Deriving an effective lake depth from satellite lake surface temperature: A feasibility study with MODIS data

G. Balsamo[1], E. Dutra[2,3], V.M. Stepanenko[4], P. Viterbo[5], P.M. Miranda[2] and D. Mironov[6]

Research Department

[1] ECMWF, Reading, United Kingdom

[2] CGUL, ILD, Faculdade de Ciencias Lisbon, Portugal

[3] Institute for Atmospheric and Climate Science, ETH, Zurich, Switzerland

[4] Moscow State University, Scientific Research Computing Center; Moscow, Russia

[5] Institute of Meteorology, Lisbon, Portugal

[6] Deutscher Wetterdienst, Offenbach, Germany

December 2009

Abstract

Modelling lakes in Numerical Weather Prediction (NWP) is important to produce accurate evaporation rates and surface temperature forecasts. Lake depth is a crucial external parameter for the implementation of lake models into NWP systems, since it controls the dynamical range of lake temperature amplitudes on diurnal to seasonal time scales. However, a global lake-depth dataset does not exist at present. A novel method to derive an effective lake depth on the basis of the remotely-sensed lake water-surface temperature (LWST) is presented. A technique is proposed to adjust a simple two-layer Fresh-water Lake model (FLake) depth such that simulated annual cycle of LWST matches satellite-based LWST climatology as closely as possible. The method is applied to 47 European lakes and the results show convergence of the solution. Merits and limitations of this approach are discussed. Preliminary validation of a derived bathymetry of the American Great Lakes is presented.

1   Introduction

Lakes are becoming an important component of land surface parameterisation schemes used in Numerical Weather Prediction (NWP) as the horizontal resolution increases. Despite the large contrast between land and inland water in terms of the surface temperature and its evolution on both diurnal and seasonal time-scales, lake models are not yet an integral part of operational NWP.

Lakes need to be adequately represented in NWP systems to accurately predict evaporation rates (Dutra *et al*. 2009, Mironov *et al*. 2009) and the water-ice phase transitions (Duguay *et al*. 2003, Lenters *et al*. 2005). It has been clearly shown that lakes may strongly affect the results of climate simulations (Bonan 1995, Lofgren 1997, Krinner 2003, Long *et al*. 2007).

The evolution of temperature and mixing conditions in lakes is primarily driven by the surface fluxes of heat and momentum. According to sensitivity studies, a crucial parameter, which strongly affects the temperature dynamical range of a lake and its capacity to mix and to freeze, is the lake depth. A good knowledge of the lake depth is therefore of paramount importance for the implementation of lake models into NWP systems. A clear link between bathymetry and surface water temperature cycle has been demonstrated in shallow ocean waters (Xie *et al*. 2002, Park *et al*. 2005). Bathymetry inversion techniques using remote sensing data have been successfully applied (Vrbancich *et al*. 2000). However, no study has tried to exploit the relationship between lake depth and surface temperature for a lake bathymetry inversion. The interest for such a technique originates from the lack of a comprehensive global lake-depth dataset.

A few global datasets available at present (Loveland *et al*. 2000, Masson *et al*. 2003, Lehner and Döll 2004) provide information on the lake cover fraction. A compilation of available regional to continental estimates of lake depth is collected by E. Kourzeneva in a lake-based mean and maximum depth dataset (http://www.flake.igb-berlin.de/ep-data.shtml). This data source is useful, and it is shown in the present study for comparison. However, it provides at best a static estimate of the mean and maximum depth of several lakes but not an actual bathymetry, necessary for representing lakes into grid-point models.

Lake physics is also influenced by other parameters such as lake water optical characteristics and sediment properties. Such information is available only for a very limited number of lakes (e.g. Arst *et al*. 2008). A global data set is not likely to be developed over the next few years. Many aspects of lake dynamics, such as water inflow from tributaries and outflow to effluents, are neglected in simplified lake models used for NWP and related applications. For these reasons, the lake depth obtained by the inversion technique is referred to as an 'effective depth' as it optimizes the simulated lake surface temperature with respect to the observations.

Oesch *et al*. (2005) have demonstrated the capability to retrieve the lake surface temperature from satellites at a reasonably high spatial resolution and with acceptable accuracy (of the order of 1-2 K). The availability of multi-annual satellite-derived lake water-surface temperature (LWST) covering the entire globe at a kilometre resolution allows for further exploitation of the information on the LWST annual cycle in order to obtain an effective lake depth.

In this paper, a novel method to derive an effective lake depth from observed LWST is presented. A simple two-layer Fresh-water Lake model (FLake) is forced with near-surface atmospheric variables in a year-round simulation, repeated for several times in order to obtain a perpetual year solution. Application of this procedure generates a model-based lake climate that can be compared to an observation-based climatology.

A minimisation technique was implemented to adjust the model lake depth to provide the best fit to a MODIS-derived climatology of LWST. This depth is a yearly average (with no account for lake water balance) and at a given location specified by latitude and longitude. It may differ from the observed depth due to errors caused by various simplifying assumptions (e.g. inaccurate estimates of the water transparency and no account for the interaction of the lake water with bottom sediments) and inaccuracy in the observations, forcing and model physics component.

Next sections briefly introduce the lake model FLake and illustrate the methodology and the datasets adopted. Results for 47 European lakes are presented together with a preliminary application of the method to the American Great Lakes region. A discussion on the potential benefits and risks involved in the lake depth inversion is presented. The last section presents the summary and conclusions from the present study and the perspectives for a global effective lake depth dataset.

2   The Fresh-water Lake model (FLake)

FLake is a Fresh-water Lake model (Mironov *et al*. 2005, Mironov 2008, Mironov *et al*. 2009) capable of predicting the vertical temperature structure and mixing conditions in lakes of various depths on time-scales from a few hours to a few years. The model is particularly suitable for NWP and climate modelling due to its low computational cost.

FLake is based on a two-layer parametric representation of the evolving temperature profile and on the integral budgets of heat and kinetic energy. The structure of the stratified layer between the upper mixed layer and the basin bottom, the lake thermocline, is described using the concept of self-similarity (assumed shape) of the temperature-depth curve. The same concept is used to describe the temperature structure of the thermally active upper layer of bottom sediments and of the ice.

FLake incorporates a flexible parameterization of the temperature profile in the thermocline and an advanced formulation to compute the mixed-layer depth, including the equation of convective entrainment and a relaxation-type equation for the depth of a wind-mixed layer. Both mixing regimes are treated with due regard for solar radiation heating. The model incorporates a module to describe the vertical temperature structure of the thermally active layer of bottom sediments and the interaction of the water column with bottom sediments, and a snow-ice module. FLake parameters are estimated using independent empirical and numerical data and are not lake-dependent. Thus FLake does not require re-tuning. With the integral approach used in FLake, the problem of solving partial differential equations (in depth and time) for the

temperature and turbulence quantities is reduced to solving ordinary differential equations for the time-dependent quantities that specify the temperature profile.

These are

- mixed-layer temperature,
- mixed-layer depth,
- bottom temperature (temperature at the water-bottom sediment interface),
- mean temperature of the water column,
- shape factor with respect to the temperature profile in the thermocline,
- temperature at the ice upper surface, and
- ice thickness

There is no water balance equation; the lake depth is kept constant. Provision is made to explicitly account for the layer of snow above the lake ice. Then, prognostic equations are carried for the temperature at the snow upper surface and for the snow thickness. In order to be used as a lake parameterisation module in an NWP or climate modelling system, FLake needs two physiographic fields. These are

- lake cover (fraction of a given grid box of an atmospheric model covered by lake water), and
- lake depth (mean depth of lakes present in a given grid box).

The lake cover is specified by the Global Land Cover Characterization (GLCC) dataset (Loveland *et al*. 2000). A parameterization scheme to compute turbulent fluxes at the lake-atmosphere interface is taken from the ECMWF model which is common to the HTESSEL land surface scheme (Viterbo and Beljaars 1995, van den Hurk *et al*. 2000, Balsamo *et al*. 2009). In the present configuration, the bottom sediment module of FLake is switched off and the heat flux at the water-bottom sediment interface is set to zero.

3   Methodology

The working hypothesis of the lake depth inversion relies on the assumption that, for a given climate and meteorological forcing, the lake depth is the dominant parameter that controls the amplitude of the LWST seasonal cycle. The link between bathymetry and surface temperature cycle has been studied for shallow ocean waters (Xie *et al*. 2002, Park *et al*. 2005) and high correlations of surface temperature patterns and bathymetry were found. The main mechanism can be easily understood by considering fully-mixed late autumn and winter conditions when the rate of cooling is inversely proportional to the lake depth so that shallower lakes cool down faster than deeper ones. Following Park *et al*. (2005), the ratio of sea surface water temperature (SST) change can be expressed as:

$$\Delta \text{SST} \ = \ Q_{\text{net}} \ / \ \left(\rho_{\text{w}} \ C_{\text{p}} \ d \right) \tag{1}$$

where Q_{net} is the net surface heat flux (W m^{-2}), ρ_{w} and C_{p} are, respectively, the water density (kg m^{-3}) and heat capacity (J kg^{-1} K^{-1}), and d is the water depth (m). Assuming that an equal atmospheric forcing is imposed to two adjacent water grid points A and B, the amplitude of the annual cycle of surface temperature will be larger at the shallower point as expressed by the following relationship:

$$\Delta SST^A \ / \ \Delta SST^B \ \approx \ d^B \ / \ d^A \tag{2}$$

Provided that a lake model has sufficient skill in reproducing the observed lake surface temperatures for a given lake depth and that reasonably accurate atmospheric forcing is available, a data assimilation scheme can be used to derive an effective lake depth by minimising the difference between the modelled and the observed LWST. This inversion method can then be applied to any lake for which observed LWST is available.

Similar procedures are known to be used with the Land Data Assimilation Systems (LDAS) which are generally capable of ingesting observations that are only indirectly related to the variables to be optimised. Soil moisture analysis is a typical example. However, this technique has never been applied to lake parameters. In data assimilation problems, the optimal estimate is obtained through the minimization of a cost function, J, as

$$J = J_b + J_o \left(\mathbf{x} \ - \ \mathbf{x_b} \right)^T \mathbf{B}^{-1} \left(\mathbf{x} - \mathbf{x_b} \right) + \left(\mathbf{y_o} - \mathbf{H(x)} \right)^T \mathbf{R}^{-1} \left(\mathbf{y_o} - \mathbf{H(x)} \right) \tag{3}$$

where the subscripts o and b indicate the observation and the model a-priori knowledge (also known as background), respectively. The vectors $\mathbf{x}$ and $\mathbf{y_o}$ contain, respectively, the model control variables and the observations, and the matrices $\mathbf{B}$ and $\mathbf{R}$ are the model a-priori error and the observation error covariance matrices. $H(\mathbf{x})$ is the so-called observation operator, responsible for mapping a model state onto the observation space. A 'well-posed' assimilation problem has an optimal solution only if the number of unknowns (degrees of freedom of the system) is matched by the number of independent observations. In the case at hand, we have observations of the seasonal evolution of the LWST and only one unknown, namely, the lake depth. It is assumed that the LWST seasonal changes in a given climate are dominated by the lake depth; secondary effects due to radiative properties or advection can be neglected. The observed LWST with assigned observational error are assimilated into a minimisation scheme. The lake model is driven offline with near-surface atmospheric forcing, producing a modelled LWST. An effective lake depth for any given lake point is found when a best match of modelled and observed LWSTs is obtained. This can be formulated by analogy with Eq. (3) in terms of a cost function as

$$J = \tfrac{1}{2}\left[d \ - d_b \right]^2 / (s_{b\_d})^2 + \tfrac{1}{2} \ \textstyle\sum_i \left[LWST_o^i - \ LWST_o^i (d) \right]^2 / (s_{o\_LWST})^2 \tag{4}$$

where the control vector $\mathbf{x}=\{d\}$ is the lake depth sought for, and the background vector $\mathbf{x_b}=\{d_b\}$ is an a-priori given lake depth. Observations are represented by the $\mathbf{y_o}=\{LWST^i_o,\}$. σ_{b_d} and σ_{o_LWST} are the background and observation errors, respectively. The MODIS lake surface temperature is available once every eight days from the satellite-based products with the i index going from 1 to n (with $n=42$, the total number of observations over the year). The observation operator $H(\mathbf{x})$ is represented by the lake model FLake that maps a lake depth onto a LWST annual cycle for a given atmospheric forcing, so that $H(\mathbf{x})=\{LWST^i_b(d)\}$. The modelled LWST was obtained by repeating a year-long run several times using the same atmospheric forcing until an equilibrium solution was achieved (this is also called perpetual-year solution). This procedure is assumed to lead to the 'lake climate' for a given forcing and a given lake depth d. Then, the minimisation of the cost function J leads to an optimal value of the lake depth.

In the experiments described below, only J_o is minimised, assuming no prior knowledge of the lake depth d_b. A J_b term may be used by successive application of this methodology, for instance, to obtain a high-

resolution lake-depth dataset, for which a lower resolution lake-depth map can provide an a-priori knowledge term. For this application, the observation errors σ_{o_LWST} would simply be scaling the value of cost function without having any impact on the derived lake depth and hence are neglected. The J_o term is then slightly modified and is taken to be equal to the LWST Mean Absolute Error (MAE):

$$J_o = (1/n) \sum_i \left| LWST_o^i - LWST_b^i(d) \right| \tag{5}$$

This methodology presents some caveats, in particular, if the working hypothesis Eq. 2 is not satisfied due to model inaccuracies, observational bias, or forcing errors. Some risks and benefits are discussed following the results obtained.

3.1 Atmospheric forcing data

The atmospheric forcing necessary to drive the lake model FLake were extracted from the ECMWF Re-Analysis ERA-Interim (Simmons *et al*. 2007) which covers the period from 1989 until present (see http://www.ecmwf.intresearch/era/do/get/era-interim) and is the follow-up reanalysis of ERA-40 (Uppala *et al*. 2005). The atmospheric forcing data are three hours apart at T255 horizontal resolution (about 80 km) globally. They include the surface pressure, the 2-m air temperature and the 2-m specific humidity, and the 10-m wind. The downward fluxes of long-wave and short-wave (solar) radiation and the precipitation flux were also taken from ERA-Interim. Data from the nearest point to the lake in question (the list of lakes is given in Table 1) were extracted as time-series for the year 2002. These data were used to examine the working hypotheses on European lakes in the results. A longer time-series was used for the lake depth inversion on the American Great Lakes.

Table 1: Subset of European lake-depth dataset (developed by E. Kourzeneva) for all lakes with areas larger than 100 km2 and at least 80% coverage by MODIS data. Subset of European lake depth dataset v1.0 by E. Kourzeneva

N	latitude	longitude	d_{mean}	d_{max}	Area*	Name of the lake (or reservoir)
1	63.12	25.67	6.50	66.20	502.00	Nimetön
2	62.80	29.72	11.80	58.60	279.70	Höytiäinen
3	66.16	28.73	4.60	42.00	240.20	Yli-Kitka
4	62.60	26.54	10.80	56.00	186.90	Konnevesi
5	62.97	30.79	10.80	47.40	169.60	Koitere
6	62.82	26.98	17.20	34.50	164.90	Iisvesi
7	63.09	25.16	8.40	45.00	156.20	Kivijärvi
8	60.99	22.32	5.50	26.20	154.00	Pyhäjärvi
9	65.03	29.11	21.50	43.00	152.70	Kiantajärvi (N43199.30)
10	62.72	29.30	6.00	58.30	135.40	Viinijärvi
11	64.13	29.19	14.50	29.00	102.00	Ontojärvi
12	63.30	25.74	8.70	66.00	101.30	Kolima
13	61.98	35.22	30.00	120.00	9690.00	Onega
14	58.12	38.67	5.60	30.40	4550.00	Rybinskoe
15	58.02	28.07	7.10	16.60	3512.00	Chudsko-Pskovskoe
16	63.60	34.60	7.40	20.50	1270.00	Vygozerskoe
17	58.35	31.63	2.70	4.50	1110.00	Ilmen

25

N	latitude	longitude	d_{mean}	d_{max}	Area*	Name of the lake (or reservoir)
18	65.70	32.00	16.00	56.00	986.00	Topozero (Tuoppajärvi)
19	51.00	30.50	4.00	18.40	922.00	Kievskoe
20	67.93	32.97	12.80	67.00	876.00	Imandra
21	63.17	33.93	29.00	103.00	815.00	Segozerskoe
22	65.05	31.57	11.00	36.00	655.00	Yushkozerskoe
23	62.00	34.00	6.00	12.00	547.00	Vygozero
24	62.23	36.88	2.20	18.00	367.00	Vodlozerskoe
25	60.20	37.70	1.80	5.60	351.00	Lacha
26	67.55	34.33	57.50	115.00	313.00	Umbozero
27	61.88	33.32	6.50	24.00	266.00	Syanmozero
28	58.30	26.00	2.50	5.60	259.00	Vyrtsyarv
29	64.95	30.55	8.50	44.00	240.00	Verhneekujto
30	67.98	35.08	5.70	35.00	223.00	Lovozero
31	66.05	31.80	8.00	41.00	209.00	Tikshezero
32	61.00	31.00	46.9	230.00	17800.00	Ladoga
33	62.60	33.75	9.50	58.00	185.00	Sandalskoe
34	63.83	30.82	8.50	34.00	166.00	Leksozero
35	59.70	39.30	1.30	2.00	166.00	Kubenskoe
36	65.78	33.93	4.50	18.00	122.00	Engozero
37	62.52	33.67	18.50	74.00	109.00	Pale (Palzerskoe)
38	53.48	12.44	6.60	28.10	102.70	Fleesensee
39	53.37	-9.09	5.00	46.00	173.80	Corrib
40	52.92	-8.33	7.60	36.00	117.50	Derg
41	60.50	10.36	65.50	131.00	140.10	Randsfjorden
42	53.74	21.75	5.80	23.40	113.40	Sniardwy
43	58.90	13.25	27.00	106.00	5648.00	Vänern
44	58.32	14.56	39.90	128.00	1856.00	Vättern
45	59.52	17.03	11.90	61.00	1140.00	Mälaren
46	59.22	15.77	6.10	22.00	478.00	Hjäelmaren
47	54.60	-6.50	8.90	34.00	385.00	Neagh

Area>100km^2 only and >80% valid MODIS observations

3.2   Synthetic data

The feasibility of the lake depth tuning was investigated using synthetic (or simulated) data on the LWST generated by Flake forced by atmospheric re-analysis data for a specified lake depth. The synthetic data are often used to verify the properties of a data assimilation system in conditions where the 'truth' is known. In this case they were generated by daily mean modelled LWST, which were then averaged over an eight-day window in order to be consistent with the MODIS LWST in terms of sampling interval. Three types of experiments were designed to evaluate i) the convergence to perpetual-year solutions, ii) the impact of observation/model biases and iii) the impact of missing observations. An 'erroneous' initial depth (with respect to the true depth which has been used to generate the synthetic observations) was then applied and the analysis technique was tested. This allows to estimate the number of iterations needed in the perpetual-
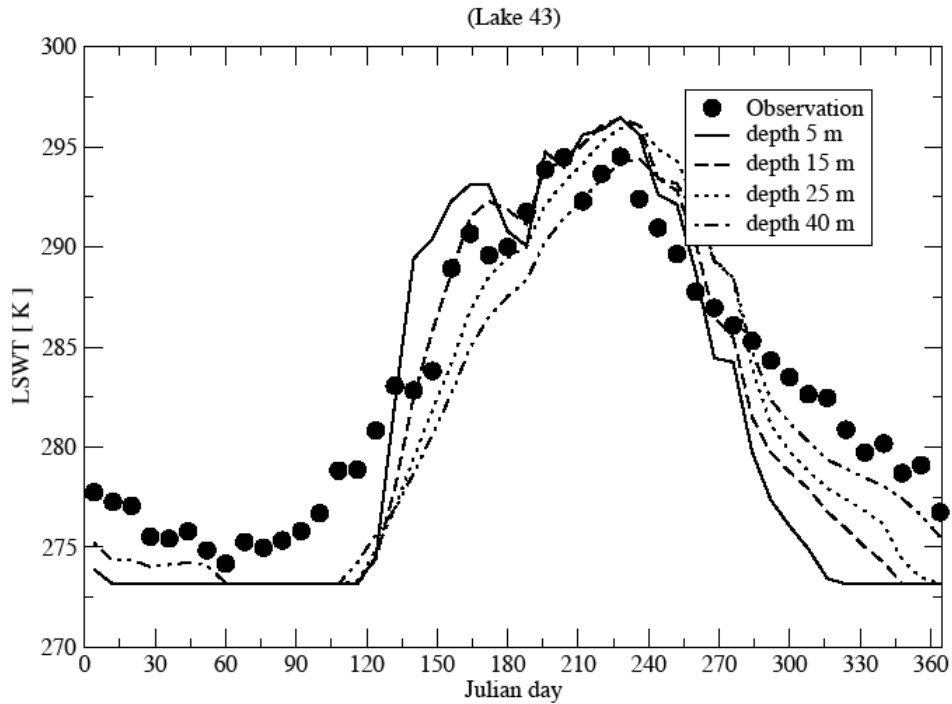
Figure 1 Mean annual cycle of LWST simulated by FLake on Lake Vanern (No. 43 in Table 1) for specified lake depths of 5, 15, 25 and 40 m.
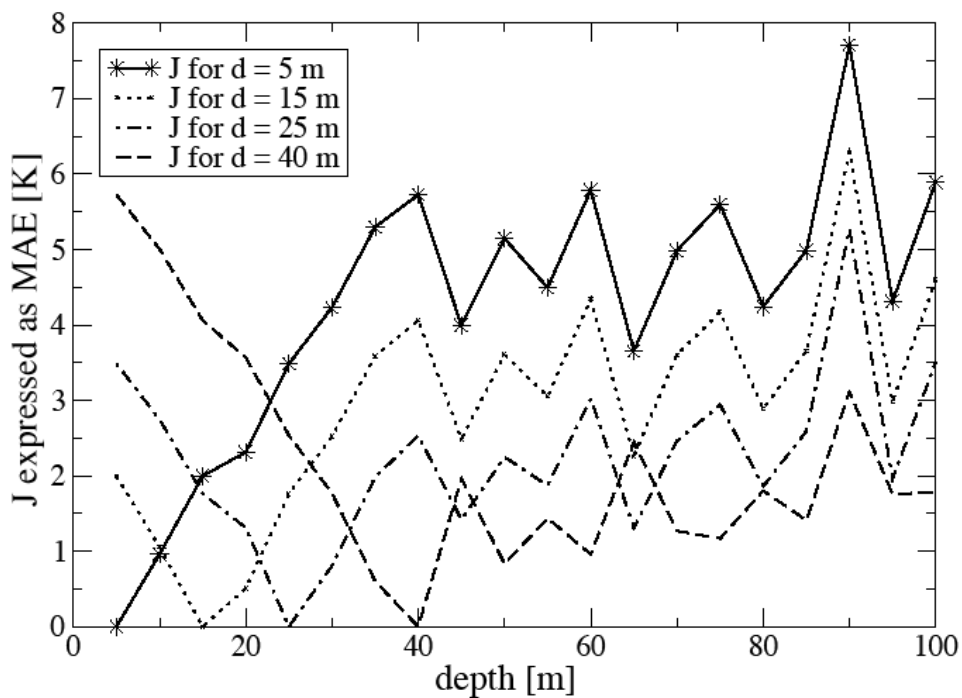

Figure 2 Cost function J expressed as Mean Absolute Error (MAE) between the model LWST and the simulated LWST observations for Lake Vanern (No. 43 in Table 1). Simulated LWST observations are generated assuming a 'true' lake depth of 5, 15, 25 and 40 m (consistently with Fig. 1). The existence of a single minimum of the cost function coincident with the 'true' lake depth is verified for all four depths considered.

year runs to achieve an equilibrium state representation of the LWST climate for a given lake depth. For shallow lakes (up to 25 m) 3-year cycles were typically sufficient to obtain a perpetual-year solution, while for deeper lakes (e.g. 40 m) 5 to 7-year cycles were sometimes needed.

Simulations with different lake depths are shown for illustration (*see* Fig.1). Synthetic observations were extracted from a model simulation at a given depth. The cost function was calculated as Mean Absolute Error (MAE), Eq. 5, between the model LWST and the synthetic observations and its value as a function of lake depth is shown in Fig. 2.

The existence of a single minimum was verified for the 4 depths considered. Convergence is ensured also in the presence of a reduced dataset (in which 20% of synthetic data were removed). The presence of LWST bias can be detrimental if it leads to the cost function minimum that corresponds to an erroneous estimate for the lake depth. Assuming that a large portion of the annual cycle is covered by observations, simple bias removal techniques (subtracting the annual mean difference) can effectively minimise the impact of temperature bias.

3.3   MODIS data

The MODIS Terra/Aqua global composite climatology was generated using Level 3 Mapped Thermal IR SST product. It consists of sea/lake water surface temperature derived from the NASA MODIS sensor on-board the Terra satellite using the 11 and 12 µm IR bands (MODIS channels 31 and 32). The SST data are available from http://oceancolor.gsfc.nasa.gov/.

The eight-day climatology at 4-km resolution built on the 2003-2006 radiances was used in this study. Due to the elongated form of many lakes, not all lakes can be seen from the satellite. In Fig. 3, an example of the lake/sea surface temperature product for Western Europe is shown. A time-series of LWST is taken from the eight-day climatology maps for lakes given in Table 1.
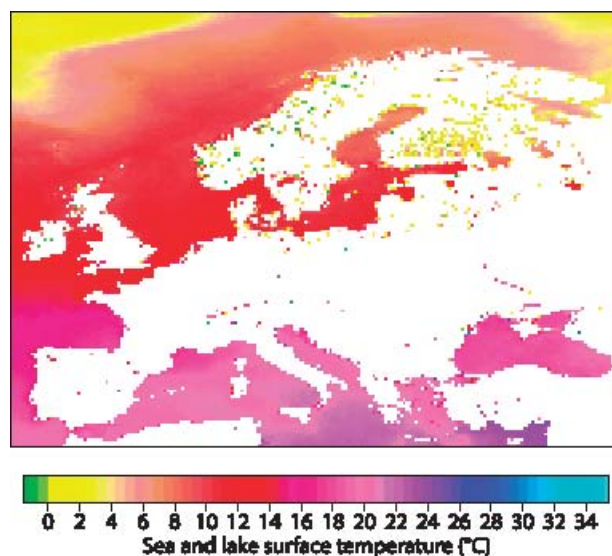

Figure 3 Example of LWST/SST of L3 MODIS product at 4-km resolution

4   Results

4.1   European lakes

The method described above was applied to a large number of European lakes for which the mean and the maximum depth are available. Lakes with less than 80% data coverage from the satellite products (due to missing data) were excluded from the analysis. This reduced the number of lakes to 47. The list of lakes is provided in Table 1. FLake was run in a way to achieve a perpetual-year state with different prescribed depth ranging from 5 m to 100 m, 5 m apart. For each simulation with the assigned lake depth, a cost function J_o given by Eq. 5 was evaluated. The effective lake depth was then determined by the minimum of the cost function J_o.

The method was successfully applied to all 47 lakes given in Table 1. Three lakes were chosen to illustrate the performance of the method. Fig. 4 shows results of simulations for Lake Konnevesi (No. 4 in Table 1) as compared to observations. Different 'simulated lake climates' correspond to different values of the lake depth. The best fit proves to coincide with the in situ estimate of the mean lake depth (Table 1) as indicated by the shape of the cost function in the lower panel of Fig. 4. However, this match is only qualitative and not a necessary condition since the lake portion represented by the simulation may be a smaller portion of the lake, and mean and maximum depth are only given as an indication for the effective depth.

In Fig. 5, results of simulations for Lake Ladoga (No. 32 in Table 1) illustrate performance of the method for a well observed large lake (the largest European lake), for which a deeper value of 40 m depth produces the best match to the observed LWST. It is clear from Fig. 5 that assuming a shallower lake would lead to an overestimation of summer temperatures and the onset and duration of ice cover.

The mean estimated depth for Ladoga is 46.9 m. In Fig. 6, results of simulations for Lake Sniardwy (N. 42 in Table 1) illustrate the case of partially missing observation data. The cost function has a minimum that corresponds to the lake depth of 25 m, which seems to visually match the observed LWST.

In general, the effective lake depth ranges from 10 to 70 m. About 69% of the lakes considered have an effective lake depth lower than the maximum lake depth. Only 15% of lakes have an effective lake depth lower that the in situ estimate of mean depth (in all these cases the effective depth exceeds 10 m).

The mean absolute error in the LWST averaged over all lakes considered was 2.46 K if the effective lake depth is used. In case a fixed lake depth of 25 m was used (this value minimizes the cost function if all lakes listed in Table 1 are taken together), the mean absolute error in LWST is 2.79 K. The effective lake depth brings an average improvement of 12% in terms of the lake surface temperature simulation error as compared to using a fixed depth of 25 m. Errors in the near surface temperature evaluated for the ECMWF operational deterministic forecast (Richardson *et al*. 2008) indicate that over the past 10 years the errors of a three-day forecast ranges from 2K to 3 K. The magnitude of errors in LWST found in this work is within the current accuracy.
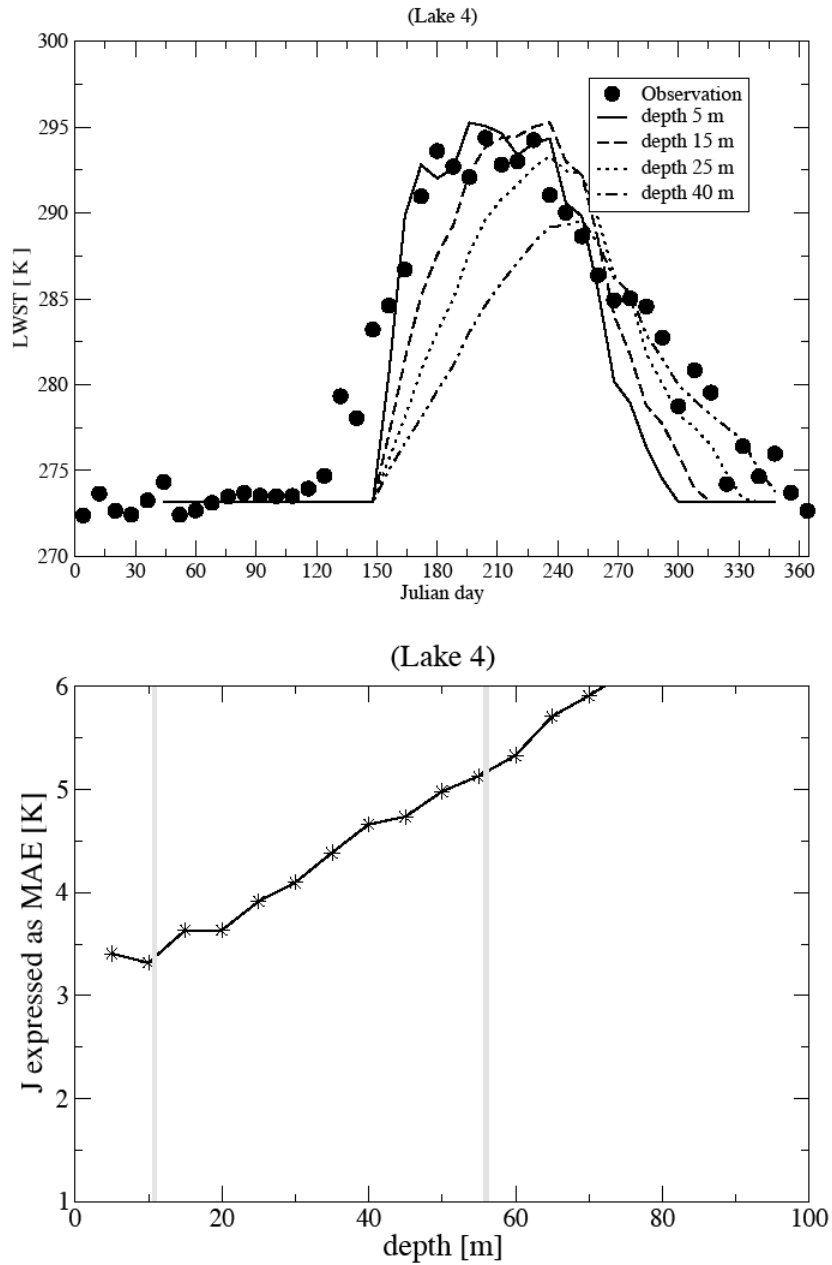
Figure 4: Example of shallow lake LWST simulations with four different lake depths, viz., 5 m, 15 m, 25 m, and 40 m for Lake Konnevesi (No. 4 in Table 1). In the upper panel, lines show simulation results, and circles show observations. The lower panel shows the cost function for different lake depths based on the LWST Mean Absolute Error (MAE). The vertical lines indicate mean and maximum lake depths, respectively, based on in situ measurements.
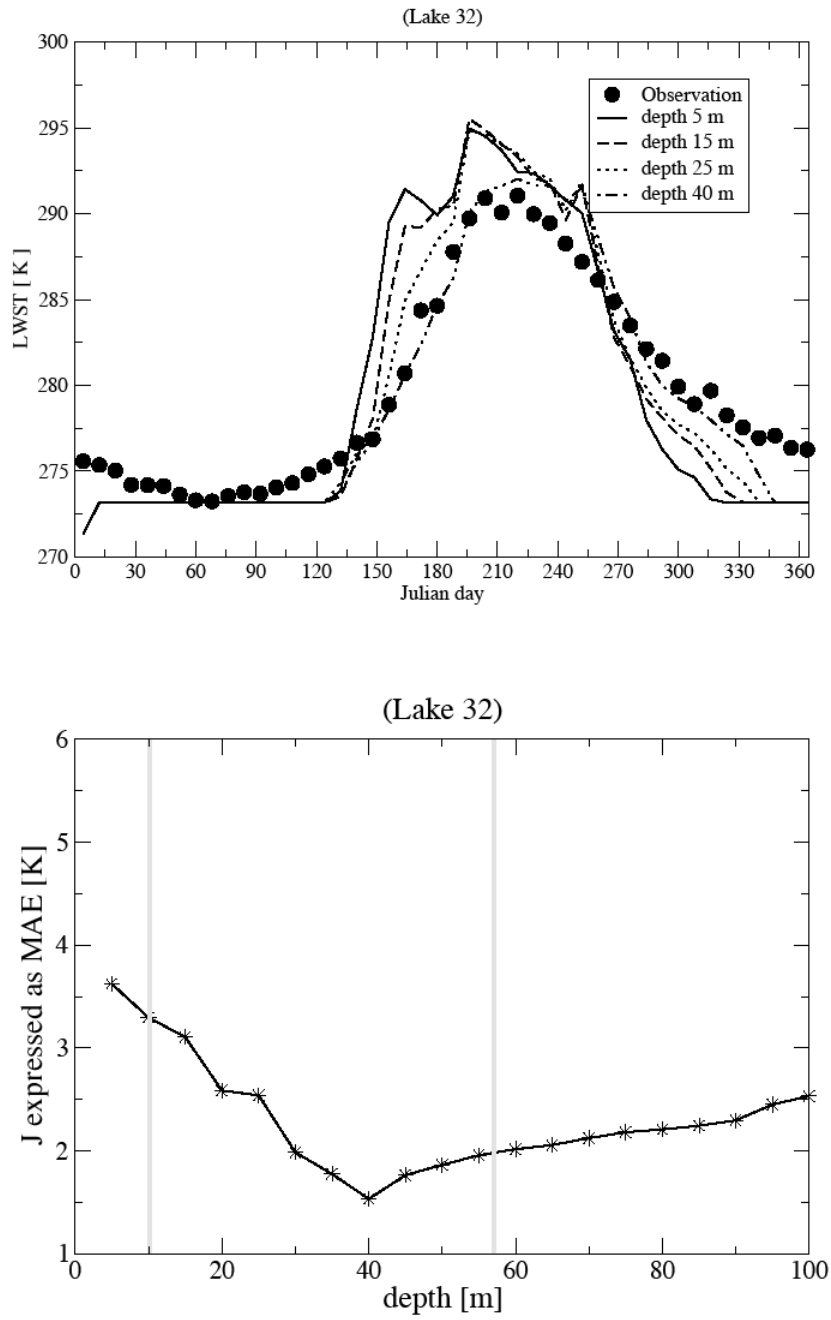
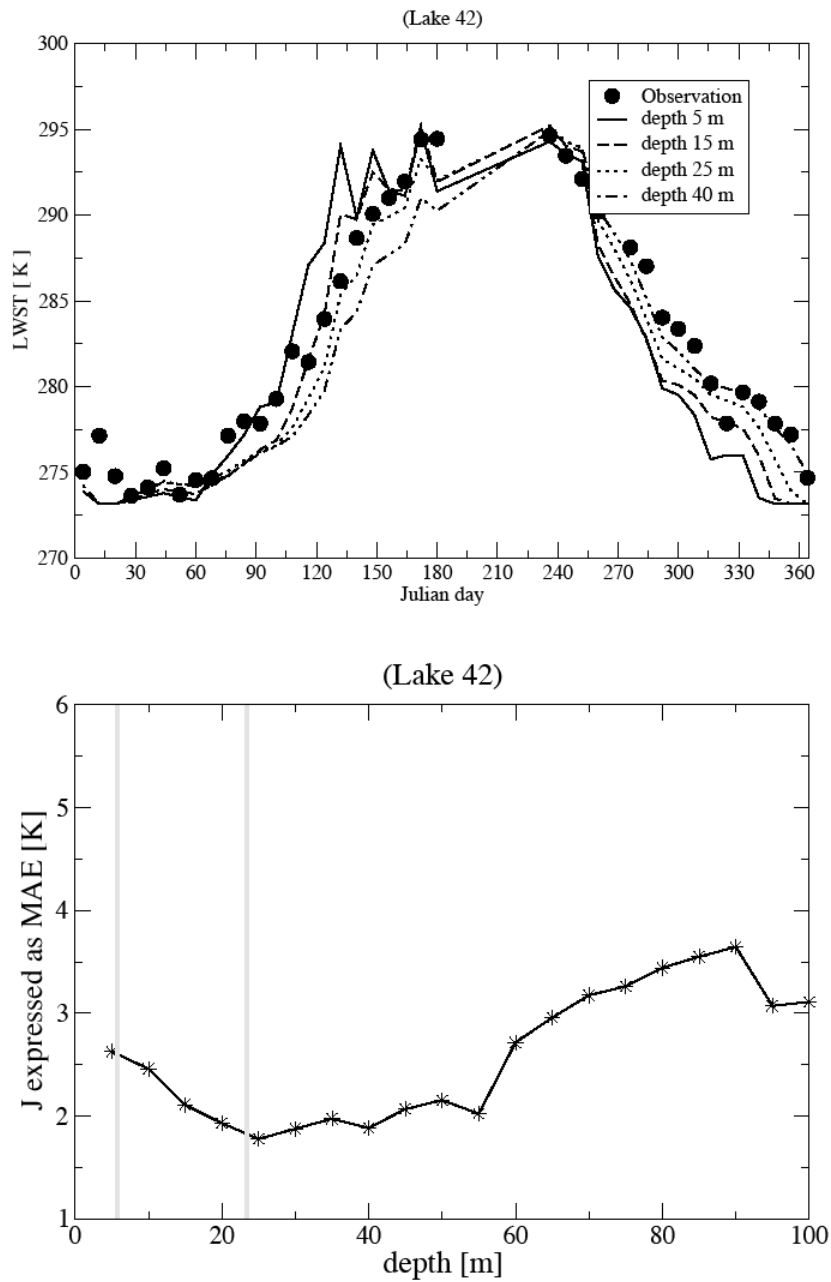*Figure 5: Same as in Fig. 4 but for Lake Ladoga (No. 32 in Table 1)*

Figure 6: Same as in Fig. 4 but for Lake Sniardwy (No. 42 in Table 1).

4.2 American Great Lakes derived bathymetry

A preliminary lake bathymetry map of the American Great Lakes was produced at ERA-Interim horizontal resolution (about 80 km). For this application, the lake depth was minimized using the 8-day MODIS LWST available over the period from 2000 through 2008. The model was initialised in 1989 and the first 10 years of simulation were found to be enough for the model to reach a perpetual-year state. The model was run for a set of depths ranging from 5 and 70 m, 5 m apart. In this case the LWST observations were spatially aggregated to the ERA-Interim model grid (at about 80 km resolution) and the inversion method was then applied over the whole period 2000-2008. The advantage of this procedure is that a much smoother behaviour of the cost function is obtained (not shown).
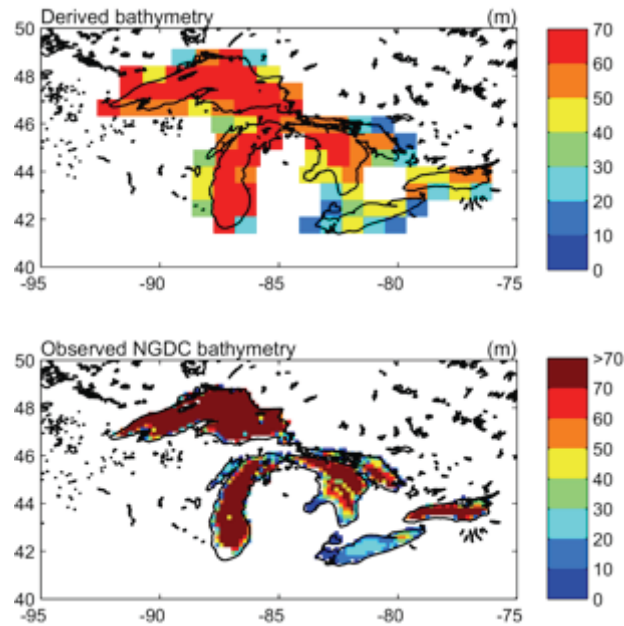
Figure 7: Derived bathymetry of the American Great Lakes at ERA-Interim model resolution (top panel) as compared to the NGDC observed bathymetry, limited to 70 m (bottom panel).

The derived lake bathymetry was compared (Fig. 7) with the actual lake bathymetry taken from the NOAA National Geophysical Data Centre (NGDC, http://www.ngdc.noaa.gov/mgg/greatlakes/greatlakes.html) at a resolution of 10' (about 18 km).

The derived lake bathymetry compares favourably to observations. The lake depth inversion concentrates mostly towards the deep limit for sensitivity with FLake. There is an indication that the shallower parts of Lake Erie (western part) and Lake Huron (southern part) are identified and the deeper parts are well matching the FLake derived maximum depth of 70 m.

5   Discussion

The methodology described above has some caveats, particularly if the working hypotheses described in the introduction break down due to model inaccuracies, observational bias, or forcing errors. The fact that the derived bathymetry may reflect the inability of the lake model (FLake) to correctly represent the LWST may lead to the use of lake depth as an ad-hoc tuning parameter (heavily model dependent), which is surely a limitation to produce a realistic bathymetry. These arguments are typical for any inversion problem. Nonetheless the method described in the present study is a valid alternative to an arbitrary, globally fixed lake depth.

It is reminded that the lake depth is largely unknown at present for thousands of lakes which represent a sizeable fraction of the Earth's surface (Dutra *et al*., 2009). Despite considerable effort being devoted to collecting lake depth estimates, those data cannot yet provide a global coverage and cannot be easily converted into bathymetry fields. Moreover the lake depth is changing over time and the lake water balance is complicated by several factors, as e.g. tributaries and effluents. For lake depths larger that 70 m, the accuracy of the lake depth estimate has no appreciable effect on the LWST cycle simulated by FLake. However, for the lake-depth less than 70 m a large sensitivity to the bathymetry suggests that the

performance of FLake is crucially dependent on the lake depth and simplifications (e.g. the use of fixed depth for all lakes present in the atmospheric model domain instead of their actual depth) may have a detrimental effect on the quality of the forecast. Then, a methodology that tunes the lake depth to obtain the 'best' LWST should lead to an improved forecast, e.g. in terms of evaporation fluxes that are one of the primary concerns in NWP.

6 Summary and conclusions

The present paper illustrates a novel technique to estimate an effective lake depth to be used as the lake-depth physiographic (or bathymetry) field in NWP models. The effective lake depth was estimated by minimizing the annual-mean difference between the observed LWST and the LWST simulated by the lake model FLake. FLake was run in a perpetual-year mode, driven by the near-surface meteorological data provided by the ERA-Interim atmospheric reanalysis. This technique leads to a lake-depth dataset that is optimized for the adopted lake model. Conclusions from the present study are summarised as follows:

- The LWST is highly sensitive to the lake depth for the lake-depth ranging from 10 m to 70 m.

- 'Tuning' of the lake depth is feasible as it leads to a unique lake-depth estimate for all lakes considered in the present study (the cost function has a single minimum), even in presence of missing LWST observation data (up to 20 %).

- For the majority of lakes considered, the effective lake depth appears to be close to the lake depth estimate based on in situ measurements, although a strict comparison is difficult to make (e.g. due to the fact that the observational data may not cover the entire lake surface, making the effective lake depth uncertain). An observational-based lake-depth data set should therefore provide a lake bathymetry in order to prove useful for the implementation of lake models into grid-point atmospheric models.

- A mean 'effective' lake depth, which minimizes the error for all the 47 lakes, appears to be equal to 25 m. This value is recommended for simulations with a fixed lake depth; it can be used as an a priori default estimate when neither in situ measurements of lake depth nor LWST data needed to determine the effective lake depth are available.

- The effective lake depth is shown to bring an average improvement of 12% in term of the LWST simulation error as compared to the fixed depth of 25 m.

- The mean errors of the LWST produced by FLake proved to be within the current accuracy expectations from operational NWP experience.

- A comparison of the effective lake depth derived within ERA-Interim (at about 80 km resolution) is shown to be in qualitative good agreement with the NGDC observed bathymetry (limited to 70 m) for the American Great Lakes. Future work will extend the methodology discussed in the present study to produce the FLake-optimised bathymetry at global scale.

Acknowledgments

References

Arst H., Erm A., Herlevi A., Kutser T., Leppäranta M., Reinart A. and Virta J. 2008. Optical properties of boreal lake waters in Finland and Estonia. *Boreal Env. Res*. **13**: 133-158.

Balsamo G., Viterbo P., Beljaars A., van den Hurk B., Hirschi M., Betts A.K. and Scipal K. 2009. A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the Integrated Forecast System. *J. Hydrometeor.* **10**: 623-643.

Bonan G.B. 1995. Sensitivity of a GCM simulation to inclusion of inland water surfaces. *J. Clim*. **8**: 2691-2704.

Duguay C.R., Flato G.M., Jeries M.O., Ménard P., Morris K. and Rouse W.R. 2003. Ice-cover variability on shallow lakes at high latitudes: model simulations and observations. *Hydrol. Proc.* **17** (17). pp. 3465-3483.

Dutra E., Stepanenko V.M., Balsamo G., Viterbo P., Miranda P.M.A., Mironov D. & Schär C. 2009. An offline study of the impact of lakes on the performance of the ECMWF surface scheme. *Boreal. Env. Res.* Special issue on Lakes in Numerical Weather Prediction (to apear).

Krinner G. 2003. Impact of lakes and wetlands on boreal climate. *J. Geophys. Res*. **108**(D16), 4520, doi:10.1029/2002JD002597.

Lehner B. and Döll P. 2004. Development and validation of a global database of lakes, reservoirs and wetlands. *J. Hydrol*. **296**: 1-22.

Lenters J.D., Kratz T.K. and Bowser C.J. 2005. Effects of climate variability on lake evaporation: Results from a long-term energy budget study of Sparkling Lake, northern Wisconsin (USA). *J. Hydrol*. **308**: 168-195.

Lofgren B.M. 1997. Simulated effects of idealized Laurentian Great Lakes on regional and large-scale climate. *J. Climate* **10**: 2847-2858.

Long Z., Perrie W., Gyakum J., Caya D. and Laprise R. 2007. Northern lake impacts on local seasonal climate. *J. Hydrometeor.* **8**: 881-896.

Loveland T.R., Reed B.C., Brown J.F., Ohlen D.O., Zhu Z., Yang L. and Merchant J.W. 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *Int. J. Remote Sens*. **21** : 1303-1330.

Masson V., Champeaux J.-L., Chauvin F., Meriguet C. and Lacaze R. 2003. A global database of land surface parameters at 1-km resolution in meteorological and climate models. *J. Climate* **16**: 1261-1282.

Mironov D.V. 2008. Parameterization of lakes in numerical weather prediction. Description of a lake model. *COSMO Technical Report*, No. 11, Deutscher Wetterdienst, Offenbach am Main, Germany, 41 pp.

Mironov D., Schneider N., Ritter B. and Heise E. 2005. Implementation of a Lake Model FLake into the Limited-Area NWP System LM of the German Weather Service: Preliminary *Results. Research Activities in Atmospheric and Oceanic Modelling*, J. Cote, Ed., Report No. 35, July 2005, WMO/TD - No. 1220, 4.15-4.16.

Mironov D., Heise E., Kourzeneva E., Ritter B., Schneider N. and Terzhevik A. 2009. Implementation of the lake parameterisation scheme FLake into numerical weather prediction model COSMO. *Boreal. Env. Res.* Special issue on Lakes in Numerical Weather Prediction (to apear).

Oesch, D.C., Jaquet, J.-M., Hauser A., and Wunderle S., 2005. Lake surface water temperature retrieval using advanced very high resolution radiometer and Moderate Resolution Imaging Spectroradiometer data: Validation and feasibility study, *J. Geophys. Res.* **110**, doi 10.1029/2004JC002857.

Park K.-A., Jong Y.C., Kim K. & Cornillon P.C. 2005. Wind and bathymetric forcing of the annual sea surface temperature signal in the East (Japan) Sea. *Geophy. Res. Lett.* **32**, L05610, doi:10.1029/2004GL022197.

Simmons A.J., Uppala S.M., Dee D. and Kobayashi S. 2007. ERA-Interim: New ECMWF reanalysis product from 1989 onwards. *ECMWF Newsletter*, 110: 25-35. http://www.ecmwf.int/publications/newsletters/

Uppala S.M., Kallberg P.W., Simmons A.J., Andrae U., Costa Bechtold V. Da Fiorino M., Gibson J.K., Haseler J., Hernandez A., Kelly G.A., Li X., Onogi K., Saarinen S., Sokka N., Allan R.P., Andersson E., Arpe K., Balmaseda M.A., Beljaars A.C.M., Berg L., Van De Bidlot J., Bormann N., Caires S., Chevallier F., Dethof A., Dragosavac M., Fisher M., Fuentes M., Hagemann S., Holm E., Hoskins B.J., Isaksen L., Janssen P.A.E.M., Jenne R., McNally A.P., Mahfouf J.-E., Morcrette J.-J., Rayner N.A., Saunders R.W., Simon P., Steri A., Trenberth K.E., Untch A., Vasiljevic D. 2005. The ERA-40 re-analysis. *Quart. J. Roy. Meteor. Soc*. **131**: 2961-3012.

Van den Hurk B., Viterbo P., Beljaars A.C.M. & Betts A.K., 2000. Offline validation of the ERA40 surface scheme. *ECMWF Technical Memorandum* 295: 42 pp http://www.ecmwf.int/publications/

Richardson D.S., Bidlot J., Ferranti L., Ghelli A., Janousek M., Leutbecher M., Prates F., Vitart F. & Zsoter E. 2008. Verification statistics and evaluations of ECMWF forecasts in 2007-2008. *ECMWF Technical Memorandum* 578, p. 45. 230 http://www.ecmwf.int/publications/

Viterbo P. and Beljaars A.C.M. 1995. An Improved Land-Surface Parameterization Scheme in the Ecmwf Model and Its Validation. *J. Climate* **8**: 2716-2748.

Vrbancich J., Fullagar P.K. & Macnae J. 2000. Bathymetry and seafloor mapping via one dimensional inversion and conductivity depth imaging of AEM. *Explorat. Geophys*. **31**: 603-610.

Xie S., Hafner J., Tanimoto Y., Liu W.T., Tokinaga H. & Xu H. 2002. Bathymetric effect on the winter sea surface temperature and climate of the Yellow and East China Seas. *Geophys. Res. Lett.* **29**(24), 2228, doi:10.1029/2002GL015884.