station.

A specific feature of the lake surface temperature OI is that the interpolation is performed not only within the (large) lakes but also across the lakes: within a statistically pre-defined radius, the observations affect all gridpoints containing a fraction of lake. This ensures that the analysed LSWT on lakes without own observations may also be influenced by observations from neighbouring lakes, not only by the first guess provided by FLake forecast.

The relations between the OI analysis and the prognostic FLake in HIRLAM are schematically illustrated in Figure 1. Within the present HIRLAM setup, the background for the analysis is provided by the short (6-hour) FLake forecast . However, but the next forecast is not initialised initialized from the analysis, see Figure 1. Instead, FLake continues running from the previous forecast, driven by the atmospheric state given by HIRLAM at each time step. This means that FLake does not benefit from the result of OI analysis does not benefit FLake but the analysis remains to some extent as an extra diagnostic field, to some extent independent of the LSWT forecast. Note that However, FLake background has a large influence in the analysis, especially over distant lakes where neighbouring observations are not available. The diagnostic LSWT analysis, available at every gridpoint of HIRLAM, might be useful e.g. for hydrological, agricultural or road weather applications.

Missing LSWT observations in spring and early winter are interpreted to represent presence of ice and given a flag value of -1.2°C. If, however, the results of the statistical moving-average-type LSWT model (Elo, 2007), provided by SYKE along with the real-time observations, indicate unfrozen conditions, the observations are considered missing. This prevents appearance of ice in summer when observations are missing but leads to a misinterpretation of data in spring if the SYKE model indicates too early melting. In the analysis, fraction of ice is diagnosed from the LSWT field in a simple way. The lake surface within a gridsquare is assumed fully ice-covered when LSWT falls below -0.5°C and fully ice-free when LSWT is above 0°C. Between these temperature thresholds, the fraction of ice changes linearly (Kheyrollah Pour et al., 2014).

The HIRLAM surface data assimilation system produces comprehensive feedback information from every analysis-forecast cycle. The feedback consists of the observed value and its deviations from the background and from the final analysis at the observation point. Bilinear interpolation of the analysed and forecast values is done to the observation location from the

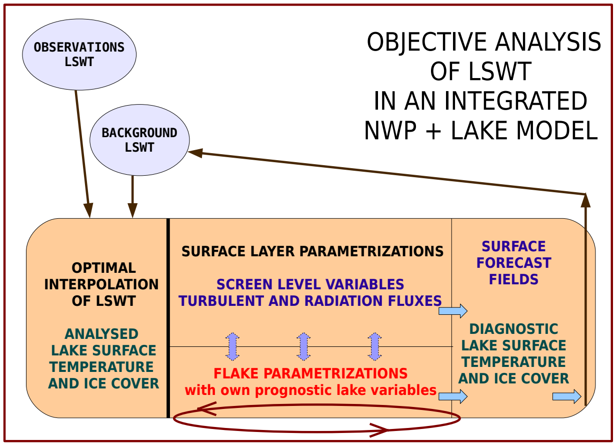


Figure 1. Coexistence of the independent objective analysis of the observed LSWT and prognostic FLake parametrizations in HIRLAM. The thin arrows are related to data flow between HIRLAM analysis-forecast cycles while the thick arrows describe processes within each cycle.

nearest gridpoints that contain a fraction of lake. In addition, information about the quality check and usage of observations is provided. Fractions of land and lake in the model grid as well as the weights, which were used to interpolate gridpoint values to the observation location, are given. ~~We use this information as basic material~~ This information is the basis of the present study (see sections 3.3 and 4).

3 Model-observation intercomparison 2012-2018

In this intercomparison we validated HIRLAM ~~/FLake~~ results against observations about the lake surface state. The impact of FLake parametrizations to the weather forecast by HIRLAM ~~is~~ was not considered. This is because ~~the archived observations and the operational HIRLAM results were used during the period from spring 2012 to summer 2018 when FLake was always an integral part of HIRLAM. This means that there are no~~ no non-FLake ~~weather forecasts to compare with~~weather forecasts exist for comparison with the operational forecasts during the validation period.

Throughout the following text, the analysed LSWT refers to the result of OI analysis, where FLake forecast has been used as background (Section 2.2) while the forecast LSWT refers to the value diagnosed from the mixed layer water temperature predicted by FLake (Section 2.1). Observed LSWT refers to the measured by SYKE lake water temperature (Section 3.2).

3.1 FMI operational HIRLAM

FMI operational HIRLAM is based on the last reference version (v.7.4), implemented in spring 2012. (~~Eerola (2013)~~ Eerola, 2013 and references therein). FLake was introduced into this version. After that the development of HIRLAM was frozen. Thus, during the years of the present comparison, the FMI operational HIRLAM system remains unmodified, which offers a clean

Table 1. FMI operational HIRLAM

Domain	From Atlantic to Ural, from North Africa beyond North Pole
Model horizontal / vertical resolution	7 km / 65 levels
HIRLAM version	7.4
Model dynamics	Hydrostatic, semi-Lagrangian, grid-point
Atmospheric physical parametrizations	Savijärvi radiation, CBR turbulence,
	Rasch-Kristiansson cloud microphysics + Kain-Fritsch convection
Surface physical parametrizations	ISBA-newsnow for surface, FLake for lakes
Data assimilation	Default atmospheric (4DVAR) and surface (OI) analysis
Lateral boundaries	ECMWF forecast
Forecast	Up to +54 h initiated every 6h (00, 06, 12, 18 UTC)

time series of data for the model-observation intercomparison. The general properties of the system are summarised in Table 1. In the present study, a coding error in FLake implementation was revealed in the reference HIRLAM v.7.4. A too large critical value to diagnose snow existence prevented practically all accumulation of the forecast snowfall on lake ice in the FMI HIRLAM-FLake operational system.

3.2 SYKE lake observations

In this study we used three different types of SYKE lake observations: LSWT, freeze-up and break-up dates and ice thickness and snow depth on lake ice. In total, observations on 45 lakes listed in Appendix (Table A2) were included as detailed in the following. The lake depths and surface areas given in Table A2 are based on he updated lake list of GLDB v.3 (Margarita Choulga, personal communication).

3.2.1 Lake temperature measurements

Regular *in-situ* lake water temperature (LWT) measurements are performed by SYKE. SYKE operates 32 Currently SYKE operates 34 regular lake and river water temperature measurement sites in Finland. The temperature of the lake water is measured every morning at 8.00 AM local time, close to shore, at 20 cm below the water surface. The measurements are recorded either automatically or manually and are performed only during the ice-free season (?Rontu et al., 2012)(Korhonen, 2019). Further, we will for simplicity denote also these data as LSWT observations although they do not represent exactly the same surface water temperature (skin temperature, radiative temperature) that could be estimated by satellite measurements. These data are available in the SYKE open data archive (SYKE, 2018). Measurements from 27 of these 34 lakes (Figure 2, white dots) used by were selected for use in the FMI operational HIRLAM , were included in 2011, and the list has been kept unmodified since that. The set of 27 daily observations, quality-controlled by HIRLAM, were obtained from the analysis feedback files

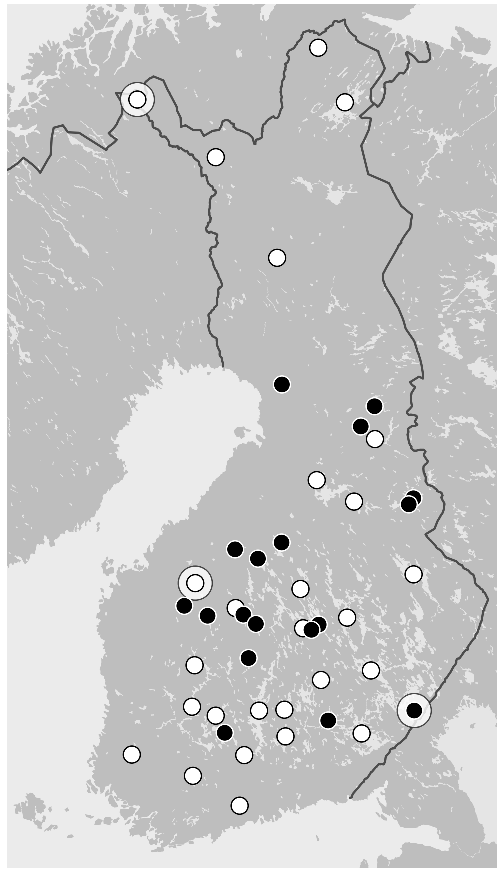


Figure 2. Map of SYKE observation points used in this study: lakes with both lake surface water temperature (LSWT) and lake ice date (LID) observations (white), lakes where only LID is available (black). On Lakes Lappajärvi, Kilpisjärvi and Simpelejärvi also ice thickness and snow ~~thickness~~ depth measurements were used (Section 4.3), they are surrounded with a large white circle. List of the lakes with coordinates is given in Appendix A2.

and used in all comparisons reported in this study. ~~These data are also available in the SYKE open data archive (SYKE, 2018)~~ ~~.~~

3.2.2 ~~Freezing~~ Freeze-up and ~~melting~~ break-up dates

Regular visual observations of ~~freezing and melting~~ freeze-up and break-up of lakes have been recorded in Finland for centuries, the longest time series starting in the middle of the 19th century ~~(?).~~ (Korhonen, 2019). Presently, dates of ~~freezing and melting~~ freeze-up and break-up are available from SYKE (2018) on 123 lakes, but the time series for many lakes are discontinuous. Further, we will denote the ~~melting and freezing~~ break-up and freeze-up dates together by “lake ice dates” (LID). ~~For both freezing and melting~~ LID observations aim at representing conditions on entire lakes. For both freeze-up and break-up the dates are available in two categories ~~: for freezing “freezing of the visible area”~~ (terminology from Korhonen, 2019): "freeze-up of the lake within sight" (code 29 by SYKE) and "permanent freezing of the visible area" "freeze-up of the whole lake" (code 30). For melting break-up the dates are defined as "no ice visible from the observation site" "no ice within sight" (code 28) and "no ice on the outer open water areas thaw areas out of the shore" (code 27). LID observations by SYKE are made independently of their LSWT measurements and possibly from different locations on the same lakes. The LSWT measurements may be started later than the date of reported lake ice break-up or end earlier than the reported freeze-up date.

LID from the 27 lakes whose LWT LSWT measurements are used in HIRLAM were available and selected for this study. In addition, 18 lakes with only LID available (Figure 2, black dots) were chosen for comparison with HIRLAM /FLake LID.

3.2.3 Ice thickness and snow thickness depth on lakes

SYKE records In the period 2012-2018 SYKE recorded the lake ice and snow thickness thickness and snow depth on around 50 locations in Finland, archived . (Archived historical data are available in total from 160 measurement sites). The manual measurements are done three times a month during the ice season. Thickness of ice and the snow snow depth on ice are measured by drilling holes through snow and ice layers along chosen tracks, normally at least 50 m from the coast (?)(Korhonen, 2019). The locations may differ from those of the LSWT measurement or LID observation over the same lakes. In this study, measurements from lakes Lappajärvi, Kilpisjärvi and Simplejärvi were utilised as additional data for validation in Section 4.3. These lakes, sufficiently large in order to fit well the HIRLAM grid, represent the western, northern and south-eastern Finland.

3.3 Lake Validation of HIRLAM lake surface statederived from HIRLAM output

3.3.1 Lake surface water temperature

Diagnosed LSWT from HIRLAM/FLake analysis and forecastcycles LSWT by HIRLAM, resulting from the objective analysis or diagnosed from the forecast, was compared with the observed LWT LSWT by SYKE using data extracted from the analysis feedback files (Section 2.2) at the observation locations on 06 UTC every day. , excluding the winter periods 1 December - 31 March. The observations (ob) at 27 SYKE stations were assumed to represent the true value, while the analysis (an) is the result of OI that combines the background forecast (fc) with the observations. Time-series, maps and statistical scores, to be presented in Section 4.1, were derived from these.

3.3.2 Freezing and melting dates

3.3.2 Lake ice conditions

Both the analysed LSWT and the lake ice thickness forecast by FLake were separately used to define LID. The values For this study, the observed LID, ice and snow thickness observations were obtained from SYKE open data base, relying on their quality control. The analysed LSWT as well as the predicted ice thickness and snow depth were picked afterwards from the HIRLAM archive for a single gridpoint nearest to each of the 45 observation locations (not interpolated as in the analysis feedback file that was used for the LSWT comparison). ~~For the definition of LID, it~~ It was assumed that the gridpoint value nearest to the location of the LSWT observation represents the ice conditions over the chosen lake.

LID given by HIRLAM were defined in two independent ways: from the analysed LSWT and from the forecast lake ice thickness. Note that the ice thickness and snow depth on ice are not analysed variables in HIRLAM. In autumn a lake can freeze and melt several times before final ~~freezing~~freeze-up. The last date when the forecast ice thickness crossed a critical value of 1 mm or the analysed LSWT fell below freezing point was selected as the date of ~~freezing. To decrease the effect of oscillation of the gridpoint values between the HIRLAM forecast-analysis cycles, the mean of the four daily ice thickness forecasts or analysed LSWT values was used~~freeze-up. In the same way, the last date when the forecast ice thickness fell below the critical value of 1 mm or the analysed LSWT value crossed the freezing point was selected as ~~melting day.~~

3.4 ~~Validation methods~~

~~For LSWT statistics we used data collected during the HIRLAM surface analysis at each active observation location (Section 2.2), excluding the winter periods 1 December - 31 March. The observations (ob) at 27 SYKE stations were assumed to represent the true value, while the analysis (an) is the result of OI that combines the background forecast (fc) with the observations. Time-series, maps and statistical scores, to be presented in Section 4.1, were derived from these~~break-up date. To decrease the effect of oscillation of the gridpoint values between the HIRLAM forecast-analysis cycles, the mean of the four daily ice thickness forecasts or analysed LSWT values was used.

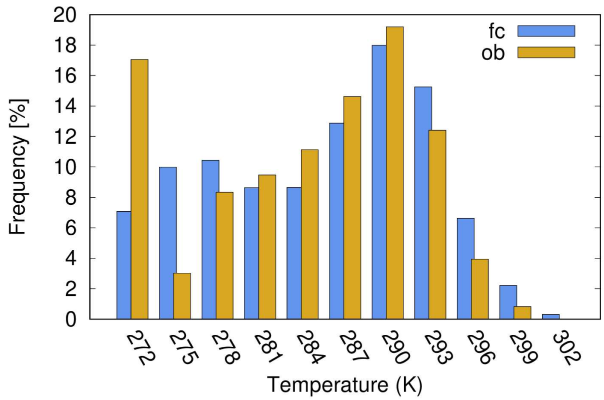
LID by HIRLAM ~~/FLake~~ were compared to the observed dates during 2012-2018~~, including in the comparison data over all months~~. In this comparison we included data also during the winter period. The category 29 observations ("freeze-up of the lake within sight", see Section 3.2.2) were used. In this category the time series were the most complete at the selected stations. For the same reason, the ~~melting~~ break-up observations of category 28 ("no ice within sight") were used for comparison. Furthermore, using a single gridpoint value for the calculation of LID also seems to correspond best the observation definition based on what is visible from the observation site. The statistics were calculated as ~~ob~~fc - ~~fc and ob~~ ob and an - ~~an~~ob. Hence, positive values mean that ~~melting or freezing~~ break-up or freeze-up takes place too late in the model as compared to the observations.

Lake ice thickness and snow depth measurements from lakes Lappajärvi, Kilpisjärvi and Simpelejärvi were utilised as additional data for validation of predicted by HIRLAM ice thickness and snow depth (Section 4.3). These lakes, representing the western, northern and south-eastern Finland, were selected for illustration based on the best data availability during the study years. They are also sufficiently large in order to fit well the HIRLAM grid.

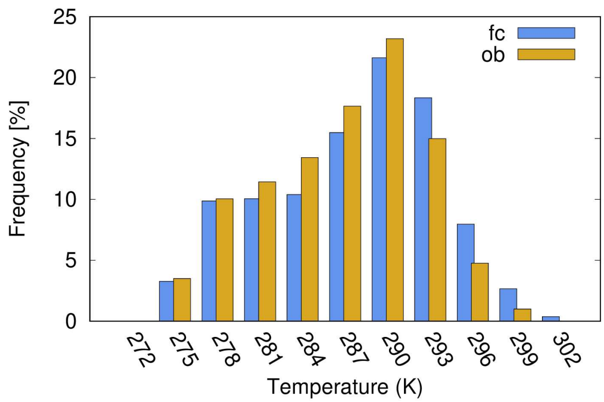
4 Results

4.1 Analysed and forecast LSWT at observation points

Figure 3 shows the frequency distribution of LSWT according to FLake forecast and SYKE observations. It is evident that the amount of data in the class of temperatures which represents frozen conditions (LSWT flag value 272 K) ~~is~~ was underestimated by the forecast (Figure 3a). When subzero temperatures ~~are~~ were excluded from the comparison (Figure 3b), underestimation in the colder temperature classes and overestimation in the warmer classes still remains.



(a) with all temperatures (also frozen conditions) included



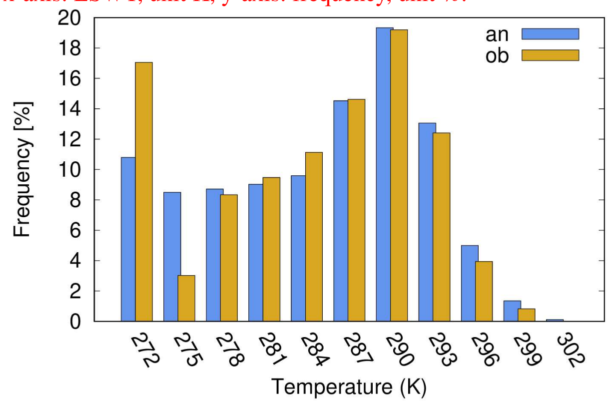
(b) only open water temperatures included

Figure 3. Frequency of observed (ob, yellow) and forecast (fc, blue) LSWT over all 27 SYKE lakes 2012-2018. x-axis: LSWT, unit K, y-axis: frequency, unit %.

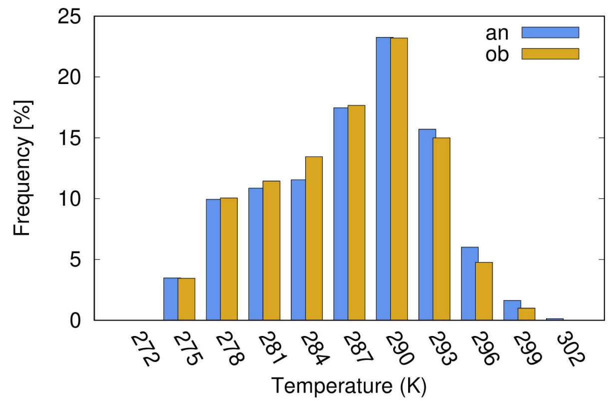
LSWT analysis (Figure 4) ~~improves~~ improved the situation somewhat but the basic features remain. This is due to the dominance of FLake forecast via the background of the analysis. In Section 4.3, we will show time-series illustrating the physics behind these LSWT statistics.

~~with all temperatures (also frozen conditions) included~~

~~only open water temperatures included Frequency of observed (yellow) and forecast (blue) LSWT over all 27 SYKE lakes 2012-2018. x-axis: LSWT, unit K, y-axis: frequency, unit %.~~



(a) with all temperatures (also frozen conditions) included



(b) only open water temperatures included

Figure 4. As for Figure 3 but for observed and analysed (an) LSWT.

Table 2 confirms the warm bias by FLake in the unfrozen conditions. Similar results were obtained for all stations together and also for our example lakes Lappajärvi and Kilpisjärvi, to be discussed in detail in Section 4.3. There were three lakes with negative LSWT bias according to FLake forecast, namely the large lakes Saimaa and Päijänne ~~,~~ and the smaller Ala-Rieveli. After the correction by objective analysis, a small positive bias converted to negative over 6 additional lakes, among them the large lakes Lappajärvi in the west and Inari in the north. The mean absolute error decreased from forecast to analysis ~~in~~ on every lake.

In the frequency distributions, the warm temperatures ~~are~~ were evidently related to summer. For FLake, the overestimation of maximum temperatures, especially in shallow lakes, is a ~~knowm~~ known feature (e.g. Kourzeneva 2014). It is related to the

Table 2. Statistical scores for LSWT at all stations and at two selected stations

station	fc or an	mean ob	bias	mae	stde	N
unit		K	K	K	K	
ALL	fc	286.3	0.91	1.94	2.34	30877
	an	286.3	0.35	1.32	1.72	30861
Lappajärvi	fc	286.9	0.33	1.23	1.62	1243
	an	286.9	-0.65	1.06	1.10	1243
Kilpisjärvi	fc	281.7	1.82	2.13	2.15	780
	an	281.7	1.10	1.42	1.51	780

Statistics over days when both forecast/analysis and observation indicate unfrozen conditions. bias = systematic difference fc/an - ob, mae = mean absolute error, stde = standard deviation of the error, N = number of days (06 UTC comparison, no ice).

difficulty of forecasting the mixed layer thermodynamics under strong solar heating. Cold and subzero temperatures ~~occur~~ occurred in spring and autumn. In a few large lakes like Saimaa, Haukivesi, Pielinen, LSWT ~~tends~~ tended to be slightly underestimated in autumn both according to the FLake and the analysis (not shown). ~~However, as will be shown in Sections 4.2 and 4.3, the~~ The cold left-hand side columns in the frequency distributions (Figures 3a and 4a) are mainly related to spring, when HIRLAM ~~/FLake tends~~ tended to melt the lakes significantly too early (Sections 4.2 and 4.3).

There are problems, especially in the analysed LSWT, over (small) lakes of irregular form that fit poorly the HIRLAM grid and where the measurements may represent more the local than the mean or typical conditions over the lake. These are the only ones where an underestimation of summer LSWT ~~can be~~ was seen. Cases ~~occur~~ occurred where FLake results differ so much from the observations that the ~~quality control of the HIRLAM surface data assimilation rejects~~ HIRLAM quality control against background values rejected the observations, forcing also the analysis to follow the incorrect forecast (not shown).

4.2 ~~Freezing~~ Freeze-up and ~~melting~~ break-up dates

In this section the freeze-up and break-up dates from HIRLAM are verified against corresponding observed dates over 45 lakes (Appendix Table A2). In the following, 'LSWT an' refers to the LID estimated from analysed LSWT and 'IceD fc' to those estimated from the forecast ice thickness by FLake. The time period contains six freezing periods (from autumn 2012 to autumn 2017) and seven melting periods (from spring 2012 to spring 2018). Due to some missing data the number of freeze-up cases was 233 and break-up cases 258. The 'IceD fc' data for the first melting period in spring 2012 was missing. The overall statistics of the error in freeze-up and break-up dates are shown in Table 3. In most cases the difference in error between the dates based on forecast and analysis was small. This is natural as the first guess of the LSWT analysis is the forecast LSWT by FLake. We will discuss next the freeze-up, then the break-up dates.

Table 3. Statistical measures of the error of ~~freezing~~ freeze-up and ~~melting~~ break-up date

		bias	~~sde~~ stde	max	min	N
unit		days	days	days	days	
~~Freezing~~ Freeze-up	LSWT an	-3.5	17.9	64	-52	233
	IceD fc	-0.3	17.8	67	-41	233
~~Melting~~ Break-up	LSWT an	-15.2	8.5	2	-54	288
	IceD fc	-20.5	9.2	-1	-56	258

Denotation: LSWT an - LID estimated from analysed LSWT, IceD fc - LID estimated from forecast ice thickness.

~~Statistics of the error in melting and freezing dates are shown in Table 3. 'LSWT an' refers to the melting/freezing dates computed from analysed lake surface temperature and~~ The bias in the error of freeze-up dates was small according to both 'IceD fc' ~~to those estimated from the forecast ice thickness. Over the 45 lakes included in this comparison, the number of cases of melting was 288 as estimated from the analyzed LSWT and 258 as estimated from the forecast ice thickness. The difference is due to starting time of our data. When the data started at the 1st of April 2012, at several stations the lake was already open according to FLake forecast while the analysed LSWT still indicated frozen conditions. For freezing, the number of cases was 233 according to both estimates. As the data contains the time period from the 1st April 2012 to the 30th June 2018, the maximum number of freezing events on an individual lake is six and that of melting events seven. In practice, the number may be less for some lakes because of missing observations~~and 'LSWT an', -0.3 and -3.5 days, respectively. The minimum and maximum errors were large in both cases: the maximum freeze-up date occurred about two months too late, the minimum about one and a half months too early. However, as will be shown later, the largest errors mostly occurred on a few problematic lakes while in most cases the errors were reasonable.

Figure 5a) shows the frequency distribution of the error of ~~freezing dates. Definition of the freezing date from the ice thickness by FLake gave slightly more occurrences~~ freeze-up dates. Forecast freeze-up dates occurred slightly more often in the unbiased class (error between -5 - +5 days), compared to the ~~estimate from the analysed LSWT~~estimated dates from the analysis. Of all cases 48 %~~and~~ / 40 % ~~fell in this class according to ice thickness and LSWT, respectively. In 16~~(percentages here and in the following are given as 'IceD fc' / 'LSWT an') fell into this class. In 20% / ~~20~~26% of cases the ~~freezing occurs~~ freeze-up occurred more than five days too late and only in ~~9~~11% / ~~11~~9% cases more than two weeks too late. ~~This class of more than two weeks too late freezing consists of~~ In case of 'IceD fc', the class of freeze-up more than 15 days too late comprised 25 cases ~~which are~~ distributed over 15 lakes, thus ~~in most cases one event~~ mostly one or two events per lake. This suggests that the error ~~is~~ was related more to individual years than to systematically problematic lakes. It is worth noting, that of the eight cases where the error ~~is~~ was over 45 days, ~~are all but one due to one lake,~~ six cases were due to a single lake, Lake Kevojärvi~~which~~ . This lake is situated in the very north of Finland. ~~This lake is~~ It is very small and narrow, with an area of

1 km~~$^{-2}$, and situated~~ ^{2}, and located in a steep canyon. Therefore it is poorly represented by the HIRLAM grid and ~~both FLake and analysis~~ the results seem unreliable.

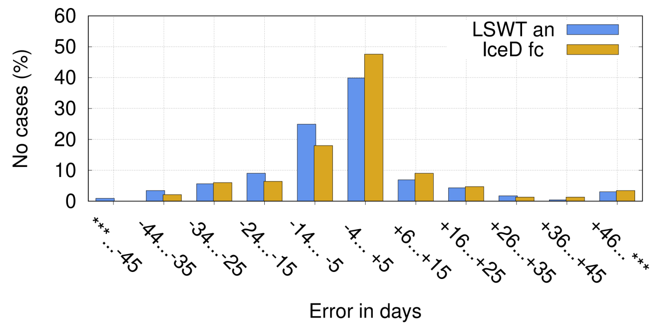
Concerning ~~the cases of~~ too early freezing, in ~~44~~33% / ~~32~~44% of the cases ~~freezing occurs~~ freeze-up occurred more than five days too early and in ~~19~~15% / ~~15~~19% more than two weeks too early. ~~The last mentioned~~ According to the forecast, these 15% (34 cases) ~~are~~ were distributed over 19 lakes. Each of the five large lakes Pielinen, Kallavesi, Haukivesi, Päijänne and Inari ~~occur~~ occurred in this category three times while all other lakes together ~~share~~ shared the remaining 19 cases during the six winters.

~~Looking at the errors in melting dates (Figure 5b) , both estimates indicate too earlymelting and~~ The break-up dates (Table 3) show a large negative bias, about two ('LSWT an') or three weeks ('IceD fc'), indicating that lake ice break-up was systematically forecast to occur too early. However, the standard deviation of the error was only about half of that of the error of freeze-up dates and there were no long tails in the distribution (Figure 5b). Hence the distribution is strongly skewed towards too early ~~dates. Based on the LSWT analysis, the~~ break-up, but much narrower than that of freeze-up (Figure 5a). The large bias was most probably due missing snow over lake ice in this HIRLAM version (see Section 5). The maximum frequency (~~52~~47 %) ~~occurs~~ was in the class ~~-14~~-24 - ~~-5 days while based on the ice thickness~~-15 days for 'IceD fc', while in case of 'LSWT an', the maximum frequency (~~47~~52 %) ~~is~~ occurred in the class ~~-24~~-14 - ~~-15 days. The mean values are -15.2 and -20.5 days and the standard deviations are 8.5 and 9.2 days, respectively. FLake suggests~~ -5 days. FLake forecast 'IceD fc' suggested only three cases in the unbiased class -4 - +5 while according to ~~the LSWT analysis there are~~ 'LSWT an' there were 12 cases in this class. Hence, the ~~melting~~ break-up dates derived from analysed ~~LWST correspond~~ LSWT corresponded the observations better than those derived from FLake ice thickness forecast~~but both are strongly biased towards too early melting. In the class where the error is over 35 days too early there are 19 cases on 12 different lakes. Four cases of these occured in the largest error category (over 45 days) on Lake Kilpisjärvi.~~

~~If we compare the error in freezing or melting dates based on analysed LSWT on those (27) lakes where SYKE temperature observations are available and used in the analysis to the rest (18) of lakes with no observations, it appears that the differences are small (not shown). Furthermore, similar differences appear also on the error estimates based on ice thickness from FLake. This suggests that the differences between these groups are related to the individual properties of the lakes: their depth, size, shape etc. rather than to the usage of LSWT observations in the analysis~~.

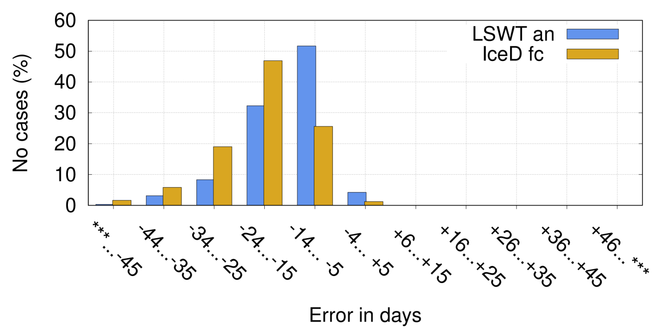
~~We can conclude that HIRLAM /FLake succeeds~~ Note that this kind of method of verifying LID compares two different types of data. The observations by SYKE are visual observations from the shore of the lake (see Section 3.2.2), while the freeze-up and break-up dates from HIRLAM are based on single-gridpoint values of LSWT or ice thickness (see Section 3.3.2). In addition, the resulting freeze-up and break-up dates from HIRLAM are somewhat sensitive to definition of the freezing and melting tresholds. Here we used 1 mm for the forecast ice thickness and the freezing point for the LSWT analysis as the critical values.

In conclusion, the validation statistics show that HIRLAM succeeded rather well in predicting ~~the~~ freezing of Finnish lakes. Almost in half of the cases the error ~~is~~ was less than ± 5 days. Some bias towards too early ~~freezing~~ freeze-up can be seen ~~. Melting is~~ both in forecast and in the analysis. Melting was more difficult. FLake ~~predicts melting~~ predicted lake ice break-up



Frequency distribution of the difference between analysed/forecast and observed freezing days over all lakes 2012-2018. Variables used in diagnosis of ice existence: analysed LSWT crossing the freezing point (blue) and forecast ice thickness > 1 mm (orange). Observed variable: freezing date by SYKE. x-axis: difference (fc-ob), unit day, y-axis: percentage of all cases.

(a) error of freeze-up dates



As for Figure 5a but for melting days.

(b) error of break-up dates

Figure 5. Frequency distribution of the difference between analysed/forecast and observed freeze-up and break-up dates over all lakes 2012-2018. Variables used in diagnosis of ice existence: analysed LSWT crossing the freezing point (blue) and forecast ice thickness > 1 mm (magenta). Observed variable: freeze-up date by SYKE. x-axis: difference (fc-ob), unit day, y-axis: percentage of all cases.

always too early, with a mean error of over two weeks, and the LSWT analysis mostly follows it. The statistics suggest that only on a few stations the freezing or melting dates were systematically wrong during most of the years. Instead, most of the large errors were distributed among many lakes. The result of the freezing or melting dates diagnostics is somewhat sensitive to how the tresholds for freezing and melting are set. Here we used 1 mm for ice thickness and the freezing point for the LSWT analysis as the critical values. analysis mostly followed it.

4.3 Comparisons on three lakes

In this section we ~~combine the analysis of LWST time-series and LID~~ present LSWT and LID time-series for two representative lakes, Kilpisjärvi in the north and Lappajärvi in the west (see the map in Figure 2). Observed and forecast ice and snow thickness are discussed, using also additional data from Lake Simpelejärvi in ~~the south-east of~~ southeastern Finland.

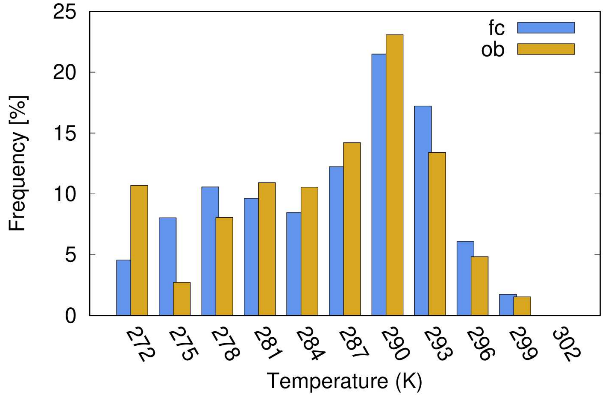
Lake Kilpisjärvi is an Arctic lake at the elevation of 473 m, surrounded by fells. ~~Its~~ The lake occupies 40 % of the area of HIRLAM gridsquare covering it (the mean elevation of the gridsquare is 614 m). The average/maximum ~~depth is 22.5~~depths of the lake are 19.5/57 m and the surface area ~~37.33~~is 37.3 km^2. The heat balance as well as the ice and snow conditions on Lake Kilpisjärvi have been ~~a subject of~~ subject to several studies (Leppäranta et al., 2012; Lei et al., 2012; Yang et al., 2013). Typically, the ice season lasts there seven months from November to May. Lake Lappajärvi is formed from a 23 km wide meteorite impact crater, which is estimated to be 76 million years old. It is Europe's largest crater lake with a surface area of 145.5 km^2 and an average/maximum depth of ~~12~~6.9/36 m. Here the climatological ice season is shorter, typically about five months from December to April. The average/maximum depth of Lake Simpelejärvi is ~~9.3~~8.7/34.4 m and the surface area 88.2 km^2. This lake is located at the border between Finland and Russia and belongs to the catchment area of Europe's largest lake, Lake Ladoga in Russia.

~~Figures 6a - 7b~~

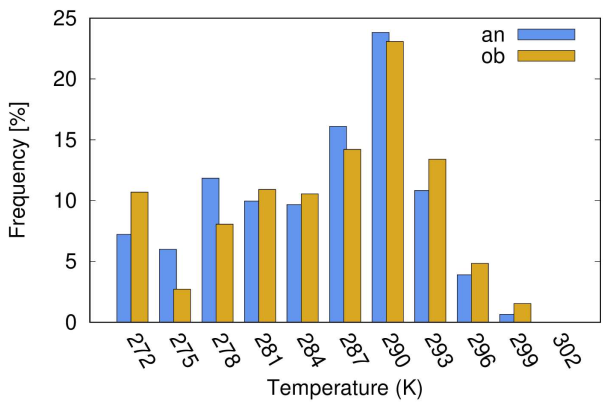
Figures 6 and 7 show the frequency distributions of LSWT according to ~~the observations~~ forecast v.s. ~~forecast and analysis for these lakes. Features common to the majority of~~ observation and analysis v.s. observation for Lappajärvi and Kilpisjärvi. Features similar to the results averaged over all lakes (Section 4.1, Figures 3 and 4) are seen, i.e. underestimation of the amount of cold temperature cases and overestimation of the warmer temperatures by the forecast and analysis. On Lake Lappajärvi, only the amount of below-freezing temperatures ~~is~~ was clearly underestimated, otherwise the distributions look quite balanced. According to the observations, on Lake Kilpisjärvi ~~the days with frozen surface dominate during the April–November periods~~ ice-covered days dominated during the periods from November to May. According to both ~~FLake forecast and HIRLAM LSWT analysis~~ LSWT analysis and forecast the amount of these days ~~is clearly smaller~~ was clearly smaller in HIRLAM.

Yearly time series of the observed, forecast and analysed LSWT, with the observed LID marked, are shown in Figures 8 and 9. In the absence of observations, the HIRLAM analysis ~~follows~~ followed the forecast. Missing data in the time series close to ~~freezing and melting~~ freeze-up and break-up are due to missing observations, hence missing information in the feedback files (see Section 2.2). Differences between the years due to the different prevailing weather conditions ~~can be~~ are seen in the temperature variations.

Generally, ~~in spring FLake tends~~ FLake tended to melt the lakes too early in spring, as already indicated by the LID statistics (Section 4.2). The too early ~~melting~~ break-up and too warm LSWT in summer show up clearly in Kilpisjärvi (Figure 9). In Lappajärvi, the model and analysis ~~are~~ were able to follow even quite large and quick variations of LSWT in summer, but ~~tend~~ tended to somewhat overestimate the maximum temperatures. Overestimation of the maximum temperatures by FLake ~~is~~ was still more prominent in shallow lakes (not shown). In autumn over Lakes Lappajärvi and Kilpisjärvi, the forecasts and analyses



(a) forecast v.s. observation



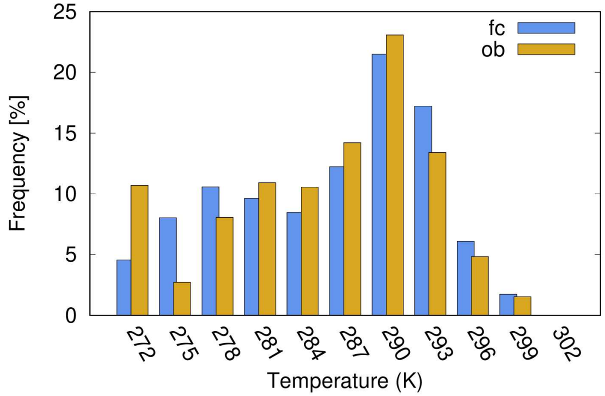
(b) analysis v.s. observation

Figure 6. Frequency of observed (yellow) and forecast or analysed (blue) LSWT over Lake Lappajärvi 2012-2018, all temperatures included. x-axis: LSWT, unit K, y-axis: frequency, unit %.

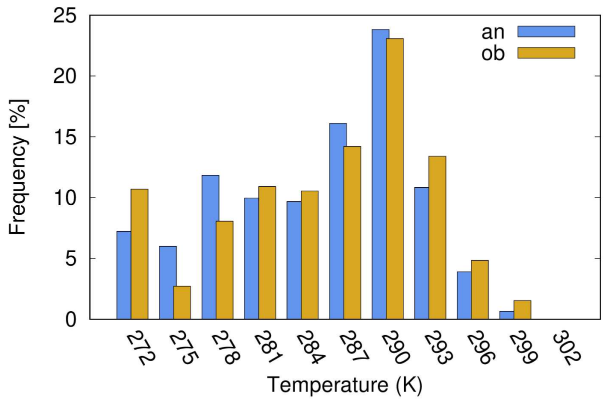
~~follow~~ followed closely the LSWT observations and ~~reproduce the freezing date~~ reproduced the freeze-up dates within a few days, which ~~is~~ was also typical to the majority of lakes.

Figure 10 shows a comparison of forecast and observed evolution of ice ~~and snow thickness~~ thickness and snow depth on Lappajärvi, Kilpisjärvi and Simpelejärvi in winter 2012-2013, typical also for the other lakes and years studied. ~~In~~ The most striking feature is that there was no snow in the HIRLAM forecast.

On all three lakes, the ice thickness ~~starts~~ started to grow after ~~freezing~~ freeze-up both according to the forecast and the observations. In the beginning HIRLAM ~~/FLake ice grows~~ ice grew faster than observed. However, according to the forecast ice thickness ~~starts~~ started to decrease in March of every year but according to the observations only a month or two later. ~~The most remarkable feature is that there is no snow in the FLake forecast. It was found that this was due to a coding error in the HIRLAM reference version 7.4 which is applied operationally in FMI.~~



(a) ~~As for Figure 3a) but for Lake Lappajärvi, with all temperatures included~~forecast v.s. observation
~~As for Figure 4a) but for Lake Lappajärvi~~
~~As for Figure 6a but for Lake Kilpisjärvi.~~



(b) ~~As for Figure 6b but for Lake Kilpisjärvi~~analysis v.s. observation

Figure 7. As for Figure 6 but for Lake Kilpisjärvi.

The too early ~~melting of~~ break-up of lake ice in the absence of snow could be explained by the wrong absorption of the solar energy in the model. In reality, the main factor of snow and ice melt in spring is the increase of daily solar radiation. In HIRLAM, the downwelling short-wave irradiance at the surface is known to be reasonable, with some overestimation of the largest clear-sky fluxes and all cloudy fluxes (Rontu et al., 2017). Over lakes, HIRLAM ~~/FLake~~ uses constant values for the snow and ice shortwave reflection, with albedo values of 0.75 and 0.5, correspondingly. When there was no snow, the lake surface was thus assumed too dark. 25 % more absorption of an assumed maximum solar irradiance of 500 Wm^{-2} (valid for the latitude of Lappajärvi in the end of March) would mean availability of extra 125 Wm~~-2~~ $^{-2}$ for melting of the ice, which corresponds the magnitude of increase of available maximum solar energy within a month at the same latitude.

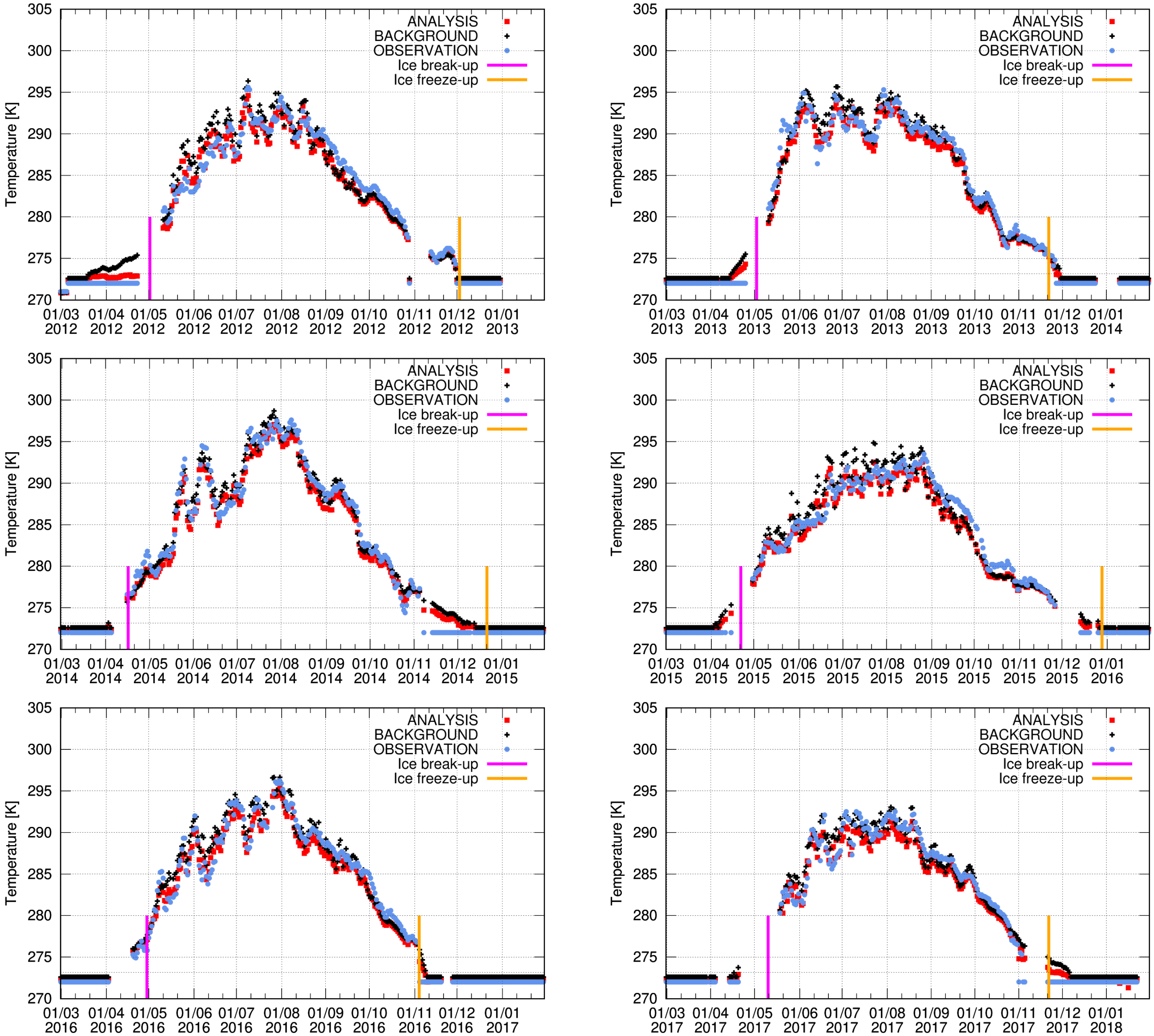


Figure 8. Time-series of the observed, analysed and forecast LSWT at the Lappajärvi observation location 23.67 E, 63.15 N for the years 2012-2018 based on 06 UTC data. Markers are shown in the inserted legend. Observed ~~freezing~~ freeze-up date (blue) and ~~melting~~ break-up date (red) are marked with vertical lines.

The forecast of too thick ice can also be explained by the absence of snow in the model. When there is no insulation by the snow layer, the longwave cooling of the ice surface in clear-sky conditions is more intensive and leads to faster growth of ice compared to the situation of snow-covered ice. In nature, ice growth can also be due to the snow transformation, a process whose parametrization in the models is demanding (Yang et al., 2013; Cheng et al., 2014).

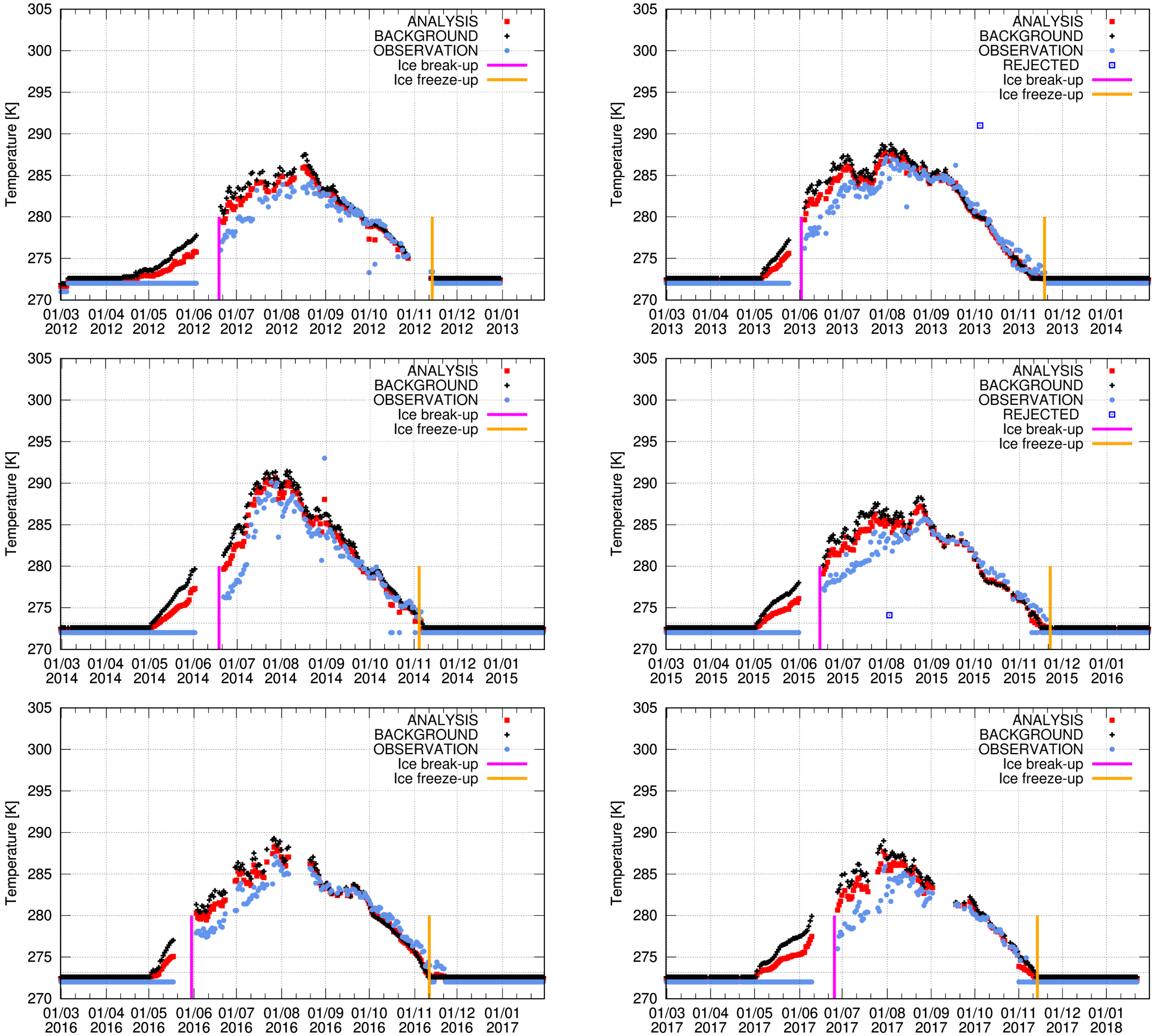


Figure 9. As for Figure 8 but for lake Kilpisjärvi, 20.82 E, 69.01 N.

Also the downwelling longwave radiation plays a role in the surface energy balance. We may expect values from 150 Wm^{-2} to 400 Wm^{-2} in the Nordic spring conditions, with the largest values related to cloudy and the smallest to clear-sky situations. The standard deviation of the predicted by HIRLAM downwelling longwave radiation fluxes has been shown to be of the order of 20 Wm^{-2}, with a positive systematic error of a few Wm^{-2} (Rontu et al., 2017). Compared to the systematic effects related to absorption of the solar radiation, the impact of the longwave radiation variations on lake ice evolution is presumably small.

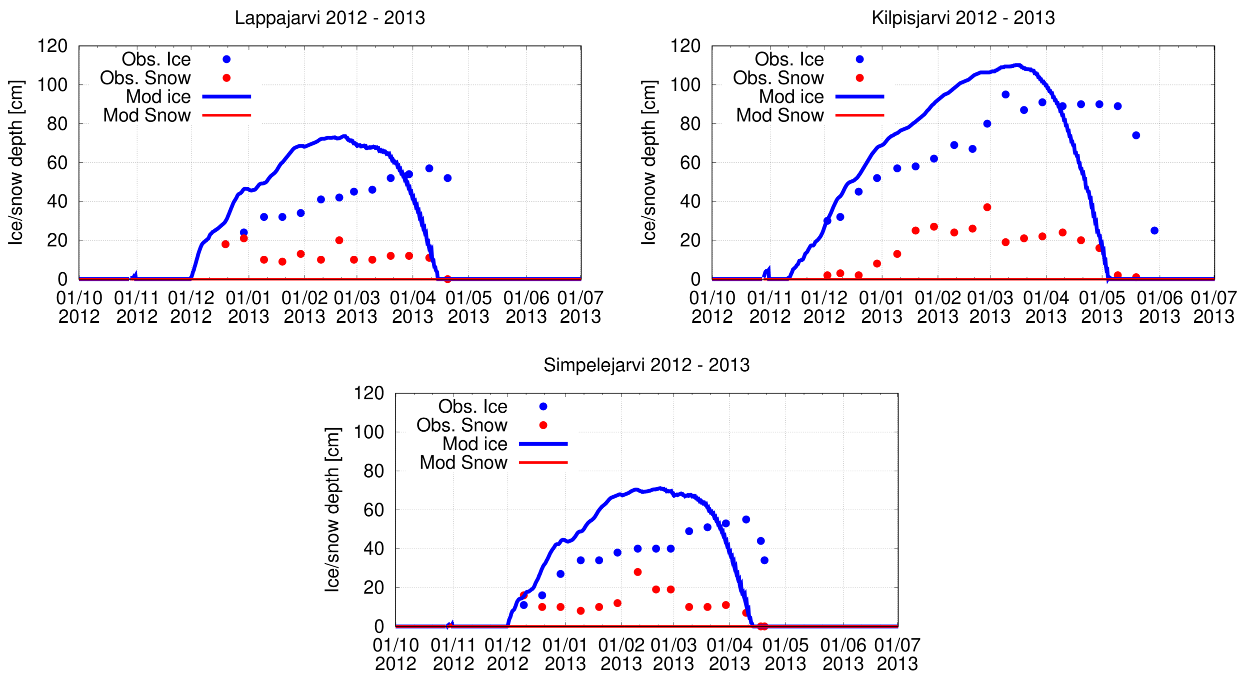


Figure 10. Evolution of ice (blue) and snow (red) thickness at Lakes Lappajärvi, Kilpisjärvi and Simpelejärvi during winter 2012-2013.

5 Discussion: snow on lake ice

The most striking result reported in Section 4 was the too early melting of the lake ice predicted by FLake in HIRLAM as compared to observations. We suggested that the early break-up is related to the missing snow on lake ice in HIRLAM. It was detected that a too large critical value to diagnose snow existence prevented practically all accumulation of the forecast snowfall on lake ice in the reference HIRLAM v.7.4, used operationally at FMI.

In general, handling of the snow cover on lake and sea ice is a demanding task for the NWP models. In HIRLAM, snow depth observations are included into the objective analysis over the land areas, but not over ice where no observations are widely available in real time. Snow depth and temperature over land are treated prognostically using dedicated parametrizations (in HIRLAM, similar to Samuelsson et al., 2006, 2011, see also Boone et al., 2017). Over the sea, a simple prognostic parametrization of sea ice temperature is applied in HIRLAM but neither the thickness of ice nor the depth or temperature of snow on ice are included (Samuelsson et al., 2006). Batrak et al. (2018) provide a useful review and references concerning prognostic sea ice schemes and their snow treatment in NWP models. An essential difference between the simple sea ice scheme and the lake ice scheme applied in HIRLAM is that the former relies on external data on the existence of sea ice cover, provided by the objective analysis, while the latter includes prognostic treatment of the lake water body also. This means that the lake ice freezes and melts in the model depending on the thermal conditions of lake water, evolving throughout the seasons.

The ice thickness, snow depth and ice and snow temperatures are prognostic variables of FLake. When the FLake parametrizations were introduced into HIRLAM (Kourzeneva et al., 2008; Eerola et al., 2010), parametrization of the snow thickness and snow temperature was first excluded. In the COSMO NWP model, snow is implicitly accounted for by modifying ice albedo using empirical data on its temperature dependence (Mironov et al., 2010). This way was applied also e.g. in a recent study over the Great Lakes (Baijnath-Rodino and Duguay, 2019).

Semmler et al. (2012) performed a detailed winter-time comparison between FLake and a more complex snow and ice thermodynamic model (HIGHTSI) on a small lake in Alaska. FLake includes only one ice and one soil layer, while HIGHTSI represents a more advanced multilayer scheme. Atmospheric forcing for the stand-alone experiments was provided by HIRLAM. Based on their sensitivity studies, Semmler et al. (2012) suggested three simplifications to the original, time-dependent snow-on-ice parametrizations of FLake: use a prescribed constant snow density, modify the value of the prescribed molecular heat conductivity and use prescribed constant albedos of dry snow and ice. Later, a similar comparison was performed over Lake Kilpisjärvi (Yang et al., 2013), confirming the improvements due to the updated snow parametrizations in FLake. Implementation of these modifications allowed to include the parametrization of snow on lake ice also into HIRLAM (Section 2.1).

In FLake, snow on lake ice is accumulated from the predicted snowfall. Snow melt on lake ice is related to snow and ice temperatures. In case of FLake integrated into HIRLAM, accumulation and melt are updated at every time step of the advancing forecast. Very small amounts of snow are considered to fall beyond the accuracy of parametrizations and removed. This is controlled by a critical limit, which was set too large (one millimeter instead of ten micrometers) in HIRLAM v.7.4. Due to the incorrect critical value, practically no snow accumulated on lake ice in the FMI operational HIRLAM, validated in this study. In a HIRLAM test experiment, where the original smaller value was used, up to 17 cm of snow accumulated on lake ice within a month (Janurary 2012, not shown).

6 Conclusions and outlook

In this study, *in-situ* lake observations from the Finnish Environment Institute were used for validation of the HIRLAM NWP model, which is applied operationally in the Finnish Meteorological Institute. ~~It~~ HIRLAM contains Freshwater Lake prognostic parametrizations and an independent objective analysis of lake surface state. We focused on comparison of observed and forecast lake surface water temperature, ice ~~and snow thickness~~ thickness and snow depth in the years 2012 - 2018. Because the HIRLAM ~~/FLake~~ system was unmodified during this period, a long uniform dataset was available for evaluation of the performance of FLake integrated ~~in~~ into an operational NWP model. On the other hand, no conclusions about the impact of the lake surface state on the operational forecast of the near-surface temperatures, cloudiness or precipitation can be drawn because of the lack of alternative (without FLake) forecasts for comparison.

On average, the forecast and analysed LSWT were warmer than observed with systematic errors of 0.91 K and 0.35 K, correspondingly. The mean absolute errors were 1.94 and 1.32 K. Thus, the independent observation-based analysis of *in-situ* LSWT observations was able to improve the FLake +6 h forecast used as the first guess. However, the resulting analysis is by definition not used for correction of the FLake forecast but remains an independent by-product of HIRLAM.

An overestimation of the FLake LSWT summer maxima was found, especially for the shallow lakes. This behaviour of FLake is well known, documented earlier e.g. by Kourzeneva (2014). It arises due to the difficulty to handle correctly the mixing in the near-surface water layer that is intensively heated by the sun.

Forecast ~~freezing~~ freeze-up dates were found to correspond the observations well, typically within a week. The forecast ice thickness tended to be overestimated, still the ~~melting~~ break-up dates over most of the lakes occured systematically several weeks too early. Practically no forecast snow was on found on the lake ice, although the snow parametrization by FLake was included in HIRLAM. The reason for the ~~wrong behaviour in HIRLAM~~ incorrect behaviour was evidently related to a ~~coding error that prevented snow accumulation~~ too large critical value to diagnose snow existence that prevented the accumulation of snow on lake ice. The too early melting and overestimated ice thickness differ from the results by Pietikäinen et al. (2018); Yang et al. (2013); Kourzeneva (2014), who reported somewhat too late melting of the Finnish lakes when FLake with realistic snow parametrizations was applied within a climate model or ~~independently,~~ stand-alone driven by NWP data. It can be concluded that a ~~realiste~~ realistic parametrization of snow on lake ice is ~~importart~~ important in order to describe correctly the lake surface state in spring.

Small lakes and those of complicated geometry cause problems for the relatively coarse HIRLAM grid of 7 - kilometre resolution. The problems are related to the observation usage, forecast and validation, especially when interpolation and selection of point values are applied. The observations and model represent different spatial scales. For example, the comparison of the ~~freezing and melting~~ freeze-up and break-up dates was based on diagnostics of single-gridpoint values that were compared to observations ~~representing~~ which represent entire lakes as ~~seen~~ overseen from the observation sites. Also the results of LID diagnostics were sensitive to the criteria for definition of the ice existence in HIRLAM~~/FLake~~. All this adds unavoidable inaccuracy into the model-observation intercomparison but does not change the main conclusions of the present study.

SYKE LSWT observations used for the real-time analysis are regular and reliable but ~~did~~ do not always cover the days immediately after ~~melting~~ break-up or close to ~~freezing~~freeze-up, partly because the quality control of HIRLAM LSWT analysis utilizes the SYKE statistical lake water temperature model results in a ~~too~~ strict way. Although the 27 observations are located all over the country, they cover a very small part of the lakes and their availability is limited to Finland. SYKE observations of the ice and snow depth as well as the ~~freezing and melting~~ freeze-up and break-up dates provide valuable data for the validation purposes.

A need for minor technical corrections in the FMI HIRLAM ~~/FLake~~ system was revealed. The ~~snow accumulation bug was corrected in October 2018,~~ coefficient influencing snow accumulation on lake ice was corrected based on our findings. Further developments and modifications are not foreseen because the HIRLAM NWP systems, applied in the European weather services, are being replaced by kilometre-scale ~~HARMONIE-AROME-based operational systems (Bengtsson et al., 2017)~~ ALADIN-HIRLAM forecasting systems (Termonia et al., 2018; Bengtsson et al., 2017), where the prognostic FLake parametrizations are also available. HARMONIE/FLake uses the newest version of the global lake database (GLDB v.3) and contains updated snow and ice properties~~that were suggested by (Yang et al., 2013)~~. The objective analysis of lake surface state is yet to be implemented~~into HARMONIE-AROME~~, taking into account the HIRLAM experience summarized in this study and earlier by Kheyrollah Pour et al. (2017). In the future, an important source of wider observational information on lake surface state are the satellite measurements, whose operational application in NWP models still requires further work.

Prognostic and diagnostic lake variables within HIRLAM variable unit type temperature of snow on lake ice K prog by FLake temperature of lake ice K prog by FLakemean water temperature K prog by FLakemixed layer temperature K prog by FLakebottom temperature K prog by FLaketemperature of upper layer sediments K prog by FLakemixed layer depth m prog by FLakethickness of upper layer sediments m prog by FLakethermocline shape factor - prog by FLakelake ice thickness m prog by FLakesnow depth on lake ice m prog by FLakeLSWT K diag by FLake= mixed layer temperature if no icelake surface temperature K diag by FLake uppermost temperature: LSWT or ice or snowLSWT K anal by HIRLAMflag value 272 K when there is icelake surface roughness m diag by HIRLAMscreen level temperature over lake m diag by HIRLAMscreen level abs.humidity over lake m diag by HIRLAManemometer level u-component over lake m diag by HIRLAManemometer level v-component over lake m diag by HIRLAMlatent heat flux over lake Wm^{-2} diag by HIRLAMsensible heat flux over lake Wm^{-2} diag by HIRLAMscalar momentum flux over lake Wm^{-2} diag by HIRLAMSW net radiation over lake Wm^{-2} diag by HIRLAMLW net radiation over lake Wm^{-2} diag by HIRLAMdepth of lake m pres in HIRLAM gridfraction of lake 0-1pres in HIRLAM gridfraction of lake ice 0-1diag in HIRLAM grid

Lakes with SYKE observations used in this study NAME LON LAT MEAND HIRD HIRFR HIRID Pielinen 29.607 63.271 11.1 10.0 0.916 4001 Kallavesi 27.783 62.762 12.1 10.0 0.814 4002 Haukivesi 28.389 62.108 9.0 10.0 0.725 4003 Saimaa 28.116 61.338 17.0 10.0 0.950 4004 Pääjärvi1 24.789 62.864 3.9 3.0 0.430 4005 Nilakka 26.527 63.115 4.9 10.0 0.866 4006 Konnevesi 26.605 62.633 15.9 10.0 0.937 4007 Jääsjärvi 26.135 61.631 4.6 10.0 0.750 4008 Päijänne 25.482 61.614 14.1 10.0 0.983 4009 Ala-Rieveli 26.172 61.303 11.3 10.0 0.549 4010 Kyyvesi 27.080 61.999 4.4 10.0 0.810 4011 Tuusulanjärvi 25.054 60.441 3.2 3.0 0.174 4012 Pyhäjärvi 22.291 61.001 5.5 5.0 0.922 4013 Längelmävesi 24.370 61.535 6.8 10.0 0.875 4014 Pääjärvi2 25.132 61.064 14.8 14.0 0.350 4015 Vaskivesi 23.764 62.142 7.0 10.0 0.349 4016 Kuivajärvi 23.860 60.786 2.2 10.0 0.419 4017 Näsijärvi 23.750 61.632 14.1 10.0 0.850 4018 Lappajärvi 23.671 63.148 12.0 10.0 1.000 4019 Pesiöjärvi 28.650 64.945 7.3 7.0 0.290 4020 Rehja-Nuasjärvi 28.016 64.184 8.5 10.0 0.534 4021 Oulujärvi 26.965 64.451 7.6 10.0 1.000 4022 Ounasjärvi 23.602 68.377 6.6 10.0 0.166 4023 Unari 25.711 67.172 6.1 10.0 0.491 4024 Kilpisjärvi 20.816 69.007 22.5 22.0 0.399 4025 Kevojärvi 27.011 69.754 7.0 10.0 0.016 4026 Inarijärvi 27.924 69.082 14.4 14.0 0.979 4027 Simpelejärvi 29.482 61.601 9.3 10.0 0.548 40241 Pokkaanlahti 27.264 61.501 7.00 10.0 0.299 40261 Muurasjärvi 25.353 63.478 9.10 10.0 0.060 40263 Kalmarinjärvi 25.001 62.786 5.80 5.0 0.330 40271 Summasjärvi 25.344 62.677 6.70 10.0 0.555 40272 Iisvesi 27.021 62.679 17.2 18.0 0.456 40277 Hankavesi 26.826 62.614 7.00 18.0 0.100 40278 Petajävesi 25.173 62.255 2.70 3.0 0.245 40282 Kukkia 24.618 61.329 6.00 10.0 0.299 40308 Ähtärinjärvi 24.045 62.755 7.00 10.0 0.266 40313 Kuortaneenjärvi 23.407 62.863 7.00 10.0 0.277 40328 Lestijärvi 24.716 63.584 7.00 10.0 0.513 40330 Pyhäjärvi 25.995 63.682 7.00 10.0 0.266 40331 Lentua 29.690 64.204 7.60 7.0 0.600 40335 Lammasjärvi 29.551 64.131 4.40 3.0 0.200 40336 Naamankajärvi 28.246 65.104 7.00 7.0 0.299 40342 Korvuanjärvi 28.663 65.348 18.50 10.0 0.342 40343 Oijärvi 25.930 65.621 7.00 10.0 0.333 40345

Code and data availability. Observational data was obtained from SYKE open data archive SYKE, 2018 as follows: 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/. A supplementary file containing the freeze-up and break-up dates as picked and prepared for the lakes studied here is attached. Data picked from HIRLAM archive are attached as supplementary files: data from the objective analysis feedback files (observed, analysed, forecast LSWT interpolated to the 27 active station locations) and from the gridded output of the HIRLAM analysis (analysed LSWT, forecast ice and snow thickness from the nearest gridpoint of all locations used in the present study).

In this study, FMI operational weather forecasts resulting from use of HIRLAM v.7.4 (rc1, with local updates) were validated against lake observations. The HIRLAM reference code is not open software but the property of the international HIRLAM-C programme. For research purposes, the codes can be requested from the programme (hirlam.org). The source codes of the version operational at FMI, relevant for the present study, are available from the authors upon request.

Author contributions. Laura Rontu computed the LSWT statistics based on HIRLAM feedback files. Kalle Eerola performed the freeze-up and break-up date, snow and ice thickness comparisons based on data picked from HIRLAM grib files. Matti Horttanainen prepared observation data obtained via SYKE open data interface and lake depths from GLDB v.3. Laura Rontu composed the manuscript text based on input from all authors.

Competing interests. No competing interests are present.

Acknowledgements. Our thanks are due to Joni-Pekka Pietikäinen and Ekaterina Kourzeneva for discussions and information, to Margarita Choulga and Olga Toptunova for the support with the ~~GLDB lake depth data~~updated GLDB v.3 data, and to Emily Gleeson for advice with English language. The comments of three anonymous reviewers and the editor significantly helped to improve the presentation of our results.

References

Baijnath-Rodino, J. and Duguay, C.: Assessment of coupled CRCM5–FLake on the reproduction of wintertime lake-induced precipitation in the Great Lakes Basin. Theoretical and Applied Climatology https://doi.org/10.1007/s00704-019-02799-8, 2019.

Balsamo, G., ~~R. Salgado~~Salgado, R., Dutra, E.~~Dutra~~, Boussetta, S.~~Boussetta,~~ Stockdale, T. ~~Stockdale and M. Potes~~and Potes, M: On the contribution of lakes in predicting near-surface temperature in a global weather forecasting model. Tellus A, 64, 15829, https://doi.org/10.3402/tellusa.v64i0.15829, 2012.

Batrak, Y., Kourzeneva, E. and Homleid, M.: Implementation of a simple thermodynamic sea ice scheme, SICE version 1.0-38h1, within the ALADIN–HIRLAM numerical weather prediction system version 38h1. Geosci. Model Dev., 11, 3347–3368, https://doi.org/10.5194/gmd-11-3347-2018, 2018.

Bengtsson, L., ~~U. Andrae, T.Aspelien~~Andrae, U., Aspelien, T., Batrak, Y.~~Batrak,~~ , Calvo, J.~~Calvo~~, ~~W.~~de Rooy, W., Gleeson, E.~~Gleeson~~, ~~B.~~Hansen-Sass, ~~M.~~B., Homleid, M.~~Hortal,~~ , Hortal, M., Ivarsson, K.~~Ivarsson,~~ , Lenderink, G.~~Lenderink~~, ~~S.Niemelä, K.~~Niemelä, S., Pagh Nielsen, K., Onvlee, J.~~Onvlee,~~ , Rontu, L.~~Rontu,~~ , Samuelsson, P.~~Samuelsson~~, ~~D.~~Santos Muñoz, D., Subias, A.~~Subias,~~ , Tijm, S.~~Tijm,~~ , Toll, V.~~Toll,~~ , Yang, X. ~~Yang, and M.~~ and Ødegaard Køltzow, M.: The HARMONIE-AROME model configuration in the ALADIN-HIRLAM NWP system. Monthly Wea. Rev., 145, 1919–1935, https://doi.org/10.1175/MWR-D-16-0417.1, 2018.

. Boone, A., Samuelsson, P., Gollvik, S., Napoly, A., Jarlan, L., Brun, E. and Decharme, B.: The interactions between soil–biosphere–atmosphere land surface model with a multi-energy balance (ISBA-MEB) option in SURFEXv8 – Part 1: Model description. Geosci. Model Dev., 10, 843–872. https://doi.org/10.5194/gmd-10-843-2017, 2017.

Cheng, B, Vihma. T, Rontu, L., Kontu A., Kheyrollah Pour, H., Duguay, C. and Pulliainen, J.: Evolution of snow and ice temperature, thickness and energy balance in Lake Orajärvi, northern Finland. Tellus A: Dynamic Meteorology and Oceanography ,66, 21564, https://doi.org/10.3402/tellusa.v66.21564, 2014.

Choulga, M., Kourzeneva, E., Zakharova, E. and Doganovsky, A.: Estimation of the mean depth of boreal lakes for use in numerical weather prediction and climate modelling. Tellus A, 66, 21295, https://doi.org/10.3402/tellusa.v66.21295, 2014.

Cordeira, J. M. and Laird, N. F. 2008. The influence of ice cover on two lake-effect snow events over lake Erie. *Mon. Wea. Rev.* **136**, 2747–2763. https://doi.org/DOI: 10.1175/2007MWR2310.1.

Elo, A-R.: Effects of climate and morphology on temperature conditions of lakes. University of Helsinki, Division of Geophysics, Report series in Geophysics. Available at http://urn.fi/URN:ISBN:978-952-10-3745-0, 2007.

Eerola, K., Rontu, L., Kourzeneva, E. and Shcherbak, E.: A study on effects of lake temperature and ice cover in HIRLAM. Boreal Env. Res., 15, 130–142, 2010.

Eerola, K.: Twenty-one years of verification from the HIRLAM NWP system. Wea. Forecasting, 28, 270–285. https://doi.org/10.1175/WAF-D-12-00068.1, 2013.

Eerola, K., Rontu, L., Kourzeneva, E., Kheyrollah Pour, H. and Duguay, C.: Impact of partly ice-free Lake Ladoga on temperature and cloudiness in an anticyclonic winter situation-a case study using a limited area model. Tellus A, 66, 23929. https://doi.org/10.3402/tellusa.v66.23929, 2014.

Gandin, L.: Objective analysis of meteorological fields. Gidrometizdat, Leningrad. Translated from Russian, Jerusalem, Israel Program for Scientific Translations, 1965.

Kheyrollah Pour, H., Rontu, L., Duguay, C. R., Eerola, K. and Kourzeneva, E.: Impact of satellite-based lake surface observations on the initial state of HIRLAM. Part II: Analysis of lake surface temperature and ice cover. Tellus A., 66, 21395. https://doi.org/10.3402/tellusa.v66.21395, 2014.

Kheyrollah Pour, H., Choulga, M., Eerola, K., Kourzeneva, E., Rontu, L., Pan F. and Duguay C. R.: Towards improved objective analysis of lake surface water temperature in a NWP model: preliminary assessment of statistical properties. Tellus A., 66, 21534. https://doi.org/10.1080/16000870.2017.1313025, 2017.

Korhonen J.: Suomen vesistöjen jääolot. Suomen ympäristö, (751), 145 pp(in Finnish). Available at , 2005.

Korhonen J.: Suomen vesistöjen lämpötilaolot 1900-luvulla. Suomen ympäristö, (566), 116 pp (in Finnish)Long-term changes and variability of the winter and spring season hydrological regime in Finland. Report series in Geophysics, No. 79, 83 pp. University of Helsinki, Faculty of science, Institute for atmospheric and earth system research. Available at, 2002. : https://helda.helsinki.fi/bitstream/handle/10138/298308/longterm.pdf, 2019.

Kourzeneva, E., Samuelsson, P., Ganbat, G. and Mironov, D.: Implementation of lake model Flake into HIRLAM. HIRLAM Newsletter, (54), 54–64. Available at http://hirlam.org/, 2008.

Kourzeneva, E., Asensio, H., Martin, E. and Faroux, S.: Global gridded dataset of lake coverage and lake depth for use in numerical weather prediction and climate modelling. Tellus A., 64,15640. https://doi.org/10.3402/tellusa.v64i0.15640, 2012a.

Kourzeneva, E., Martin, E., Batrak, Y. and Moigne, P. L.:. Climate data for parameterisation of lakes in numerical weather prediction models. Tellus A., 64, 17226. https://doi.org/10.3402/tellusa.v64i0.17226, 2012b.

Kourzeneva, E.: Assimilation of lake water surface temperature observations with Extended Kalman filter. Tellus A., 66, 21510. https://doi.org/10.3402/tellusa.v66.21510, 2014.

Laird, N. F., Kristovich, D. A. R. and Walsh, J. E.: Idealized model simulations examining the mesoscale structure of winter lake-effect circulations. *Mon. Wea. Rev.* **131**, 206–221. https://doi.org/DOI: 10.1175/1520-0493(2003)131<0206:IMSETM>2.0.CO;2, 2003.

Leppäranta, M., Lindgren, E. and Shirasawa, K.: The heat budget of Lake Kilpisjarvi in the Arctic tundra. Hydrology research +research, 48, 969-980. https://doi.org/10.2166/nh.2016.171, 2012.

Lei, R., Leppäranta, M., Cheng, B., Heil P. and Li, Z.: Changes in ice-season characteristics of a European Arctic lake from 1964 to 2008. Climatic Change, 115, 725–739. https://doi.org/10.1007/s10584-012-0489-2, 2012.

Lindstedt, D., P. Lind, E. Kjellström and C. Jones.: A new regional climate model operating at the meso-gamma scale: performance over Europe. Tellus A, 67, 24138. https://doi.org/10.3402/tellusa.v67.24138, 2015.

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, J. and co authors: Development of a global land cover characteristics database and IGBP DISCover from 1-km AVHRR data. Int. J. Rem. Sens., 21, 1303–1130. Data and documentation available online at: http://edc2.usgs.gov/glcc/glcc.php, 2000.

Masson, V., P. Le Moigne, E. Martin, S. Faroux, S., and coauthorsAlias, A., Alkama, R., Belamari, S., Barbu, A., Boone, A., Bouyssel, F., Brousseau, P., Brun, E., Calvet, J.-C., Carrer, D., Decharme, B., Delire, C., Donier, S., Essaouini, K., Gibelin, A.-L., Giordani, H., Habets, F., Jidane, M., Kerdraon, G., Kourzeneva, E., Lafaysse, M., Lafont, S., Lebeaupin Brossier, C., Lemonsu, A., Mahfouf, J.-F., Marguinaud, P., Mokhtari, M., Morin, S., Pigeon, G., Salgado, R., Seity, Y., Taillefer, F., Tanguy, G., Tulet, P., Vincendon, B., Vionnet, V., and Voldoire, A.: The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes, Geosci. Model Dev., 6, 929–960.https://doi.org/:10.5194/gmd-6-929-2013, 2013.

Mironov, D., Heise, E., Kourzeneva, E., Ritter, B., Schneider, N. and Terzhevik, A.: Implementation of the lake parameterisation scheme FLake into the numerical weather prediction model COSMO. Boreal Env. Res., 15, 218–230, 2010.

Pietikäinen, J.-P., Markkanen, T., Sieck, K., Jacob, D., Korhonen, J., Räisänen, P., Gao, Y., Ahola, J., Korhonen, H., Laaksonen, A., and Kaurola, J.: The regional climate model REMO (v2015) coupled with the 1-D freshwater lake model FLake (v1): Fenno-Scandinavian climate and lakes. Geosci. Model Dev., 11, 1321–1342, https://doi.org/10.5194/gmd-11-1321-2018, 2018.

Rontu, L., Eerola, K., Kourzeneva, E. and Vehviläinen, B.: Data assimilation and parametrisation of lakes in HIRLAM. Tellus A., 64, 17611. https://doi.org/10.3402/tellusa.v64i0.17611, 2012.

Rontu L., Gleeson, E., Räisänen, P., Nielsen, K.P., Savijärvi, H. and Sass, B.H.: The HIRLAM fast radiation scheme for mesoscale numerical weather prediction models. Adv.Sci.Res., 14, 195–215, https://doi.org/10.5194/asr-14-195-2017, 2017.

Rooney, G. G. and Bornemann, F. J.: The performance of FLake in the Met Office Unified Model. Tellus A, 65, 21363, https://doi.org/10.3402/tellusa.v65i0.21363, 2013.

. Samuelsson P., Gollvik, S., and Ullerstig, A.: The land-surface scheme of the Rossby Centre regional atmospheric climate model (RCA3). Report in Meteorology 122, SMHI, SE-60176 Norrköping, Sweden, 2006.

Samuelsson, P., Kourzeneva, E. and Mironov, D.: The impact of lakes on the European climate as stimulated by a regional climate model. Boreal Env. Res., 15, 113–129, 2010.

Samuelsson, P., Jones, C., Willén, U., Ullerstig, A., Gollvik, S., Hansson, U., Jansson, C., Kjellström, E., Nikulin, G., and Wyser, K.: The Rossby Centre Regional Climate Model RCA3: Model description and performance, Tellus A, 63, 1–3, https://doi.org/10.1111/j.1600-0870.2010.00478.x, 2011.

Semmler, T., B. Cheng, Y. Yang, and L. Rontu. Snow and ice on Bear Lake (Alaska) – sensitivity experiments with two lake ice models. Tellus A, 64, 17339. https://doi.org/10.3402/tellusa.v64i0.17339, 2012.

SYKE. ISO 19115 Geographic Information Metadata on Hydrological data. Available at http://metatieto.ymparisto.fi:8080/geoportal/catalog/search/resource/details.page?uuid=%7B86FC3188-6796-4C79-AC58-8DBC7B568827%7D, 2018.

Termonia, P., Fischer, C., Bazile, E., Bouyssel, F., Brožková, R., Bénard, P., Bochenek, B., Degrauwe, D., Derkova, M., El Khatib, R., Hamdi, R., Mašek, J., Pottier, P., Pristov, N., Seity, Y., Smolíková, P., Spaniel, O., Tudor, M., Wang, Y., Wittmann, C., and Joly, A.: The ALADIN System and its canonical model configurations AROME CY41T1 and ALARO CY40T1, Geosci. Model Dev., 11, 257-281, https://doi.org/10.5194/gmd-11-257-2018, 2018.

Undén, P., Rontu, L., Järvinen, H., Lynch, P., Calvo, J., Cats, G., Cuxart, J., Eerola, K., Fortelius, C., Garcia-Moya, J. A. et al.: HIRLAM-5 scientific documentation. Available at http://www.hirlam.org, 2002.

Yang, Y., Cheng, B., Kourzeneva, E., Semmler, T., RontuL. and coauthors: 2013. , L., Leppäranta, M., Shirasawa, K. and Li, Z. J.: Modelling experiments on air-snow-ice interactions over Kilpisjärviair–snow–ice interactions over Kilpisjärvi, a lake in northern Finland. Boreal EnvironEnv. Res. ,18, 341–358. : 341–358, 2013.

Table A1. Prognostic and diagnostic lake variables within HIRLAM

variable	unit	type
temperature of snow on lake ice	K	prog by FLake
temperature of lake ice	K	prog by FLake
mean water temperature	K	prog by FLake
mixed layer temperature	K	prog by FLake
bottom temperature	K	prog by FLake
temperature of upper layer sediments	K	prog by FLake
mixed layer depth	m	prog by FLake
thickness of upper layer sediments	m	prog by FLake
thermocline shape factor	-	prog by FLake
lake ice thickness	m	prog by FLake
snow depth on lake ice	m	prog by FLake
LSWT	K	diag by FLake
		= mixed layer temperature if no ice
lake surface temperature	K	diag by FLake
		uppermost temperature: LSWT or ice or snow
LSWT	K	anal by HIRLAM
		flag value 272 K when there is ice
fraction of lake ice	[0-1]	diag fraction in HIRLAM grid
lake surface roughness	m	diag by HIRLAM
screen level temperature over lake	K	diag by HIRLAM
screen level abs.humidity over lake	$kgkg^{-1}$	diag by HIRLAM
anemometer level u-component over lake	ms^{-1}	diag by HIRLAM
anemometer level v-component over lake	ms^{-1}	diag by HIRLAM
latent heat flux over lake	Wm^{-2}	diag by HIRLAM
sensible heat flux over lake	Wm^{-2}	diag by HIRLAM
scalar momentum flux over lake	Pa	diag by HIRLAM
SW net radiation over lake	Wm^{-2}	diag by HIRLAM
LW net radiation over lake	Wm^{-2}	diag by HIRLAM
depth of lake	m	pres in HIRLAM grid
fraction of lake	[0-1]	pres in HIRLAM grid

Denotation: prog = prognostic, diag = diagnostic, pres = prescribed, anal = result of OI

Table A2. Lakes with SYKE observations used in this study.

NAME	LON	LAT	MEAND (m)	MAXD (m)	AREA (kgm^{-2})	HIRD (m)	HIRFR	HIRID
Pielinen	29.607	63.271	10.1	61.0	894.2	10.0	0.916	4001
Kallavesi	27.783	62.762	9.7	75.0	316.1	10.0	0.814	4002
Haukivesi	28.389	62.108	9.1	55.0	560.4	10.0	0.725	4003
Saimaa	28.116	61.338	10.8	85.8	1,377.0	10.0	0.950	4004
Pääjärvi1	24.789	62.864	3.8	14.9	29.5	3.0	0.430	4005
Nilakka	26.527	63.115	4.9	21.7	169.0	10.0	0.866	4006
Konnevesi	26.605	62.633	10.6	57.1	189.2	10.0	0.937	4007
Jääsjärvi	26.135	61.631	4.6	28.2	81.1	10.0	0.750	4008
Päijänne	25.482	61.614	14.1	86.0	864.9	10.0	0.983	4009
Ala-Rieveli	26.172	61.303	11.3	46.9	13.0	10.0	0.549	4010
Kyyvesi	27.080	61.999	4.4	35.3	130.0	10.0	0.810	4011
Tuusulanjärvi	25.054	60.441	3.2	9.8	5.9	3.0	0.174	4012
Pyhäjärvi	22.291	61.001	5.5	26.2	155.2	5.0	0.922	4013
Längelmävesi	24.370	61.535	6.8	59.3	133.0	10.0	0.875	4014
Pääjärvi2	25.132	61.064	14.8	85.0	13.4	14.0	0.350	4015
Vaskivesi	23.764	62.142	7.0	62.0	46.1	10.0	0.349	4016
Kuivajärvi	23.860	60.786	2.2	9.9	8.2	10.0	0.419	4017
Näsijärvi	23.750	61.632	14.7	65.6	210.6	10.0	0.850	4018
Lappajärvi	23.671	63.148	6.9	36.0	145.5	10.0	1.000	4019
Pesiöjärvi	28.650	64.945	3.9	15.8	12.7	7.0	0.290	4020
Rehja-Nuasjärvi	28.016	64.184	8.5	42.0	96.4	10.0	0.534	4021
Oulujärvi	26.965	64.451	6.9	35.0	887.1	10.0	1.000	4022
Ounasjärvi	23.602	68.377	6.6	31.0	6.9	10.0	0.166	4023
Unari	25.711	67.172	5.0	24.8	29.1	10.0	0.491	4024
Kilpisjärvi	20.816	69.007	19.5	57.0	37.3	22.0	0.399	4025
Kevojärvi	27.011	69.754	11.1	35.0	1.0	10.0	0.016	4026
Inarijärvi	27.924	69.082	14.3	92.0	1,039.4	14.0	0.979	4027

Denotation: LON and LAT are the longitude E and latitude N in degrees, MEAND and MAXD are the mean and maximum depths and AREA is the water surface area from the updated lake list of GLDB v.3 (Margarita Choulga, personal communication), HIRD and HIRFR are the mean lake depth and fraction of lakes [0. . . 1] interpolated to the selected HIRLAM gridpoint, taken from the operational HIRLAM that uses GLDB v.2 as the source for lake depths. HIRID is the lake index used by HIRLAM and in this study. Above the middle line are the 27 lakes with both LSWT and LID observations, below the 18 lakes where only LID was available.

Table A3. Lakes with SYKE observations used in this study. Part 2

NAME	LON	LAT	MEAND (m)	MAXD (m)	AREA (kgm^{-2})	HIRD (m)	HIRFR	HIRID
Simpelejärvi	29.482	61.601	9.3	34.4	88.2	10.0	0.548	40241
Pökkäänlahti	27.264	61.501	8.0	84.3	58.0	10.0	0.299	40261
Muurasjärvi	25.353	63.478	9.0	35.7	21.1	10.0	0.060	40263
Kalmarinselkä	25.001	62.786	5.7	21.9	7.1	5.0	0.330	40271
Summasjärvi	25.344	62.677	6.7	40.5	21.9	10.0	0.555	40272
Iisvesi	27.021	62.679	17.2	34.5	164.9	18.0	0.456	40277
Hankavesi	26.826	62.614	7.0	49.0	18.2	18.0	0.100	40278
Petajävesi	25.173	62.255	4.2	26.6	8.8	3.0	0.245	40282
Kukkia	24.618	61.329	5.2	35.6	43.9	10.0	0.299	40308
Ähtärinjärvi	24.045	62.755	5.2	27.0	39.9	10.0	0.266	40313
Kuortaneenjärvi	23.407	62.863	3.3	16.2	14.9	10.0	0.277	40328
Lestijärvi	24.716	63.584	3.6	6.9	64.7	10.0	0.513	40330
Pyhäjärvi	25.995	63.682	6.3	27.0	121.8	10.0	0.266	40331
Lentua	29.690	64.204	7.4	52.0	77.8	7.0	0.600	40335
Lammasjärvi	29.551	64.131	4.3	21.0	46.8	3.0	0.200	40336
Naamankajärvi	28.246	65.104	2.9	14.0	8.5	7.0	0.299	40342
Korvuanjärvi	28.663	65.348	6.0	37.0	15.4	10.0	0.342	40343
Oijärvi	25.930	65.621	1.1	2.4	21.0	10.0	0.333	40345

Denotation: LON and LAT are the longitude E and latitude N in degrees, MEAND and MAXD are the mean and maximum depths and AREA is the water surface area from the updated lake list of GLDB v.3 (Margarita Choulga, personal communication), HIRD and HIRFR are the mean lake depth and fraction of lakes [0...1] interpolated to the selected HIRLAM gridpoint, taken from the operational HIRLAM that uses GLDB v.2 as the source for lake depths. HIRID is the lake index used by HIRLAM and in this study. Above the middle line are the 27 lakes with both LSWT and LID observations, below the 18 lakes where only LID was available.

Validation of lake surface state in the HIRLAM v.7.4 NWP model against *in-situ* measurements in Finland

Laura Rontu[1], Kalle Eerola[1], and Matti Horttanainen[1]

[1]Finnish Meteorological Institute, P.O. Box 503, 00101 Helsinki, Finland

Correspondence: laura.rontu@fmi.fi

Abstract. High Resolution Limited Area Model (HIRLAM), used for the operational numerical weather prediction in the Finnish Meteorological Institute (FMI), includes prognostic treatment of lake surface state since 2012. Forecast is based on the Freshwater Lake (FLake) model integrated into HIRLAM. Additionally, an independent objective analysis of lake surface water temperature (LSWT) combines the short forecast of FLake to observations from the Finnish Environment Institute (SYKE). The resulting description of lake surface state - forecast FLake variables and analysed LSWT - was compared to SYKE observations of lake water temperature, ~~freezing and melting~~ freeze-up and break-up dates as well as the ice thickness and snow depth for 2012-2018 over 45 lakes in Finland. During the ice-free period, the predicted LSWT corresponded to the observations with a slight overestimation, with a systematic error of + 0.91 K. The colder temperatures were underrepresented and the maximum temperatures were too high. The objective analysis of LSWT was able to reduce the bias to + 0.35 K. The predicted ~~freezing~~ freeze-up dates corresponded well the observed dates, mostly within the accuracy of a week. The forecast ~~melting~~ break-up dates were far too early, typically several weeks ahead of the observed dates. The growth of ice thickness after ~~freezing~~ freeze-up was generally overestimated. However, practically no predicted snow appeared on lake ice. The absence of snow, ~~found to~~ presumably be due to ~~a technical error in HIRLAM~~an incorrect security coefficient value, is suggested to be also the main reason of the inaccurate simulation of the lake ice ~~melt~~ melting in spring.

1 Introduction

Lakes influence the energy exchange between the surface and the atmosphere, the dynamics of the atmospheric boundary layer and the near-surface weather. This is important for weather forecasting over the areas where lakes, especially those with a large yearly variation of the water temperature, freezing in autumn and melting in spring, cover a significant area of the surface (Kheyrollah Pour et al., 2017; Laird et al., 2003 and references therein). Description of the lake surface state influences the numerical weather prediction (NWP) results, in particular in the models whose resolution is high enough to account for even the smaller lakes (Eerola et al., 2014 and references therein). Especially, the existence of ice can be important for the numerical forecast (Eerola et al., 2014; Cordeira and Laird, 2008).

In the Finnish Meteorological Institute (FMI), the High Resolution Limited Area Model HIRLAM (Undén et al., 2002; Eerola, 2013) has been applied since 1990 for the numerical short-range weather forecast. In the beginning, the monthly climatological water surface temperature for both sea (sea surface temperature SST) and lakes (Lake Surface Water Temperature LSWT) was used. Since 2012, HIRLAM includes a prognostic lake temperature parameterization based on the Freshwater Lake Model (FLake, Mironov et al., 2010). An independent objective analysis of observed LSWT (Kheyrollah Pour et al., 2017 and references therein) was implemented in 2011. The fractional ice cover (lake ice concentration in each gridsquare of the model) is diagnosed from the analysed LSWT.

FLake was designed to be used as a parametrization scheme for the forecast of the lake surface state in NWP and climate models. It allows to predict the lake surface state in interaction with the atmospheric processes treated by the NWP model. The radiative and turbulent fluxes as well as the predicted snow precipitation from the atmospheric model are combined with FLake processes at each time-step of the model integration in the model grid, where the fraction and depth of lakes are prescribed.

FLake has been implemented into the other main European NWP and regional climate models, first into COSMO (Mironov et al., 2010) then into ECMWF (Balsamo et al., 2012), Unified Model (Rooney and Bornemann, 2013), SURFEX surface modelling framework (Masson et al., 2016), regional climate models RCA (Samuelsson et al., 2010), HCLIM (Lindstedt et al., 2015) and REMO (Pietikäinen et al., 2018), among others. Description of lake surface state and its influence in the numerical weather and climate prediction has been validated in various ways. Results of case studies, e.g. Eerola et al. (2014) and shorter-period NWP experiments, e.g. Eerola et al. (2010); Rontu et al. (2012); Kheyrollah Pour et al. (2014); Kheyrollah Pour et al. (2017) as well as climate model results, e.g. Samuelsson et al. (2010); Pietikäinen et al. (2018), have been compared with remote-sensing satellite data and *in-situ* lake temperature and ice measurements as well as validated against the standard weather observations. In general, improvement of the scores has been seen over regions where lakes occupy a significant area. However, specific features of each of the host models influence the results of the coupled atmosphere-lake system as FLake appears to be quite sensitive to the forcing by the atmospheric model.

The aim of the present study is to ~~use~~ validate the lake surface state forecast by the operational HIRLAM NWP model using the *in-situ* LSWT measurements, lake ice ~~freezing and melting~~ freeze-up and break-up dates and measurements of ice and snow thickness by the Finnish Environment Institute (Suomen Ympäristökeskus = SYKE)~~for validation of the lake surface state forecast by the operational HIRLAM NWP model~~. For this purpose, HIRLAM analyses and forecasts archived by ~~the Finnish Meteorological Institute (FMI)~~ FMI were compared with the observations by SYKE over the lakes of Finland from spring 2012 to summer 2018. To our knowledge, this is the longest available detailed dataset that allows to evaluate how well the lake surface state is simulated by an operational NWP model that applies FLake parametrizations.

2 Lake surface state in HIRLAM

FLake was implemented in the HIRLAM forecasting system in 2012 (Kourzeneva et al., 2008; Eerola et al., 2010). The model utilizes external datasets on the lake depth (Kourzeneva et al., 2012a; Choulga et al., 2014) and the lake climatology (Kourzeneva et al., 2012b). The latter is only needed in order to provide initial values of FLake prognostic variables in the very first forecast (so-called cold start). Real-time The use of real-time *in-situ* LSWT observations by SYKE for 27 Finnish lakes were obtained was introduced in 2011 to be used for into the operational LSWT analysis in HIRLAM (Eerola et al., 2010; Rontu et al., 2012). In the current operational HIRLAM at FMIof FMI, FLake provides the background for the optimal interpolation analysis (OI, based on Gandin, 1965) of LSWT. However, the prognostic FLake variables are not corrected using the analysed LSWT. This would require more advanced data assimilation methods based on e.g. the extended Kalman filter (Kourzeneva, 2014).

2.1 Freshwater lake model in HIRLAM

FLake is a bulk model capable of predicting the vertical temperature structure and mixing conditions in lakes of various depths on time-scales from hours to years (Mironov et al., 2010). The model is based on two-layer parametric representation of the evolving temperature profile in the water and on the integral budgets of energy for the layers in question. Bottom sediments and the thermodynamics of the ice and snow on ice layers are treated separately. FLake depends on prescribed lake depth information. The prognostic and diagnostic variables of HIRLAM FLake together with the analysed lake surface variables in HIRLAM are listed in the Appendix (Table A1).

At each time step of during the HIRLAM forecast, FLake is driven by the atmospheric radiative and turbulent fluxes as well as the predicted snowfall, provided by the physical parameterisations in HIRLAM. This couples the atmospheric variables over lakes with the lake surface properties as provided by FLake parametrization. Most importantly, FLake provides HIRLAM with the evolving lake surface (water, ice, snow) temperature , that influences and radiative properties, that influence the HIRLAM forecast of the grid-average near-surface temperatures.

Implementation of FLake model as a parametrization scheme in HIRLAM was based on the experiments described by Rontu et al., 2012Rontu et al. (2012). Compared to the reference version of FLake (Mironov et al., 2010), minor modifications were introduced, namely, use of constant snow density = 300 kgm^{-3}, molecular heat conductivity = 1 Jm^{-1}s^{-1}K^{-1}, constant albedos of dry snow = 0.75 and ice = 0.5. Bottom sediment calculations were excluded. Global lake depth database (GLDB v.2, Choulga et al., 2014) was used for derivation of mean lake depth in each gridsquare. Fraction of lake was taken from HIRLAM physiography database, where it originates from GLCC (Loveland et al., 2000).

Lake surface temperature is diagnosed from the mixed layer temperature for the unfrozen lake gridpoints and from the ice or snow-on-ice temperature for the frozen points. In FLake, ice starts to grow from an assumed value of one millimeter when temperature reaches the freezing point. The whole lake tile in a gridsquare is considered by FLake either frozen or unfrozen. Snow on ice is accumulated from the model's snowfall at each time step during the numerical integration.

2.2 Objective analysis of LSWT observations

A comprehensive description of the optimal interpolation (OI) of the LSWT observations in HIRLAM is given by (Kheyrollah Pour et al., 2017). Shortly, LSWT analysis is obtained by correcting the FLake forecast at each gridpoint by using the weighted average of the deviations of observations from their background values. Prescribed statistical information about the observation and background error variance as well as the distance-dependent autocorrelation between the locations (observations and gridpoints) are applied. The real-time observations entering the HIRLAM surface analysis system are subject to quality control in two phases. First, the observations are compared to the background, provided by the FLake short forecast. Second, optimal interpolation is done at each observation location, using the neighbouring observations only (excluding the current observation) and comparing the result to the observed value at the station.

A specific feature of the lake surface temperature OI is that the interpolation is performed not only within the (large) lakes but also across the lakes: within a statistically pre-defined radius, the observations affect all gridpoints containing a fraction of lake. This ensures that the analysed LSWT on lakes without own observations may also be influenced by observations from neighbouring lakes, not only by the first guess provided by FLake forecast.

The relations between the OI analysis and the prognostic FLake in HIRLAM are schematically illustrated in Figure 1. Within the present HIRLAM setup, the background for the analysis is provided by the short (6-hour) FLake forecast but the next forecast is not initialized from the analysis. Instead, FLake continues running from the previous forecast, driven by the atmospheric state given by HIRLAM at each time step. This means that FLake does not benefit from the result of OI analysis but the analysis remains as an extra diagnostic field, to some extent independent of the LSWT forecast. However, FLake background has a large influence in the analysis, especially over distant lakes where neighbouring observations are not available. The diagnostic LSWT analysis, available at every gridpoint of HIRLAM, might be useful e.g. for hydrological, agricultural or road weather applications.

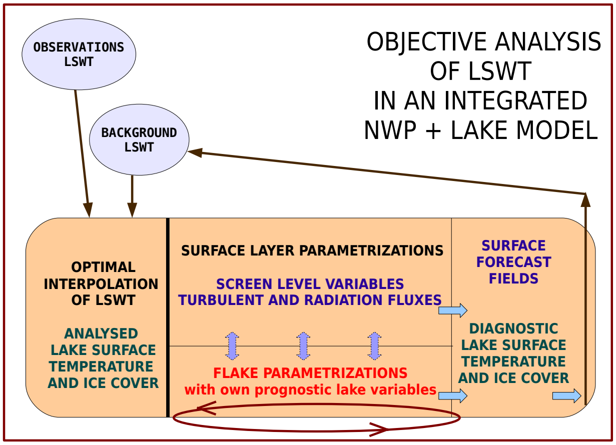


Figure 1. Coexistence of the independent objective analysis of the observed LSWT and prognostic FLake parametrizations in HIRLAM. The thin arrows are related to data flow between HIRLAM analysis-forecast cycles while the thick arrows describe processes within each cycle.

Missing LSWT observations in spring and early winter are interpreted to represent presence of ice and given a flag value of -1.2°C. If, however, the results of the statistical LSWT model (Elo, 2007), provided by SYKE along with the real-time observations, indicate unfrozen conditions, the observations are considered missing. This prevents appearance of ice in summer when observations are missing but leads to a misinterpretation of data in spring if the SYKE model indicates too early melting.

In the analysis, fraction of ice is diagnosed from the LSWT field in a simple way. The lake surface within a gridsquare is assumed fully ice-covered when LSWT falls below -0.5°C and fully ice-free when LSWT is above 0°C. Between these temperature thresholds, the fraction of ice changes linearly (Kheyrollah Pour et al., 2014).

The HIRLAM surface data assimilation system produces comprehensive feedback information from every analysis-forecast cycle. The feedback consists of the observed value and its deviations from the background and from the final analysis at the observation point. Bilinear interpolation of the analysed and forecast values is done to the observation location from the nearest gridpoints that contain a fraction of lake. In addition, information about the quality check and usage of observations is provided. Fractions of land and lake in the model grid as well as the weights, which were used to interpolate gridpoint values to the observation location, are given. This information is the basis of the present study (see sections 3.3 and 4).

3 Model-observation intercomparison 2012-2018

In this intercomparison we validated HIRLAM /FLake results against observations about the lake surface state. The impact of FLake parametrizations to the weather forecast by HIRLAM was not considered. This is because no non-FLake weather forecasts exist for comparison with the operational forecasts during the validation period.

Throughout the following text, the analysed LSWT refers to the result of OI analysis, where FLake forecast has been used as background (Section 2.2) while the forecast LSWT refers to the value diagnosed from the mixed layer water temperature predicted by FLake (Section 2.1). Observed LSWT refers to the measured by SYKE lake water temperature (Section 3.2).

3.1 FMI operational HIRLAM

FMI operational HIRLAM is based on the last reference version (v.7.4), implemented in spring 2012. (Eerola, 2013 and references therein). FLake was introduced into this version. After that the development of HIRLAM was frozen. Thus, during the years of the present comparison, the FMI operational HIRLAM system remains unmodified, which offers a clean time series of data for the model-observation intercomparison. The general properties of the system are summarised in Table 1. In the present study, a coding error in FLake implementation was revealed in the reference HIRLAM v.7.4. A too large critical value to diagnose snow existence prevented practically all accumulation of the forecast snowfall on lake ice in the FMI HIRLAM-FLake operational system.

3.2 SYKE lake observations

In this study we used three different types of SYKE lake observations: LSWT, lake ice dates (LID) freeze-up and break-up dates and ice thickness and snow depth on lake ice. In total, observations on 45 lakes listed in Appendix (Table A2) were included as detailed in the following. The lake depths and surface areas given in Table A2 are based on he updated lake list of GLDB v.3 (Margarita Choulga, personal communication).

Table 1. FMI operational HIRLAM

Domain	From Atlantic to Ural, from North Africa beyond North Pole
Model horizontal / vertical resolution	7 km / 65 levels
HIRLAM version	7.4
Model dynamics	Hydrostatic, semi-Lagrangian, grid-point
Atmospheric physical parametrizations	Savijärvi radiation, CBR turbulence,
	Rasch-Kristiansson cloud microphysics + Kain-Fritsch convection
Surface physical parametrizations	ISBA-newsnow for surface, FLake for lakes
Data assimilation	Default atmospheric (4DVAR) and surface (OI) analysis
Lateral boundaries	ECMWF forecast
Forecast	Up to +54 h initiated every 6h (00, 06, 12, 18 UTC)

3.2.1 Lake temperature measurements

Regular *in-situ* lake water temperature measurements are performed by SYKE. Currently SYKE operates 34 regular lake and river water temperature measurement sites in Finland. The temperature of the lake water is measured every morning at 8.00 AM local time, close to shore, at 20 cm below the water surface. The measurements are recorded either automatically or manually and are performed only during the ice-free season ~~(?Rontu et al., 2012)~~(Korhonen, 2019). Further, we will for simplicity denote also these data as LSWT observations although they do not represent exactly the same surface water temperature (skin temperature, radiative temperature) that could be estimated by satellite measurements. These data are available in the SYKE open data archive (SYKE, 2018). Measurements from 27 of these 34 lakes (Figure 2, white dots) were selected for use in the FMI operational HIRLAM in 2011, and the list has been kept unmodified since that. The set of 27 daily observations, quality-controlled by HIRLAM, were obtained from the analysis feedback files and used in all comparisons reported in this study.

3.2.2 ~~Freezing~~ Freeze-up and ~~melting~~ break-up dates

Regular visual observations of ~~freezing and melting~~ freeze-up and break-up of lakes have been recorded in Finland for centuries, the longest time series starting in the middle of the 19th century ~~(?).~~ (Korhonen, 2019). Presently, dates of ~~freezing and melting~~ freeze-up and break-up are available from SYKE (2018) on 123 lakes, but the time series for many lakes are discontinuous. Further, we will denote the ~~melting and freezing~~ break-up and freeze-up dates together by “lake ice dates” (LID). ~~For both freezing and melting~~ LID observations aim at representing conditions on entire lakes. For both freeze-up and break-up the dates are available in two categories ~~: for freezing “freezing of the visible area”~~ (terminology from Korhonen, 2019): "freeze-up of the lake within sight" (code 29 by SYKE) and ~~“permanent freezing of the visible area”~~ "freeze-up of the whole lake" (code 30). For ~~melting~~ break-up the dates are defined as ~~“no ice visible from the observation site”~~ “no ice within sight” (code 28) and “~~no ice on the outer open water areas~~ thaw areas out of the shore” (code 27). LID observations ~~aim at representing conditions~~

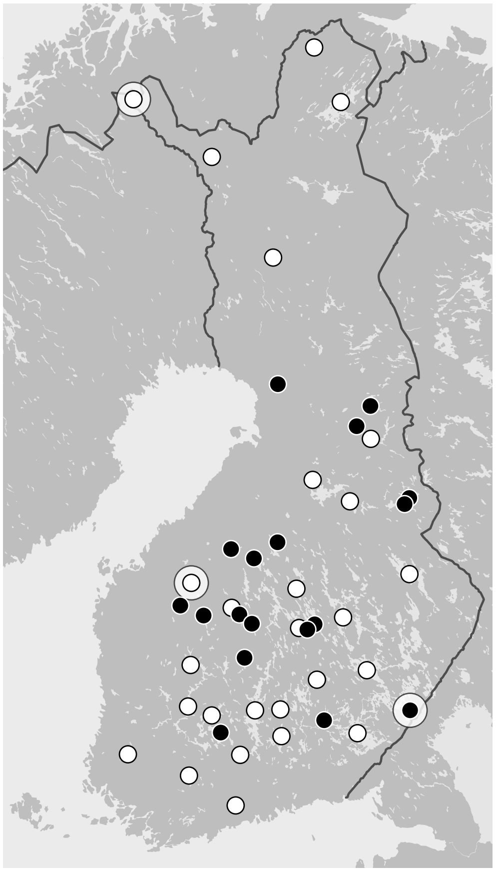


Figure 2. Map of SYKE observation points used in this study: lakes with both lake surface water temperature (LSWT) and lake ice date (LID) observations (white), lakes where only LID is available (black). On Lakes Lappajärvi, Kilpisjärvi and Simpelejärvi also ice thickness and snow depth measurements were used (Section 4.3), they are surrounded with a large white circle. List of the lakes with coordinates is given in Appendix A2.

~~on entire lakes. LID observations~~ by SYKE are made independently of their LSWT measurements and possibly from different locations on the same lakes. The LSWT measurements may be started later than the date of reported lake ice ~~melting~~ break-up or end earlier than the reported ~~freezing~~ freeze-up date.

LID from the 27 lakes whose LSWT measurements are used in HIRLAM were available and selected for this study. In addition, 18 lakes with only LID available (Figure 2, black dots) were chosen for comparison with HIRLAM ~~/FLake~~ LID.

3.2.3 Ice thickness and snow depth on lakes

~~SYKE records~~ In the period 2012-2018 SYKE recorded the lake ice thickness and snow depth on around 50 locations in Finland. ~~Archived~~ (Archived historical data are available in total from 160 measurement sites). The manual measurements are done three times a month during the ice season. Thickness of ice and snow depth on ice are measured by drilling holes through snow and ice layers along chosen tracks, normally at least 50 m from the coast ~~(?)~~(Korhonen, 2019). The locations may differ from those of the LSWT measurement or LID observation over the same lakes.

3.3 Validation of HIRLAM ~~/FLake~~ lake surface state

3.3.1 Lake surface water temperature

LSWT by HIRLAM~~/FLake~~, resulting from the objective analysis or diagnosed from the forecast, was compared with the observed LSWT by SYKE using data extracted from the analysis feedback files (Section 2.2) at the observation locations on 06 UTC every day, excluding the winter periods 1 December - 31 March. The observations (ob) at 27 SYKE stations were assumed to represent the true value, while the analysis (an) is the result of OI that combines the background forecast (fc) with the observations. Time-series, maps and statistical scores, to be presented in Section 4.1, were derived from these.

3.3.2 Lake ice conditions

For this study, the observed LID, ice and snow thickness observations were obtained from SYKE open data base, relying on their quality control. The ~~HIRLAM/FLake~~ analysed LSWT as well as the predicted ice thickness and snow depth were picked afterwards from the HIRLAM archive for a single gridpoint nearest to each of the 45 observation locations (not interpolated as in the analysis feedback file that was used for the LSWT comparison). It was assumed that the gridpoint value nearest to the location of the LSWT observation represents the ice conditions over the chosen lake.

LID given by HIRLAM were defined in two independent ways: from the analysed LSWT and from the forecast lake ice thickness. Note that the ice thickness and snow depth on ice are not analysed variables in HIRLAM. In autumn a lake can freeze and melt several times before final ~~freezing~~freeze-up. The last date when the forecast ice thickness crossed a critical value of 1 mm or the analysed LSWT fell below freezing point was selected as the date of ~~freezing~~freeze-up. In the same way, the last date when the forecast ice thickness fell below the critical value of 1 mm or the analysed LSWT value crossed the freezing point was selected as ~~melting day~~break-up date. To decrease the effect of oscillation of the gridpoint values between the HIRLAM forecast-analysis cycles, the mean of the four daily ice thickness forecasts or analysed LSWT values was used.

LID by HIRLAM ~~/FLake~~ were compared to the observed dates during 2012-2018. In this comparison we included data also during the winter period. The category 29 observations ("~~freezing of the visible area~~freeze-up of the lake within sight", see Section 3.2.2) were used. In this category the time series were the most complete at the selected stations. For the same reason, the ~~melting~~ break-up observations of category 28 ("no ice ~~visible from the observation site~~within sight") were used for comparison. Furthermore, using a single gridpoint value for the calculation of LID also seems to correspond best the observation definition based on what is visible from the observation site. The statistics were calculated as fc - ob and an - ob. Hence, positive values mean that ~~melting or freezing~~ break-up or freeze-up takes place too late in the model as compared to the observations.

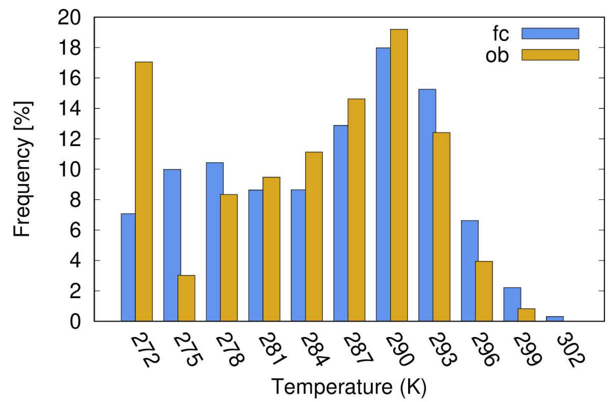
~~In this study, lake~~ Lake ice thickness and snow depth measurements from lakes Lappajärvi, Kilpisjärvi and Simpelejärvi were utilised as additional data for validation of predicted by HIRLAM ~~/FLake~~ ice thickness and snow depth (Section 4.3).

These lakes, representing the western, northern and south-eastern Finland, were selected for illustration based on the best data availability during the study years. They are also sufficiently large in order to fit well the HIRLAM grid.

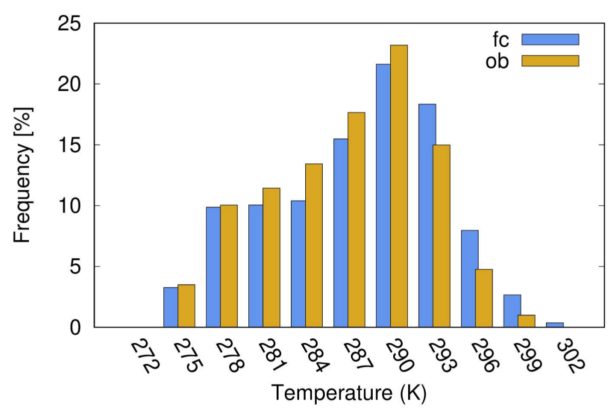
4 Results

4.1 Analysed and forecast LSWT at observation points

Figure 3 shows the frequency distribution of LSWT according to FLake forecast and SYKE observations. It is evident that the amount of data in the class of temperatures which represents frozen conditions (LSWT flag value 272 K) was underestimated by the forecast (Figure 3a). When subzero temperatures were excluded from the comparison (Figure 3b), underestimation in the colder temperature classes and overestimation in the warmer classes still remains.



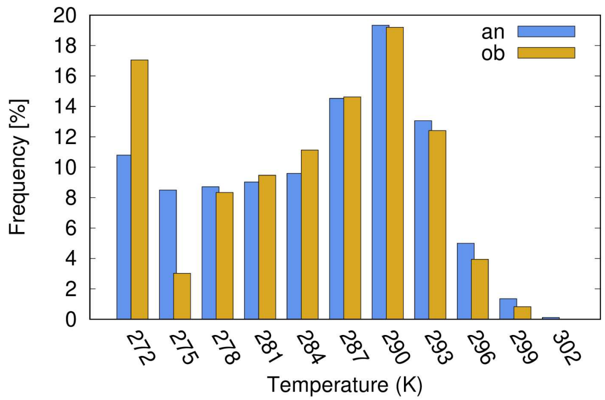
(a) with all temperatures (also frozen conditions) included



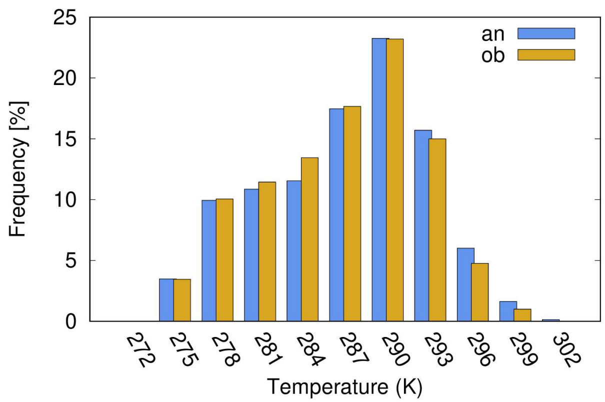
(b) only open water temperatures included

Figure 3. Frequency of observed (ob, yellow) and forecast (fc, blue) LSWT over all 27 SYKE lakes 2012-2018. x-axis: LSWT, unit K, y-axis: frequency, unit %.

LSWT analysis (Figure 4) improved the situation somewhat but the basic features remain. This is due to the dominance of FLake forecast via the background of the analysis. In Section 4.3, we will show time-series illustrating the physics behind these LSWT statistics.



(a) with all temperatures (also frozen conditions) included



(b) only open water temperatures included

Figure 4. As for Figure 3 but for observed and analysed (an) LSWT.

Table 2 confirms the warm bias by FLake in the unfrozen conditions. Similar results were obtained for all stations together and also for our example lakes Lappajärvi and Kilpisjärvi, to be discussed in detail in Section 4.3. There were three lakes with negative LSWT bias according to FLake forecast, namely the large lakes Saimaa and Päijänne and the smaller Ala-Rieveli. After the correction by objective analysis, a small positive bias converted to negative over 6 additional lakes, among them the large lakes Lappajärvi in the west and Inari in the north. The mean absolute error decreased from forecast to analysis on every lake.

In the frequency distributions, the warm temperatures were evidently related to summer. For FLake, the overestimation of maximum temperatures, especially in shallow lakes, is a known feature (e.g. Kourzeneva 2014). It is related to the difficulty

Table 2. Statistical scores for LSWT at all stations and at two selected stations

station	fc or an	mean ob	bias	mae	stde	N
unit		K	K	K	K	
ALL	fc	286.3	0.91	1.94	2.34	30877
	an	286.3	0.35	1.32	1.72	30861
Lappajärvi	fc	286.9	0.33	1.23	1.62	1243
	an	286.9	-0.65	1.06	1.10	1243
Kilpisjärvi	fc	281.7	1.82	2.13	2.15	780
	an	281.7	1.10	1.42	1.51	780

Statistics over days when both forecast/analysis and observation indicate unfrozen conditions. bias = systematic difference fc/an - ob, mae = mean absolute error, stde = standard deviation of the error, N = number of days (06 UTC comparison, no ice).

of forecasting the mixed layer thermodynamics under strong solar heating. Cold and subzero temperatures occurred in spring and autumn. In a few large lakes like Saimaa, Haukivesi, Pielinen, LSWT tended to be slightly underestimated in autumn both according to the FLake and the analysis (not shown). The cold left-hand side columns in the frequency distributions (Figures 3a and 4a) are mainly related to spring, when HIRLAM ~~/FLake~~ tended to melt the lakes significantly too early (Sections 4.2 and 4.3).

There are problems, especially in the analysed LSWT, over (small) lakes of irregular form that fit poorly the HIRLAM grid and where the measurements may represent more the local than the mean or typical conditions over the lake. These are the only ones where an underestimation of summer LSWT was seen. Cases occurred where FLake results differ so much from the observations that the HIRLAM quality control against background values rejected the observations, forcing also the analysis to follow the incorrect forecast (not shown).

4.2 ~~Freezing~~ Freeze-up and ~~melting~~ break-up dates

In this section the ~~freezing and melting~~ freeze-up and break-up dates from HIRLAM are verified against corresponding observed dates over 45 lakes (Appendix Table A2). In the following, 'LSWT an' refers to the LID estimated from analysed LSWT and 'IceD fc' to those estimated from the forecast ice thickness by FLake. The time period contains six freezing periods (from autumn 2012 to autumn 2017) and seven melting periods (from spring 2012 to spring 2018). Due to some missing data the number of ~~freezing~~ freeze-up cases was 233 and ~~melting~~ break-up cases 258. The 'IceD fc' data for the first melting period in spring 2012 was missing. The overall statistics of the error in ~~freezing and melting~~ freeze-up and break-up dates are shown in Table 3. In most cases the difference in error between the dates based on forecast and analysis was small. This is natural as the first guess of the LSWT analysis is the forecast LSWT by FLake. We will discuss next the ~~freezing~~freeze-up, then the ~~melting~~ break-up dates.

Table 3. Statistical measures of the error of ~~freezing~~ freeze-up and ~~melting~~ break-up date

unit		bias days	stde days	max days	min days	N
~~Freezing~~ Freeze-up	LSWT an	-3.5	17.9	64	-52	233
	IceD fc	-0.3	17.8	67	-41	233
~~Melting~~ Break-up	LSWT an	-15.2	8.5	2	-54	288
	IceD fc	-20.5	9.2	-1	-56	258

Denotation: LSWT an - LID estimated from analysed LSWT, IceD fc - LID estimated from forecast ice thickness.

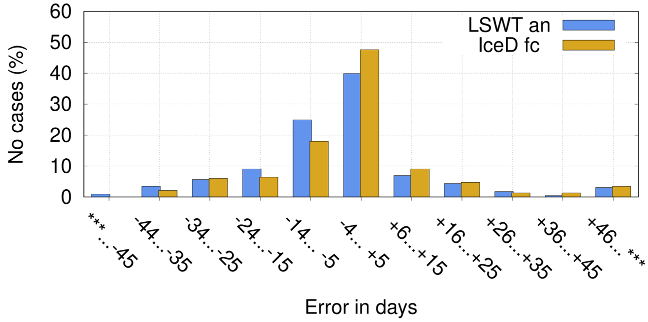
The bias in the error of ~~freezing~~ freeze-up dates was small according to both 'IceD fc' and 'LSWT an', -0.3 and -3.5 days, respectively. The minimum and maximum errors were large in both cases: the maximum ~~freezing day~~ freeze-up date occurred about two months too late, the minimum about one and a half months too early. However, as will be shown later, the largest errors mostly occurred on a few problematic lakes while in most cases the errors were reasonable.

Figure 5a) shows the frequency distribution of the error of ~~freezing~~ freeze-up dates. Forecast ~~freezing~~ freeze-up dates occurred slightly more often in the unbiased class (error between -5 - +5 days), compared to the estimated dates from the analysis. Of all cases 48 %/ 40 % (percentages here and in the following are given as 'IceD fc' / 'LSWT an') fell into this class. In 20% / 26% of cases the ~~freezing~~ freeze-up occurred more than five days too late and only in 11% / 9% cases more than two weeks too late. In case of 'IceD fc', the class of ~~freezing~~ freeze-up more than 15 days too late comprised 25 cases distributed over 15 lakes, thus mostly one or two events per lake. This suggests that the error was related more to individual years than to systematically problematic lakes. It is worth noting, that of the eight cases where the error was over 45 days, six cases were due to a single lake, Lake Kevojärvi. This lake is situated in the very north of Finland. It is very small and narrow, with an area of 1 km^2, and located in a steep canyon. Therefore it is poorly represented by the HIRLAM grid and the results seem unreliable.

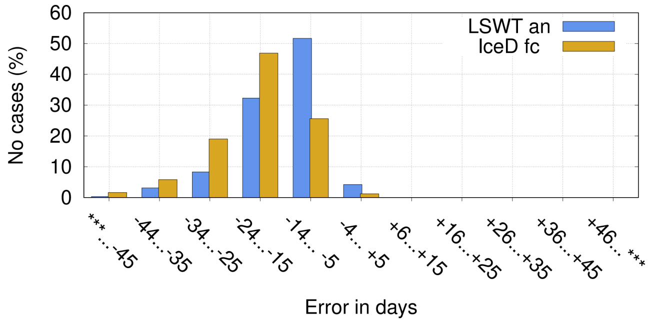
Concerning too early freezing, in 33% / 44% of the cases ~~freezing~~ freeze-up occurred more than five days too early and in 15% / 19% more than two weeks too early. According to the forecast, these 15% (34 cases) were distributed over 19 lakes. Each of the five large lakes Pielinen, Kallavesi, Haukivesi, Päijänne and Inari occurred in this category three times while all other lakes together shared the remaining 19 cases during the six winters.

The ~~melting~~ break-up dates (Table 3) show a large negative bias, about two ('LSWT an') or three weeks ('IceD fc'), indicating that lake ice ~~melting~~ break-up was systematically forecast to occur too early. However, the standard deviation of the error was only about half of that of the error of ~~freezing~~ freeze-up dates and there were no long tails in the distribution (Figure 5b). ~~The~~ Hence the distribution is strongly skewed towards too early ~~melting~~break-up, but much narrower than that of ~~freezing~~ freeze-up (Figure 5a). The large bias was most probably due ~~to the bug of this HIRLAM version that prevented the accumulation of~~ missing snow over lake ice ~~(see also Section 4.3~~in this HIRLAM version (see Section 5). The maximum frequency (47 %) was in the class -24 - -15 days for 'IceD fc', while in case of 'LSWT an', the maximum frequency (52 %)

occurred in the class -14 - -5 days. FLake forecast 'IceD fc' suggested only three cases in the unbiased class -4 - +5 while according to 'LSWT an' there were 12 cases in this class. Hence, the ~~melting~~ break-up dates derived from analysed LSWT corresponded the observations better than those derived from FLake ice thickness forecast.



(a) error ~~in freezing days~~of freeze-up dates



(b) error ~~in melting days~~of break-up dates

Figure 5. Frequency distribution of the difference between analysed/forecast and observed ~~freezing~~ freeze-up and ~~melting days~~ break-up dates over all lakes 2012-2018. Variables used in diagnosis of ice existence: analysed LSWT crossing the freezing point (blue) and forecast ice thickness $>$ 1 mm (magenta). Observed variable: ~~freezing~~ freeze-up date by SYKE. x-axis: difference (fc-ob), unit day, y-axis: percentage of all cases.

Note that this kind of method of verifying LID compares two different types of data. The observations by SYKE are visual observations from the shore of the lake (see Section 3.2.2), while the ~~freezing and melting~~ freeze-up and break-up dates from HIRLAM are based on single-gridpoint values of LSWT or ice thickness (see Section 3.3.2). In addition, the resulting ~~freezing and melting~~ freeze-up and break-up dates from HIRLAM are somewhat sensitive to definition of the freezing and melting tresholds. Here we used 1 mm for the forecast ice thickness and the freezing point for the LSWT analysis as the critical values.

In conclusion, the validation statistics show that HIRLAM ~~/FLake~~ succeeded rather well in predicting freezing of Finnish lakes. Almost in half of the cases the error was less than $\pm$ 5 days. Some bias towards too early ~~freezing~~ freeze-up can be seen both in forecast and in the analysis. Melting was more difficult. FLake predicted ~~ice melting~~ lake ice break-up always too early, with a mean error of over two weeks, and the analysis mostly followed it. These results are rather obvious because of the missing snow on ice.

4.3 Comparisons on three lakes

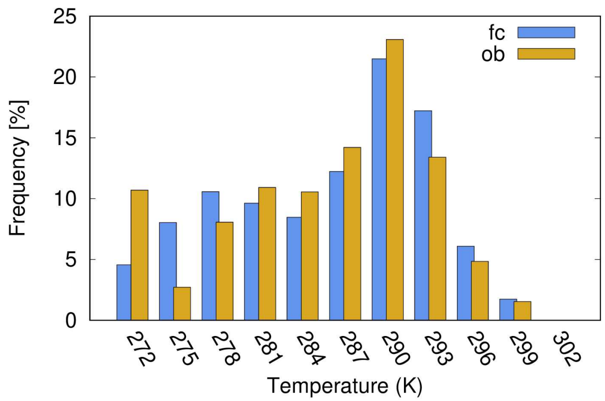
In this section we present LSWT and LID time-series for two representative lakes, Kilpisjärvi in the north and Lappajärvi in the west (see the map in Figure 2). Observed and forecast ice and snow thickness are discussed, using also additional data from Lake Simpelejärvi in southeastern Finland.

Lake Kilpisjärvi is an Arctic lake at the elevation of 473 m, surrounded by fells. The lake occupies 40 % of the area of HIRLAM gridsquare covering it (the mean elevation of the gridsquare is 614 m). The average/maximum depths of the lake are 19.5/57 m and the surface area is 37.3 km^2. The heat balance as well as the ice and snow conditions on Lake Kilpisjärvi have been subject to several studies (Leppäranta et al., 2012; Lei et al., 2012; Yang et al., 2013). Typically, the ice season lasts there seven months from November to May. Lake Lappajärvi is formed from a 23 km wide meteorite impact crater, which is estimated to be 76 million years old. It is Europe's largest crater lake with a surface area of 145.5 km^2 and an average/maximum depth of 6.9/36 m. Here the climatological ice season is shorter, typically about five months from December to April. The average/maximum depth of Lake Simpelejärvi is 8.7/34.4 m and the surface area 88.2 km^2. This lake is located at the border between Finland and Russia and belongs to the catchment area of Europe's largest lake, Lake Ladoga in Russia.

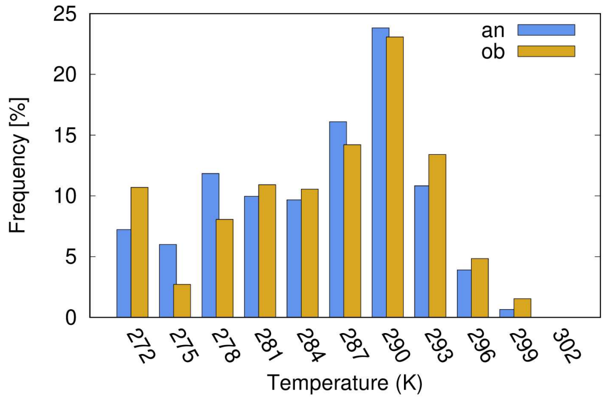
Figures 6 and 7 show the frequency distributions of LSWT according to forecast v.s. observation and analysis v.s. observation for Lappajärvi and Kilpisjärvi. Features similar to the results averaged over all lakes (Section 4.1, Figures 3 and 4) are seen, i.e. underestimation of the amount of cold temperature cases and overestimation of the warmer temperatures by the forecast and analysis. On Lake Lappajärvi, only the amount of below-freezing temperatures was clearly underestimated, otherwise the distributions look quite balanced. According to the observations, on Lake Kilpisjärvi ice-covered days dominated during the periods from April to November November to May. According to both FLake forecast and HIRLAM LSWT analysis LSWT analysis and forecast the amount of these days was clearly smaller in HIRLAM.

Yearly time series of the observed, forecast and analysed LSWT, with the observed LID marked, are shown in Figures 8 and 9. In the absence of observations, the HIRLAM analysis followed the forecast. Missing data in the time series close to freezing and melting freeze-up and break-up are due to missing observations, hence missing information in the feedback files (see Section 2.2). Differences between the years due to the different prevailing weather conditions can be are seen in the temperature variations.

Generally, FLake tended to melt the lakes too early in spring, as already indicated by the LID statistics (Section 4.2). The too early melting break-up and too warm LSWT in summer show up clearly in Kilpisjärvi (Figure 9). In Lappajärvi, the model and analysis were able to follow even quite large and quick variations of LSWT in summer, but tended to somewhat overestimate the maximum temperatures. Overestimation of the maximum temperatures by FLake was still more prominent in shallow lakes (not shown). In autumn over Lakes Lappajärvi and Kilpisjärvi, the forecasts and analyses followed closely the LSWT observations and reproduced the freezing freeze-up dates within a few days, which was also typical to the majority of lakes.



(a) forecast v.s. observation



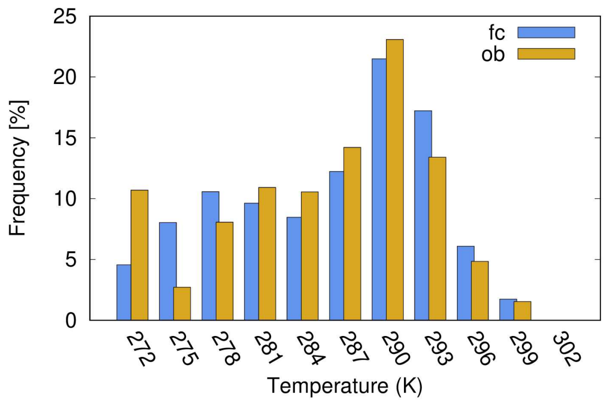
(b) analysis v.s. observation

Figure 6. Frequency of observed (yellow) and forecast or analysed (blue) LSWT over Lake Lappajärvi 2012-2018, all temperatures included. x-axis: LSWT, unit K, y-axis: frequency, unit %.

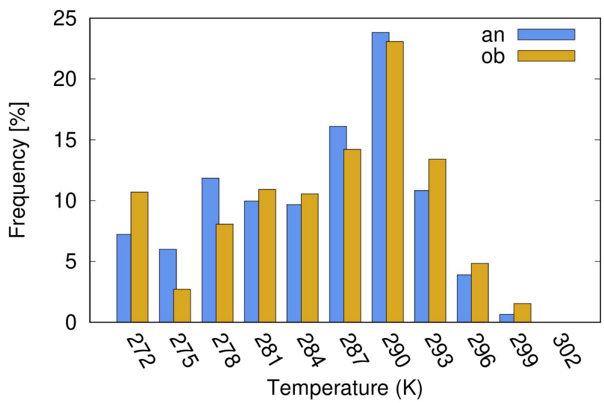
Figure 10 shows a comparison of forecast and observed evolution of ice thickness and snow depth on Lappajärvi, Kilpisjärvi and Simpelejärvi in winter 2012-2013, typical also for the other lakes and years studied. ~~In~~ The most striking feature is that there was no snow in the HIRLAM forecast.

On all three lakes, the ice thickness started to grow after ~~freezing~~ freeze-up both according to the forecast and the observations. In the beginning HIRLAM ~~/FLake~~ ice grew faster than observed. However, according to the forecast ice thickness started to decrease in March of every year but according to the observations only a month or two later. ~~The most remarkable feature is that there was no snow in the FLake forecast. It was found that this was due to a coding error in the HIRLAM reference version 7.4 which is applied operationally in FMI.~~

The too early ~~melting of~~ break-up of lake ice in the absence of snow could be explained by the wrong absorption of the solar energy in the model. In reality, the main factor of snow and ice melt in spring is the increase of daily solar radiation. In



(a) forecast v.s. observation



(b) analysis v.s. observation

Figure 7. As for Figure 6 but for Lake Kilpisjärvi.

HIRLAM, the downwelling short-wave irradiance at the surface is known to be reasonable, with some overestimation of the largest clear-sky fluxes and all cloudy fluxes (Rontu et al., 2017). Over lakes, HIRLAM ~~/FLake~~ uses constant values for the snow and ice shortwave reflection, with albedo values of 0.75 and 0.5, correspondingly. When there was no snow, the lake surface was thus assumed too dark. 25 % more absorption of an assumed maximum solar irradiance of 500 Wm^{-2} (valid for the latitude of Lappajärvi in the end of March) would mean availability of extra 125 Wm^{-2} for melting of the ice, which corresponds the magnitude of increase of available maximum solar energy within a month at the same latitude.

The forecast of too thick ice can also be explained by the absence of snow in the model. When there is no insulation by the snow layer, the longwave cooling of the ice surface in clear-sky conditions is more intensive and leads to faster growth of ice compared to the situation of snow-covered ice. In nature, ice growth can also be due to the snow transformation, a process whose parametrization in the models is demanding (Yang et al., 2013; Cheng et al., 2014).

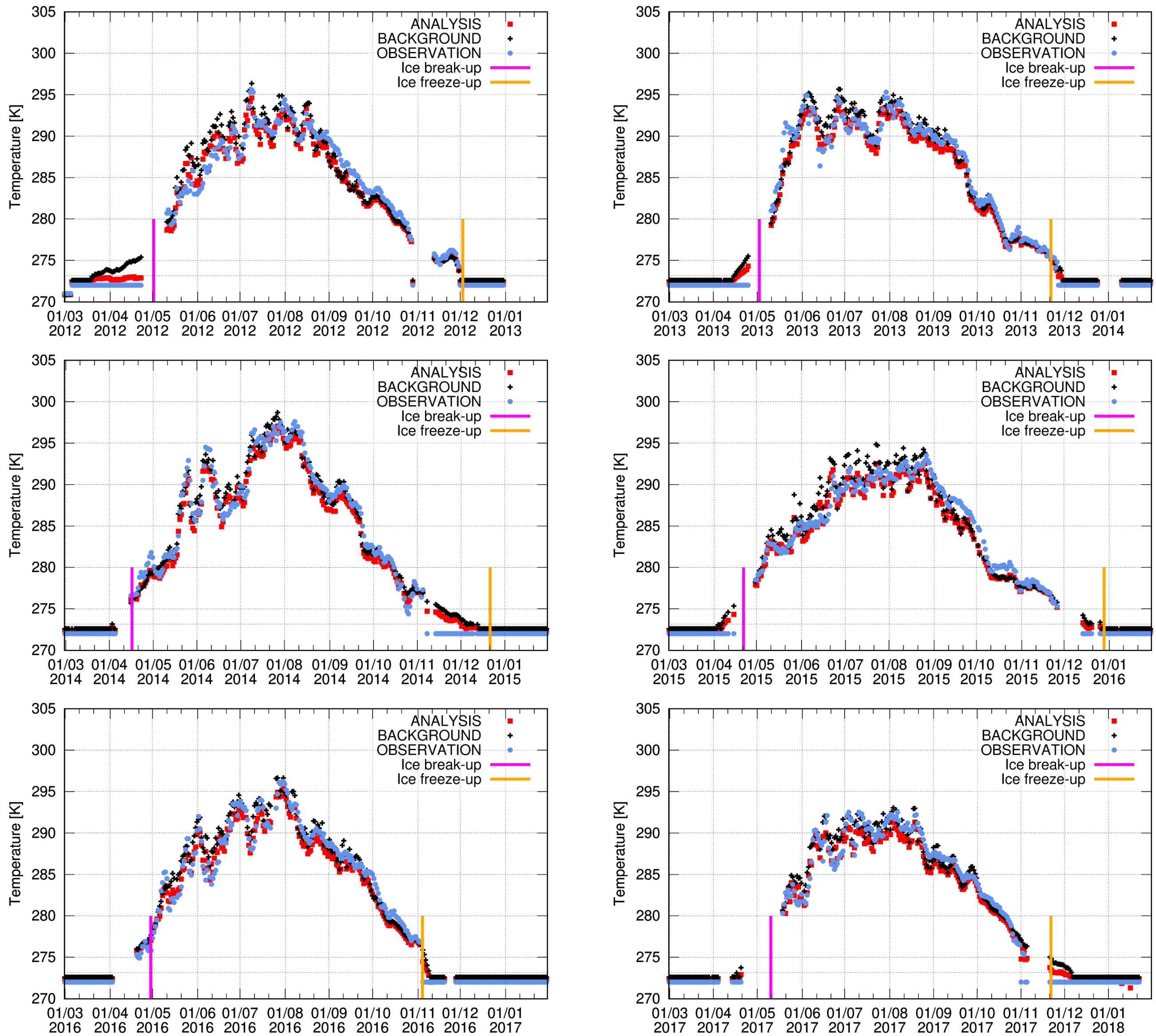


Figure 8. Time-series of the observed, analysed and forecast LSWT at the Lappajärvi observation location 23.67 E, 63.15 N for the years 2012-2018 based on 06 UTC data. Markers are shown in the inserted legend. Observed ~~freezing~~ freeze-up date (blue) and ~~melting~~ break-up date (red) are marked with vertical lines.

Also the downwelling longwave radiation plays a role in the surface energy balance. We may expect values from 150 Wm^{-2} to 400 Wm^{-2} in the Nordic spring conditions, with the largest values related to cloudy and the smallest to clear-sky situations. The standard deviation of the predicted by HIRLAM downwelling longwave radiation fluxes has been shown to be of the order

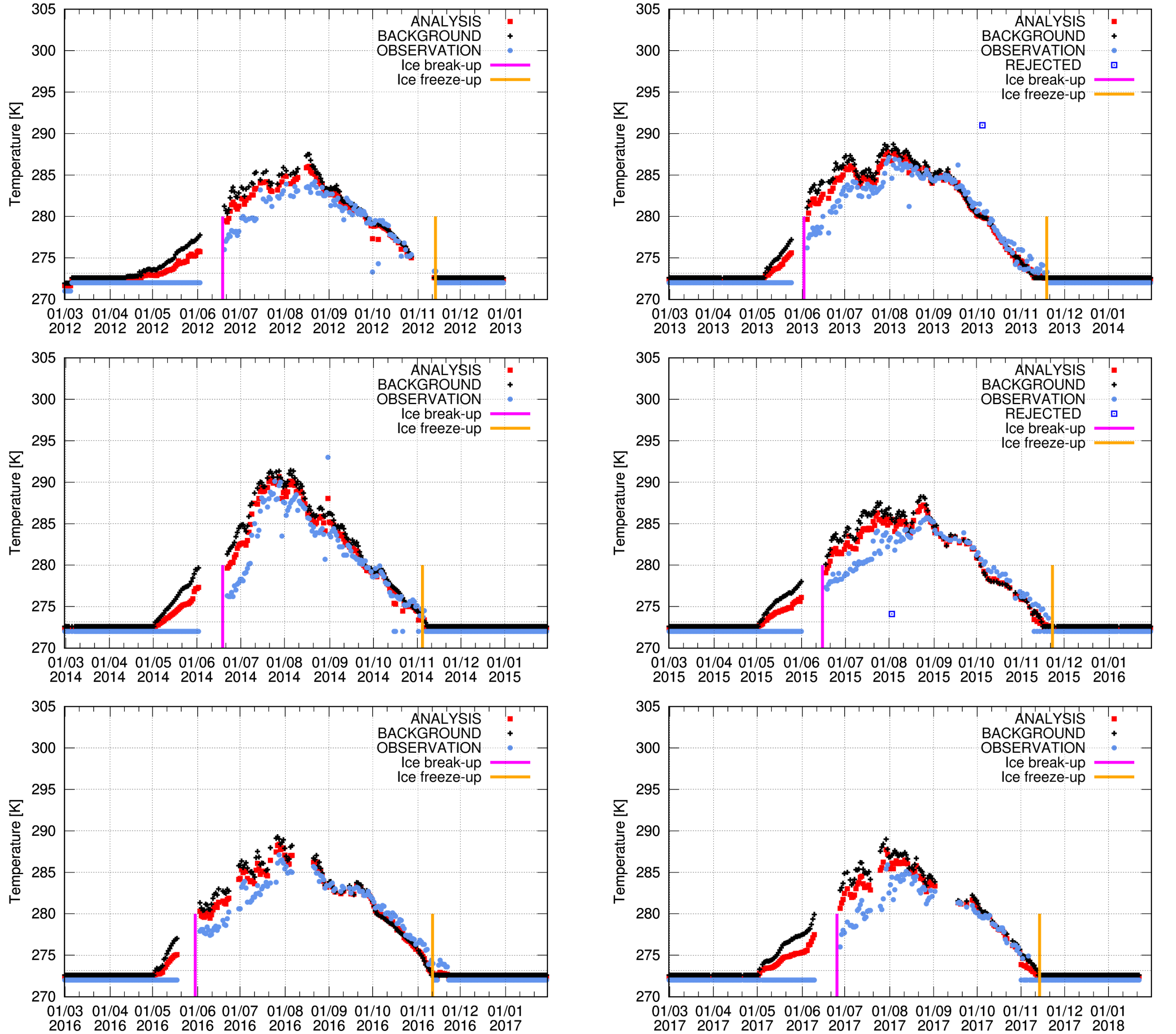


Figure 9. As for Figure 8 but for lake Kilpisjärvi, 20.82 E, 69.01 N.

of 20 Wm^{-2}, with a positive systematic error of a few Wm^{-2} (Rontu et al., 2017). Compared to the systematic effects related to absorption of the solar radiation, the impact of the longwave radiation variations on lake ice evolution is presumably small.

5 Discussion: snow on lake ice

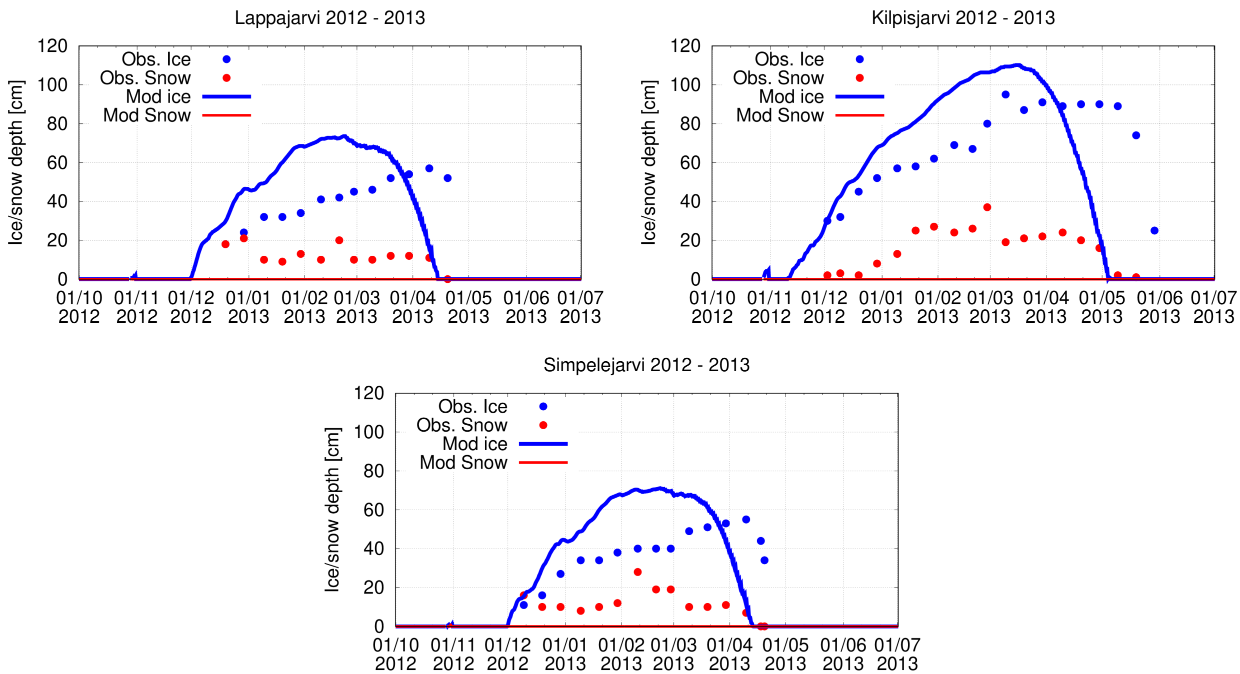


Figure 10. Evolution of ice (blue) and snow (red) thickness at Lakes Lappajärvi, Kilpisjärvi and Simpelejärvi during winter 2012-2013.

The most striking result reported in Section 4 was the too early melting of the lake ice predicted by FLake in HIRLAM as compared to observations. We suggested that the early break-up is related to the missing snow on lake ice in HIRLAM. It was detected that a too large critical value to diagnose snow existence prevented practically all accumulation of the forecast snowfall on lake ice in the reference HIRLAM v.7.4, used operationally at FMI.

In general, handling of the snow cover on lake and sea ice is a demanding task for the NWP models. In HIRLAM, snow depth observations are included into the objective analysis over the land areas, but not over ice where no observations are widely available in real time. Snow depth and temperature over land are treated prognostically using dedicated parametrizations (in HIRLAM, similar to Samuelsson et al., 2006, 2011, see also Boone et al., 2017). Over the sea, a simple prognostic parametrization of sea ice temperature is applied in HIRLAM but neither the thickness of ice nor the depth or temperature of snow on ice are included (Samuelsson et al., 2006). Batrak et al. (2018) provide a useful review and references concerning prognostic sea ice schemes and their snow treatment in NWP models. An essential difference between the simple sea ice scheme and the lake ice scheme applied in HIRLAM is that the former relies on external data on the existence of sea ice cover, provided by the objective analysis, while the latter includes prognostic treatment of the lake water body also. This means that the lake ice freezes and melts in the model depending on the thermal conditions of lake water, evolving throughout the seasons.

The ice thickness, snow depth and ice and snow temperatures are prognostic variables of FLake. When the FLake parametrizations were introduced into HIRLAM (Kourzeneva et al., 2008; Eerola et al., 2010), parametrization of the snow thickness and snow temperature was first excluded. In the COSMO NWP model, snow is implicitly accounted for by modifying ice albedo using empirical data on its temperature dependence (Mironov et al., 2010). This way was applied also e.g. in a recent study over the Great Lakes (Baijnath-Rodino and Duguay, 2019).

Semmler et al. (2012) performed a detailed winter-time comparison between FLake and a more complex snow and ice thermodynamic model (HIGHTSI) on a small lake in Alaska. FLake includes only one ice and one soil layer, while HIGHTSI represents a more advanced multilayer scheme. Atmospheric forcing for the stand-alone experiments was provided by HIRLAM. Based on their sensitivity studies, Semmler et al. (2012) suggested three simplifications to the original, time-dependent snow-on-ice parametrizations of FLake: use a prescribed constant snow density, modify the value of the prescribed molecular heat conductivity and use prescribed constant albedos of dry snow and ice. Later, a similar comparison was performed over Lake Kilpisjärvi (Yang et al., 2013), confirming the improvements due to the updated snow parametrizations in FLake. Implementation of these modifications allowed to include the parametrization of snow on lake ice also into HIRLAM (Section 2.1).

In FLake, snow on lake ice is accumulated from the predicted snowfall. Snow melt on lake ice is related to snow and ice temperatures. In case of FLake integrated into HIRLAM, accumulation and melt are updated at every time step of the advancing forecast. Very small amounts of snow are considered to fall beyond the accuracy of parametrizations and removed. This is controlled by a critical limit, which was set too large (one millimeter instead of ten micrometers) in HIRLAM v.7.4. Due to the incorrect critical value, practically no snow accumulated on lake ice in the FMI operational HIRLAM, validated in this study. In a HIRLAM test experiment, where the original smaller value was used, up to 17 cm of snow accumulated on lake ice within a month (Janurary 2012, not shown).

6 Conclusions and outlook

In this study, *in-situ* lake observations from the Finnish Environment Institute were used for validation of the HIRLAM NWP model, which is applied operationally in the Finnish Meteorological Institute. HIRLAM contains Freshwater Lake prognostic parametrizations and an independent objective analysis of lake surface state. We focused on comparison of observed and forecast lake surface water temperature, ice thickness and snow depth in the years 2012 - 2018. Because the HIRLAM ~~/FLake~~ system was unmodified during this period, a long uniform dataset was available for evaluation of the performance of FLake integrated into an operational NWP model. On the other hand, no conclusions about the impact of the lake surface state on the operational forecast of the near-surface temperatures, cloudiness or precipitation can be drawn because of the lack of alternative (without FLake) forecasts for comparison.

On average, the forecast and analysed LSWT were warmer than observed with systematic errors of 0.91 K and 0.35 K, correspondingly. The mean absolute errors were 1.94 and 1.32 K. Thus, the independent observation-based analysis of *in-situ* LSWT observations was able to improve the FLake +6 h forecast used as the first guess. However, the resulting analysis is by definition not used for correction of the FLake forecast but remains an independent by-product of HIRLAM. An overestimation of the FLake LSWT summer maxima was found, especially for the shallow lakes. This behaviour of FLake is well known, documented earlier e.g. by ~~Kourzeneva, 2014~~Kourzeneva (2014). It arises due to the difficulty to handle correctly the mixing in the near-surface water layer that is intensively heated by the sun.

Forecast ~~freezing~~ freeze-up dates were found to correspond the observations well, typically within a week. The forecast ice thickness tended to be overestimated, still the ~~melting~~ break-up dates over most of the lakes occured systematically several weeks too early. Practically no forecast snow was on found on the lake ice, although the snow parametrization by FLake was included in HIRLAM. The reason for the incorrect behaviour was evidently related to a ~~coding error in HIRLAM that prevented snow accumulation~~ too large critical value to diagnose snow existence that prevented the accumulation of snow on lake ice. The too early melting and overestimated ice thickness differ from the results by ~~Pietikäinen et al., 2018; Yang et al., 2013; Kourzeneva, 2014~~ Pietikäinen et al. (2018); Yang et al. (2013); Kourzeneva (2014), who reported somewhat too late melting of the Finnish lakes when FLake with realistic snow parametrizations was applied within a climate model or stand-alone driven by NWP data. It can be concluded that a ~~realiste~~ realistic parametrization of snow on lake ice is ~~importart~~ important in order to describe correctly the lake surface state in spring.

Small lakes and those of complicated geometry cause problems for the relatively coarse HIRLAM grid of 7 - kilometre resolution. The problems are related to the observation usage, forecast and validation, especially when interpolation and selection of point values are applied. The observations and model represent different spatial scales. For example, the comparison of the ~~freezing and melting~~ freeze-up and break-up dates was based on diagnostics of single-gridpoint values that were compared to observations ~~representing~~ which represent entire lakes as overseen from the observation sites. Also the results of LID diagnostics were sensitive to the criteria for definition of the ice existence in HIRLAM~~/FLake~~. All this adds unavoidable inaccuracy into the model-observation intercomparison but does not change the main conclusions of the present study.

SYKE LSWT observations used for the real-time analysis are regular and reliable but do not always cover the days immediately after ~~melting~~ break-up or close to ~~freezing~~freeze-up, partly because the quality control of HIRLAM LSWT analysis utilizes the SYKE statistical lake water temperature model results in a ~~too~~ strict way. Although the 27 observations are located all over the country, they cover a very small part of the lakes and their availability is limited to Finland. SYKE observations of the ice and snow depth as well as the ~~freezing and melting~~ freeze-up and break-up dates provide valuable data for the validation purposes.

A need for minor technical corrections in the FMI HIRLAM ~~/FLake~~ system was revealed. The ~~snow accumulation bug was corrected in October 2018,~~ coefficient influencing snow accumulation on lake ice was corrected based on our findings. Further developments and modifications are not foreseen because the HIRLAM NWP systems, applied in the European weather services, are being replaced by kilometre-scale ALADIN-HIRLAM forecasting systems (Termonia et al., 2018; Bengtsson et al., 2017), where the prognostic FLake parametrizations are also available. HARMONIE/FLake uses the newest version of the global lake database (GLDB v.3) and contains updated snow and ice properties~~that were suggested by Yang et al., 2013~~. The objective analysis of lake surface state is yet to be implemented, taking into account the HIRLAM experience summarized in this study and earlier by ~~Kheyrollah Pour et al., 2017~~Kheyrollah Pour et al. (2017). In the future, an important source of wider observational information on lake surface state are the satellite measurements, whose operational application in NWP models still requires further work.

Code and data availability. Observational data was obtained from SYKE open data archive SYKE, 2018 as follows: 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/. A supplementary file containing the freeze-up and break-up dates as picked and prepared for the lakes studied here is attached. Data picked from HIRLAM archive are attached as supplementary files: data from the objective analysis feedback files (observed, analysed, forecast LSWT interpolated to the 27 active station locations) and from the gridded output of the HIRLAM analysis (analysed LSWT, forecast ice and snow thickness from the nearest gridpoint of all locations used in the present study).

In this study, FMI operational weather forecasts resulting from use of HIRLAM v.7.4 (rc1, with local updates) were validated against lake observations. The HIRLAM reference code is not open software but the property of the international HIRLAM-C programme. For research purposes, the codes can be requested from the programme (hirlam.org). The source codes of the version operational at FMI, relevant for the present study, are available from the authors upon request.

Author contributions. Laura Rontu computed the LSWT statistics based on HIRLAM feedback files. Kalle Eerola performed the freeze-up and break-up date, snow and ice thickness comparisons based on data picked from HIRLAM grib files. Matti Horttanainen prepared observation data obtained via SYKE open data interface and lake depths from GLDB v.3. Laura Rontu composed the manuscript text based on input from all authors.

Competing interests. No competing interests are present.

Acknowledgements. Our thanks are due to Joni-Pekka Pietikäinen and Ekaterina Kourzeneva for discussions and information, to Margarita Choulga and Olga Toptunova for the support with the updated GLDB v.3 data, and to Emily Gleeson for advice with English language. The comments of three anonymous reviewers and the editor significantly helped to improve the presentation of our results.

References

Baijnath-Rodino, J. and Duguay, C.: Assessment of coupled CRCM5–FLake on the reproduction of wintertime lake-induced precipitation in the Great Lakes Basin. Theoretical and Applied Climatology https://doi.org/10.1007/s00704-019-02799-8, 2019.

Balsamo, G., Salgado, R., Dutra, E., Boussetta, S., Stockdale, T. and Potes, M: On the contribution of lakes in predicting near-surface temperature in a global weather forecasting model. Tellus A, 64, 15829, https://doi.org/10.3402/tellusa.v64i0.15829, 2012.

Batrak, Y., Kourzeneva, E. and Homleid, M.: Implementation of a simple thermodynamic sea ice scheme, SICE version 1.0-38h1, within the ALADIN–HIRLAM numerical weather prediction system version 38h1. Geosci. Model Dev., 11, 3347–3368, https://doi.org/10.5194/gmd-11-3347-2018, 2018.

Bengtsson, L., Andrae, U., Aspelien, T., Batrak, Y., Calvo, J., de Rooy, W., Gleeson, E., Hansen-Sass, B., Homleid, M., Hortal, M., Ivarsson, K., Lenderink, G., Niemelä, S., Pagh Nielsen, K., Onvlee, J., Rontu, L., Samuelsson, P., Santos Muñoz, D., Subias, A., Tijm, S., Toll, V., Yang, X. and Ødegaard Køltzow, M.: The HARMONIE-AROME model configuration in the ALADIN-HIRLAM NWP system. Monthly Wea. Rev., 145, 1919–1935, https://doi.org/10.1175/MWR-D-16-0417.1, 2018.

. Boone, A., Samuelsson, P., Gollvik, S., Napoly, A., Jarlan, L., Brun, E. and Decharme, B.: The interactions between soil–biosphere–atmosphere land surface model with a multi-energy balance (ISBA-MEB) option in SURFEXv8 – Part 1: Model description. Geosci. Model Dev., 10, 843–872. https://doi.org/10.5194/gmd-10-843-2017, 2017.

Cheng, B, Vihma. T, Rontu, L., Kontu A., Kheyrollah Pour, H., Duguay, C. and Pulliainen, J.: Evolution of snow and ice temperature, thickness and energy balance in Lake Orajärvi, northern Finland. Tellus A: Dynamic Meteorology and Oceanography ,66, 21564, https://doi.org/10.3402/tellusa.v66.21564, 2014.

Choulga, M., Kourzeneva, E., Zakharova, E. and Doganovsky, A.: Estimation of the mean depth of boreal lakes for use in numerical weather prediction and climate modelling. Tellus A, 66, 21295, https://doi.org/10.3402/tellusa.v66.21295, 2014.

Cordeira, J. M. and Laird, N. F. 2008. The influence of ice cover on two lake-effect snow events over lake Erie. *Mon. Wea. Rev.* **136**, 2747–2763. https://doi.org/DOI: 10.1175/2007MWR2310.1.

Elo, A-R.: Effects of climate and morphology on temperature conditions of lakes. University of Helsinki, Division of Geophysics, Report series in Geophysics. Available at http://urn.fi/URN:ISBN:978-952-10-3745-0, 2007.

Eerola, K., Rontu, L., Kourzeneva, E. and Shcherbak, E.: A study on effects of lake temperature and ice cover in HIRLAM. Boreal Env. Res., 15, 130–142, 2010.

Eerola, K.: Twenty-one years of verification from the HIRLAM NWP system. Wea. Forecasting, 28, 270–285. https://doi.org/10.1175/WAF-D-12-00068.1, 2013.

Eerola, K., Rontu, L., Kourzeneva, E., Kheyrollah Pour, H. and Duguay, C.: Impact of partly ice-free Lake Ladoga on temperature and cloudiness in an anticyclonic winter situation-a case study using a limited area model. Tellus A, 66, 23929. https://doi.org/10.3402/tellusa.v66.23929, 2014.

Gandin, L.: Objective analysis of meteorological fields. Gidrometizdat, Leningrad. Translated from Russian, Jerusalem, Israel Program for Scientific Translations, 1965.

Kheyrollah Pour, H., Rontu, L., Duguay, C. R., Eerola, K. and Kourzeneva, E.: Impact of satellite-based lake surface observations on the initial state of HIRLAM. Part II: Analysis of lake surface temperature and ice cover. Tellus A., 66, 21395. https://doi.org/10.3402/tellusa.v66.21395, 2014.

Kheyrollah Pour, H., Choulga, M., Eerola, K., Kourzeneva, E., Rontu, L., Pan F. and Duguay C. R.: Towards improved objective analysis of lake surface water temperature in a NWP model: preliminary assessment of statistical properties. Tellus A., 66, 21534. https://doi.org/10.1080/16000870.2017.1313025, 2017.

Korhonen J.: Suomen vesistöjen jääolot. Suomen ympäristö, (751), 145 pp(in Finnish). Available at , 2005.

Korhonen J.: Suomen vesistöjen lämpötilaolot 1900-luvulla. Suomen ympäristö, (566), 116 pp (in Finnish)Long-term changes and variability of the winter and spring season hydrological regime in Finland. Report series in Geophysics, No. 79, 83 pp. University of Helsinki, Faculty of science, Institute for atmospheric and earth system research. Available at, 2002. : https://helda.helsinki.fi/bitstream/handle/10138/298308/longterm.pdf, 2019.

Kourzeneva, E., Samuelsson, P., Ganbat, G. and Mironov, D.: Implementation of lake model Flake into HIRLAM. HIRLAM Newsletter, (54), 54–64. Available at http://hirlam.org/, 2008.

Kourzeneva, E., Asensio, H., Martin, E. and Faroux, S.: Global gridded dataset of lake coverage and lake depth for use in numerical weather prediction and climate modelling. Tellus A., 64,15640. https://doi.org/10.3402/tellusa.v64i0.15640, 2012a.

Kourzeneva, E., Martin, E., Batrak, Y. and Moigne, P. L.:. Climate data for parameterisation of lakes in numerical weather prediction models. Tellus A., 64, 17226. https://doi.org/10.3402/tellusa.v64i0.17226, 2012b.

Kourzeneva, E.: Assimilation of lake water surface temperature observations with Extended Kalman filter. Tellus A., 66, 21510. https://doi.org/10.3402/tellusa.v66.21510, 2014.

Laird, N. F., Kristovich, D. A. R. and Walsh, J. E.: Idealized model simulations examining the mesoscale structure of winter lake-effect circulations. *Mon. Wea. Rev.* **131**, 206–221. https://doi.org/DOI: 10.1175/1520-0493(2003)131<0206:IMSETM>2.0.CO;2, 2003.

Leppäranta, M., Lindgren, E. and Shirasawa, K.: The heat budget of Lake Kilpisjarvi in the Arctic tundra. Hydrology research, 48, 969-980. https://doi.org/10.2166/nh.2016.171, 2012.

Lei, R., Leppäranta, M., Cheng, B., Heil P. and Li, Z.: Changes in ice-season characteristics of a European Arctic lake from 1964 to 2008. Climatic Change, 115, 725–739. https://doi.org/10.1007/s10584-012-0489-2, 2012.

Lindstedt, D., P. Lind, E. Kjellström and C. Jones.: A new regional climate model operating at the meso-gamma scale: performance over Europe. Tellus A, 67, 24138. https://doi.org/10.3402/tellusa.v67.24138, 2015.

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, J. and co authors: Development of a global land cover characteristics database and IGBP DISCover from 1-km AVHRR data. Int. J. Rem. Sens., 21, 1303–1130. Data and documentation available online at: http://edc2.usgs.gov/glcc/glcc.php, 2000.

Masson, V., P. Le Moigne, E. Martin, S. Faroux, S., Alias, A., Alkama, R., Belamari, S., Barbu, A., Boone, A., Bouyssel, F., Brousseau, P., Brun, E., Calvet, J.-C., Carrer, D., Decharme, B., Delire, C., Donier, S., Essaouini, K., Gibelin, A.-L., Giordani, H., Habets, F., Jidane, M., Kerdraon, G., Kourzeneva, E., Lafaysse, M., Lafont, S., Lebeaupin Brossier, C., Lemonsu, A., Mahfouf, J.-F., Marguinaud, P., Mokhtari, M., Morin, S., Pigeon, G., Salgado, R., Seity, Y., Taillefer, F., Tanguy, G., Tulet, P., Vincendon, B., Vionnet, V., and Voldoire, A.: The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes, Geosci. Model Dev., 6, 929–960. https://doi.org/:10.5194/gmd-6-929-2013, 2013.

Mironov, D., Heise, E., Kourzeneva, E., Ritter, B., Schneider, N. and Terzhevik, A.: Implementation of the lake parameterisation scheme FLake into the numerical weather prediction model COSMO. Boreal Env. Res., 15, 218–230, 2010.

Pietikäinen, J.-P., Markkanen, T., Sieck, K., Jacob, D., Korhonen, J., Räisänen, P., Gao, Y., Ahola, J., Korhonen, H., Laaksonen, A., and Kaurola, J.: The regional climate model REMO (v2015) coupled with the 1-D freshwater lake model FLake (v1): Fenno-Scandinavian climate and lakes. Geosci. Model Dev., 11, 1321–1342, https://doi.org/10.5194/gmd-11-1321-2018, 2018.

Rontu, L., Eerola, K., Kourzeneva, E. and Vehviläinen, B.: Data assimilation and parametrisation of lakes in HIRLAM. Tellus A., 64, 17611. https://doi.org/10.3402/tellusa.v64i0.17611, 2012.

Rontu L., Gleeson, E., Räisänen, P., Nielsen, K.P., Savijärvi, H. and Sass, B.H.: The HIRLAM fast radiation scheme for mesoscale numerical weather prediction models. Adv.Sci.Res., 14, 195–215, https://doi.org/10.5194/asr-14-195-2017, 2017.

Rooney, G. G. and Bornemann, F. J.: The performance of FLake in the Met Office Unified Model. Tellus A, 65, 21363, https://doi.org/10.3402/tellusa.v65i0.21363, 2013.

. Samuelsson P., Gollvik, S., and Ullerstig, A.: The land-surface scheme of the Rossby Centre regional atmospheric climate model (RCA3). Report in Meteorology 122, SMHI, SE-60176 Norrköping, Sweden, 2006.

Samuelsson, P., Kourzeneva, E. and Mironov, D.: The impact of lakes on the European climate as stimulated by a regional climate model. Boreal Env. Res., 15, 113–129, 2010.

Samuelsson, P., Jones, C., Willén, U., Ullerstig, A., Gollvik, S., Hansson, U., Jansson, C., Kjellström, E., Nikulin, G., and Wyser, K.: The Rossby Centre Regional Climate Model RCA3: Model description and performance, Tellus A, 63, 1–3, https://doi.org/10.1111/j.1600-0870.2010.00478.x, 2011.

Semmler, T., B. Cheng, Y. Yang, and L. Rontu. Snow and ice on Bear Lake (Alaska) – sensitivity experiments with two lake ice models. Tellus A, 64, 17339. https://doi.org/10.3402/tellusa.v64i0.17339, 2012.

SYKE. ISO 19115 Geographic Information Metadata on Hydrological data. Available at http://metatieto.ymparisto.fi:8080/geoportal/catalog/search/resource/details.page?uuid=%7B86FC3188-6796-4C79-AC58-8DBC7B568827%7D, 2018.

Termonia, P., Fischer, C., Bazile, E., Bouyssel, F., Brožková, R., Bénard, P., Bochenek, B., Degrauwe, D., Derkova, M., El Khatib, R., Hamdi, R., Mašek, J., Pottier, P., Pristov, N., Seity, Y., Smolíková, P., Spaniel, O., Tudor, M., Wang, Y., Wittmann, C., and Joly, A.: The ALADIN System and its canonical model configurations AROME CY41T1 and ALARO CY40T1, Geosci. Model Dev., 11, 257-281, https://doi.org/10.5194/gmd-11-257-2018, 2018.

Undén, P., Rontu, L., Järvinen, H., Lynch, P., Calvo, J., Cats, G., Cuxart, J., Eerola, K., Fortelius, C., Garcia-Moya, J. A. et al.: HIRLAM-5 scientific documentation. Available at http://www.hirlam.org, 2002.

Yang, Y., Cheng, B., Kourzeneva, E., Semmler, T., Rontu, L., Leppäranta, M., Shirasawa, K. and Li, Z. J.: Modelling experiments on air–snow–ice interactions over Kilpisjärvi, a lake in northern Finland. Boreal Env. Res. 18: 341–358, 2013.

Table A1. Prognostic and diagnostic lake variables within HIRLAM

variable	unit	type
temperature of snow on lake ice	K	prog by FLake
temperature of lake ice	K	prog by FLake
mean water temperature	K	prog by FLake
mixed layer temperature	K	prog by FLake
bottom temperature	K	prog by FLake
temperature of upper layer sediments	K	prog by FLake
mixed layer depth	m	prog by FLake
thickness of upper layer sediments	m	prog by FLake
thermocline shape factor	-	prog by FLake
lake ice thickness	m	prog by FLake
snow depth on lake ice	m	prog by FLake
LSWT	K	diag by FLake
		= mixed layer temperature if no ice
lake surface temperature	K	diag by FLake
		uppermost temperature: LSWT or ice or snow
LSWT	K	anal by HIRLAM
		flag value 272 K when there is ice
fraction of lake ice	[0-1]	diag fraction in HIRLAM grid
lake surface roughness	m	diag by HIRLAM
screen level temperature over lake	K	diag by HIRLAM
screen level abs.humidity over lake	$kgkg^{-1}$	diag by HIRLAM
anemometer level u-component over lake	ms^{-1}	diag by HIRLAM
anemometer level v-component over lake	ms^{-1}	diag by HIRLAM
latent heat flux over lake	Wm^{-2}	diag by HIRLAM
sensible heat flux over lake	Wm^{-2}	diag by HIRLAM
scalar momentum flux over lake	Pa	diag by HIRLAM
SW net radiation over lake	Wm^{-2}	diag by HIRLAM
LW net radiation over lake	Wm^{-2}	diag by HIRLAM
depth of lake	m	pres in HIRLAM grid
fraction of lake	[0-1]	pres in HIRLAM grid

Denotation: prog = prognostic, diag = diagnostic, pres = prescribed, anal = result of OI

Table A2. Lakes with SYKE observations used in this study.

NAME	LON	LAT	MEAND (m)	MAXD (m)	AREA (kgm^{-2})	HIRD (m)	HIRFR	HIRID
Pielinen	29.607	63.271	10.1	61.0	894.2	10.0	0.916	4001
Kallavesi	27.783	62.762	9.7	75.0	316.1	10.0	0.814	4002
Haukivesi	28.389	62.108	9.1	55.0	560.4	10.0	0.725	4003
Saimaa	28.116	61.338	10.8	85.8	1,377.0	10.0	0.950	4004
Pääjärvi1	24.789	62.864	3.8	14.9	29.5	3.0	0.430	4005
Nilakka	26.527	63.115	4.9	21.7	169.0	10.0	0.866	4006
Konnevesi	26.605	62.633	10.6	57.1	189.2	10.0	0.937	4007
Jääsjärvi	26.135	61.631	4.6	28.2	81.1	10.0	0.750	4008
Päijänne	25.482	61.614	14.1	86.0	864.9	10.0	0.983	4009
Ala-Rieveli	26.172	61.303	11.3	46.9	13.0	10.0	0.549	4010
Kyyvesi	27.080	61.999	4.4	35.3	130.0	10.0	0.810	4011
Tuusulanjärvi	25.054	60.441	3.2	9.8	5.9	3.0	0.174	4012
Pyhäjärvi	22.291	61.001	5.5	26.2	155.2	5.0	0.922	4013
Längelmävesi	24.370	61.535	6.8	59.3	133.0	10.0	0.875	4014
Pääjärvi2	25.132	61.064	14.8	85.0	13.4	14.0	0.350	4015
Vaskivesi	23.764	62.142	7.0	62.0	46.1	10.0	0.349	4016
Kuivajärvi	23.860	60.786	2.2	9.9	8.2	10.0	0.419	4017
Näsijärvi	23.750	61.632	14.7	65.6	210.6	10.0	0.850	4018
Lappajärvi	23.671	63.148	6.9	36.0	145.5	10.0	1.000	4019
Pesiöjärvi	28.650	64.945	3.9	15.8	12.7	7.0	0.290	4020
Rehja-Nuasjärvi	28.016	64.184	8.5	42.0	96.4	10.0	0.534	4021
Oulujärvi	26.965	64.451	6.9	35.0	887.1	10.0	1.000	4022
Ounasjärvi	23.602	68.377	6.6	31.0	6.9	10.0	0.166	4023
Unari	25.711	67.172	5.0	24.8	29.1	10.0	0.491	4024
Kilpisjärvi	20.816	69.007	19.5	57.0	37.3	22.0	0.399	4025
Kevojärvi	27.011	69.754	11.1	35.0	1.0	10.0	0.016	4026
Inarijärvi	27.924	69.082	14.3	92.0	1,039.4	14.0	0.979	4027

Denotation: LON and LAT are the longitude E and latitude N in degrees, MEAND and MAXD are the mean and maximum depths and AREA is the water surface area from the updated lake list of GLDB v.3 (Margarita Choulga, personal communication), HIRD and HIRFR are the mean lake depth and fraction of lakes [0...1] interpolated to the selected HIRLAM gridpoint, taken from the operational HIRLAM that uses GLDB v.2 as the source for lake depths. HIRID is the lake index used by HIRLAM and in this study. Above the middle line are the 27 lakes with both LSWT and LID observations, below the 18 lakes where only LID was available.

Table A3. Lakes with SYKE observations used in this study. Part 2

NAME	LON	LAT	MEAND (m)	MAXD (m)	AREA (kgm^{-2})	HIRD (m)	HIRFR	HIRID
Simpelejärvi	29.482	61.601	9.3	34.4	88.2	10.0	0.548	40241
Pökkäänlahti	27.264	61.501	8.0	84.3	58.0	10.0	0.299	40261
Muurasjärvi	25.353	63.478	9.0	35.7	21.1	10.0	0.060	40263
Kalmarinselkä	25.001	62.786	5.7	21.9	7.1	5.0	0.330	40271
Summasjärvi	25.344	62.677	6.7	40.5	21.9	10.0	0.555	40272
Iisvesi	27.021	62.679	17.2	34.5	164.9	18.0	0.456	40277
Hankavesi	26.826	62.614	7.0	49.0	18.2	18.0	0.100	40278
Petajävesi	25.173	62.255	4.2	26.6	8.8	3.0	0.245	40282
Kukkia	24.618	61.329	5.2	35.6	43.9	10.0	0.299	40308
Ähtärinjärvi	24.045	62.755	5.2	27.0	39.9	10.0	0.266	40313
Kuortaneenjärvi	23.407	62.863	3.3	16.2	14.9	10.0	0.277	40328
Lestijärvi	24.716	63.584	3.6	6.9	64.7	10.0	0.513	40330
Pyhäjärvi	25.995	63.682	6.3	27.0	121.8	10.0	0.266	40331
Lentua	29.690	64.204	7.4	52.0	77.8	7.0	0.600	40335
Lammasjärvi	29.551	64.131	4.3	21.0	46.8	3.0	0.200	40336
Naamankajärvi	28.246	65.104	2.9	14.0	8.5	7.0	0.299	40342
Korvuanjärvi	28.663	65.348	6.0	37.0	15.4	10.0	0.342	40343
Oijärvi	25.930	65.621	1.1	2.4	21.0	10.0	0.333	40345

Denotation: LON and LAT are the longitude E and latitude N in degrees, MEAND and MAXD are the mean and maximum depths and AREA is the water surface area from the updated lake list of GLDB v.3 (Margarita Choulga, personal communication), HIRD and HIRFR are the mean lake depth and fraction of lakes [0. . . 1] interpolated to the selected HIRLAM gridpoint, taken from the operational HIRLAM that uses GLDB v.2 as the source for lake depths. HIRID is the lake index used by HIRLAM and in this study. Above the middle line are the 27 lakes with both LSWT and LID observations, below the 18 lakes where only LID was available.