Meteorologische Zeitschrift, Vol. **22**, No. 6, 675–684 (published online January 2014)

Open Access Article

Regional Evaluation of ERA-40 Reanalysis Data with Marine Atmospheric Observations in the North Sea Area

NILS H. SCHADE[1,*], HARTMUT HEINRICH[1] and GUDRUN ROSENHAGEN[2]

[1]Bundesamt für Seeschifffahrt und Hydrographie, Hamburg, Germany
[2]Deutscher Wetterdienst, Hamburg, Germany

(Manuscript received January 21, 2013; in revised form November 1, 2013; accepted November 15, 2013)

Abstract

An important task of the departmental research programme KLIWAS is the evaluation and assessment of climate model results by means of a comprehensive reference data set. For validation purposes, and to create a North Sea wide maritime atmospheric and oceanographic reference database, in-situ observations of the *Centre for Global Marine Meteorological Observations* (GZS) of the *National Meteorological Service* DWD have been compared to the ERA-40 reanalysis. ERA-40 is used as forcing for the hindcast runs of the ENSEMBLES regional climate models, which is used within the KLIWAS model chain. The GZS hosts a regularly updated, quality controlled, world-wide data bank of weather observations from the oceans. It includes data from all sorts of observation platforms as *Voluntary Observing Ships* (VOS), drifting and moored buoys, light vessels, and offshore platforms, either from real-time (RT) via the *Global Telecommunication System* (GTS) or from international exchange in delayed-mode (DM). In addition to the automated set of programs applied for high quality control, erroneous data are also manually corrected to a certain extent, if possible. To assure reliable statistics for the evaluation, the corrected observations are gridded to a resolution of 2.25 degree, so each grid box includes four ERA-40 reanalysis grid points. The temporal coverage of the grid boxes depends on shipping routes and the positions of automated systems. Observed air temperatures, covering a period of 40 years (1961–2000), show noticeable differences to the reanalysis data for all land influenced boxes, specifically in the winter months. The same differences can be found if ERA-40 data alone are compared between land- and sea facing boxes. They can not be found in GZS data. It can be assumed that the differences are not resulting from measurement errors or uncertain fraction variabilities, since they are small during the winter months. A comparison of the differences basing on the 1981–2000 period to those of the 1961–1980 period shows an increase. Further investigations reveal that differences are largest in the landward part of the grid boxes, whereas in the parts facing to the open sea, observations and ERA-40 are in fair agreement. This leads to the conclusion, that the resolution of the ERA-40 reanalysis is not sufficient for detailed analyses of air temperatures near the coasts. However, this problem does not occur for the sea level pressure. The described land influence does not interfere with parameters, that are more or less insensitive to the land-sea distribution.

Keywords: KLIWAS, ERA-40, Marine in-situ Observations, North Sea area, Regional Climate Modelling.

1 Introduction

KLIWAS is a research program of the *German Federal Ministry of Transport, Building and Urban Development* (BMVBS). It aims at the investigation of potential consequences of climate change for navigation on inland and coastal waterways and to formulate appropriate strategies for adaptation to changed environmental conditions in the future. Basic objectives of the project are (1) the evaluation and assessment of climate model results by means of a comprehensive reference database and (2) to make this data available for subsequent projects.

Current climate models are not able to sufficiently reproduce the spatial variability in the North Sea area. Especially, no regional coupled Atmosphere-Ocean model (RCAOM) exists to describe the complex non-linear interactions of air/sea fluxes properly. Nevertheless, SCHRUM et al. (2003) showed promising results of a coupled hindcast model run over a full seasonal cycle with clear improvements compared to the uncoupled run. Therefore, within KLIWAS three RCAOMs will be implemented. Until the results of these models can be investigated, the focus lies on the available uncoupled ENSEMBLES data. The ENSEMBLES project (VAN DEN LINDEN and MITCHELL, 2009) ran from 2004 to 2009, coordinated by the *Met Office Hadley Centre*. Probabilistic projections of the future climate were produced for the European region, based on an ensemble of 15 regional climate models (RCMs), nested in 5 global models (GCMs) and forced by the A1B emissions scenario. The hindcast runs have been performed using the ERA-40 reanalysis as model input. Results focussing on terrestrial regions have been discussed in the literature, e.g. by LORENZ and JACOB (2010), HAYLOCK et al. (2008), or KLOK and KLEINTANK (2009).

Up to now, no respective investigations for the North Sea region exist, so the development of a framework for the evaluation of a comprehensive marine

*Corresponding author: Nils H. Schade, Bundesamt für Seeschifffahrt und Hydrographie, Bernhard-Nocht-Strasse 78, 20359 Hamburg, Germany, e-mail: nils.schade@bsh.de

DOI 10.1127/0941-2948/2013/0471
0941-2948/2013/0471 $ 4.50

ocean-atmosphere climatology in the North Sea and adjacent parts of the North Atlantic and the Baltic Sea has become an important task. The North Sea is of particular interest and importance for marine carrier operations and off-shore activities. Especially in this context, subsequent projects within the KLIWAS model chain investigate amongst others appropriate strategies for adaption to changed environmental conditions to safeguard the efficiency of transport ways and to preserve the water quality and the habitats in coastal waters. Thus, the quality of reanalyses and operational analyses should be precisely checked and possible biases compared to in-situ observations should be addressed.

For the validation of the RCM output over the North Sea area, high quality controlled surface marine in-situ observations and/or reanalyses data are needed as reference of the actual climate state. In-situ observations are the sole ground based measurements of atmospheric parameters over the open sea. They provide information that cannot be reliably measured from satellites (Kent and Berry, 2005), e.g. surface air temperature and sea level pressure, although the data quality varies within the course of time. Several publications dealing with the popular *International Comprehensive Ocean-Atmosphere Data Set* ICOADS (Worley et al., 2005; Woodruff et al., 2011) are focusing on the ocean on a global scale. Some authors have specifically investigated the importance of ship log data in scientific research, also on a regional scale, e.g. Kuettel et al. (2009) for the reconstruction of past sea level pressure fields derived from ship logbook wind data over the North Atlantic/ European area. Nevertheless, gridded data sets are mostly of an insufficient resolution for regional climate analyses and based on terrestrial observations, containing only a small number of island stations, if any (Alexander et al., 2009; Jones et al., 1999).

Here, we introduce the database of the *Centre for Global Marine Meteorological Observations*, further referred to as GZS. This database is hosted by the *German Meteorological Service* DWD and aims to ensure the production and supply of marine climatological data.

For the evaluation of the ERA-40 data, areas were identified with a sufficient number and quality of observations in GZS. The investigated data in this study, covering a time period of 40 years beginning in 1961, are described in the following section 2. DWD's high quality control procedures for the marine atmospheric observations are described briefly in section 2.1 and the data preprocessing in section 2.2. Results of our investigations are presented in section 3, followed by a summary and conclusion in section 4.

2 Data

GZS is regularly updated with meteorological observations over the world's oceans, reaching back to 1850. It contains the available data from all sorts of measurement platforms: Data from *Voluntary Observing merchant Ships* (VOS), research, naval and light vessels, buoys, and digitized data from historical weather journals (DWD project HISTOR, Gloeden, 2011) and registrations. Altogether, GZS consists of more than 188 million real-time, distributed via the international *Global Telecommunication System* (GTS), and non real-time weather reports. These data from national sources and bilateral or international exchange are complemented by data from ICOADS. ICOADS has an overall proportion of 20,21% in GZS for the investigated period from 1961–2000, decreasing towards the end of the period (11.91% in the last decade). Since the proportion of GTS reports, inherent in both GZS and ICOADS, further increases, additional reports from ICOADS in GZS are now of about 3.5%.

The ERA-40 Reanalyses (Uppala et al., 2005) covers the time period from 1957 to 2002 and contains worldwide meteorological parameters in 60 height levels on a 6 hourly basis. It is provided by the *European Centre for Medium-Range Weather Forecast* (ECMWF) in collaboration with other institutions, e.g. the MetOffice, Exeter, UK, or the Max-Planck-Institut für Meteorologie, Hamburg, Germany. Several data types are assimilated, i.e. operational real-time data distributed via GTS (1979–2002), observations from aircrafts, buoys, offshore platforms, radiosondes and ship data, taken from ICOADS (1950–1999). Since 1973 satellite data are included also. The assimilation system uses an updated form of the older 3D variational analyses of the ECMWF with a spectral T159 resolution, corresponding to a 1.125×1.125 degree geographical grid. The model has a reduced Gaussian grid of about 1.125 degree in latitude and a reduced number of longitudinal grid points from the Equator to the Poles.

Each data set has its own particular problem: While GZS marine in-situ observations are unevenly distributed in space and time and are derived from several different observation types from different countries and with different methods for the measured parameters, the spatial resolution of the existing reanalyses data is coarse and the models suffer from biases in numerics and parameterizations, inhomogeneities of data assimilation input, etc. In this context ERA-40 is used as an example for reanalyses data, since the KLIWAS model chain is based on the ENSEMBLES Hindcast runs which are forced by ERA-40. Further investigations with ERA-Interim and other reanalyses products are planned in the future.

In the following, we describe the procedures to obtain GZS air temperatures and sea level pressures in the North Sea area of best possible quality to conduct reliable comparisons with ERA-40 reanalysis data.

2.1 High Quality Control

The high quality control (HQC) procedures used for GZS data are more or less the same as published in the literature, e.g. by GANDLIN (1987), or ISHII et al. (2003), so they are described only briefly here:

HQC is based on the program by HOEFLICH et al. (1975). In a first step, the data are checked on doubles (i.e. already existing data or simply data received twice or more often), the correct contents of each column, i.e. allowed values in the column "air temperatures" etc., and a date/time and position tracking ("cruise control") by means of reported position, course and speed. For this, all observations have to be related to previous and subsequent observation times, i.e. date and time, the geographical positions, and observed values. But also comparisons to other observations in the vicinity of the checked data, spatial as well as temporal, are carried out ("vicinity test").

Climate and consistency checks follow. All data have to be within reasonable boundaries (minimum and maximum values) depending on the geographical position. In this "climate test", sea level pressure as well as air and sea surface temperatures are compared to the marine data base of DWD (MARDAB), containing the number of observations, mean value and variation for each 2.5° or 5° grid world wide. Correspondingly, the MARDAB boxes in the North Sea area include a high number of observations and the climate test is supposed to be reliable. Here, the boundaries for air temperatures are −25 °C and +40 °C, for sea level pressure 930 hPa and 1050 hPa.

Special routines check for consistency and abidance by the laws of physics (e.g. wet-bulb temperature < air temperature). As a result of the quality control, all observations are flagged (values of 0, "no control" up to 9, "missing") and, in case of an erroneous or doubtful value, checked manually, corrected if possible and finally written in file. However, manual correction can only be applied for a limited number of non real-time weather reports, which is actual 3% of the incoming data.

2.2 Preprocessing

All those different observational data contain specific errors that can not be detected by the HQC and have to be addressed in preprocessing: Biases known to be related to the composition of nations, to which the ships belong, instrumental random errors and sampling errors.

ISHII et al. (2005) noted inhomogeneities in the quality of the historical observations due to missing call signs, ship identifications or changes in the observation methods (q.v. CARDONE et al., 1990). Low pressure biases might be caused by "suction", i.e. air flowing from inside a ship leeward to the outside (HAYASHI, 1974), and air temperatures may be to a certain extent affected by the solar heating of the instruments and their surroundings (BERRY et al., 2004). KENT and BERRY (2005) found random measurement errors averaged for each month in the period 1970 to 2002 of 1.2 ± 0.3 K (air temperature, with height correction) and of 2.4 ± 0.9 hPa (sea level pressure) in the North Sea area.

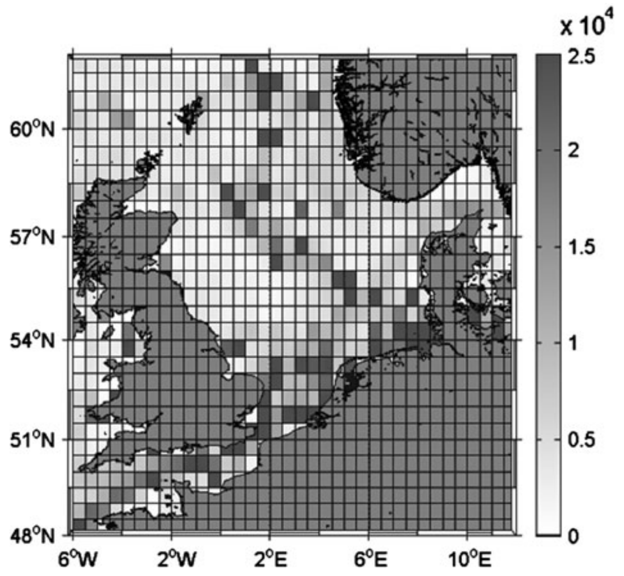

Figure 1: Distribution of the total number of air temperature observations in GZS in the North Sea area. Resolution: 0.5°. Base period: 1961–2000.

Both estimates have been found to decrease towards the end of the period, assumed to be related to an increased understanding of the importance of positioning and quality of the instruments. GULEV et al. (2003) investigated VOS wave observations and found fair weather biases due to the fact, that ships avoid stormy conditions.

Observations in the North Sea, however, are "contaminated" by discrete shipping routes, leading along the coasts and around Scotland into the North Atlantic, displayed in Fig. 1. Shown is the distribution of air temperature observations in GZS on a 0.5 degree resolution from 1961–2000 that passed the HQC. Obviously, for most grid boxes in the North Sea area the occupancy is insufficient for a robust investigation. Two criteria were therefore chosen to select the boxes of investigation: (1) Enough high quality observations of good timely distribution and (2) no remarkable land influence. The two boxes that passed these criteria are shown in Fig. 2 (left). Since the ERA-40 reanalysis data are stored on an irregular grid of 1.125 degree in longitude and between 1.121 degree and 1.125 degree in latitude, all grid boxes for the investigations are chosen thus, that four ERA-40 grid points are centred in the middle. All observations were averaged around the 6-hourly ERA-40 output and only those time points for which data of both data bases exist, further referred to as "GZS-like sampling", were evaluated.

Comparing the number of observations in box 2, ICOADS includes about twice as much air temperature data in the early 60s and 70s compared to GZS, whereas the rest of the period is in good agreement, especially since the late 70s. Box 1 shows the exact opposite (not shown). This may be due to slightly different thresholds in the different quality control procedures and to a certain extent different processed data. Since time series

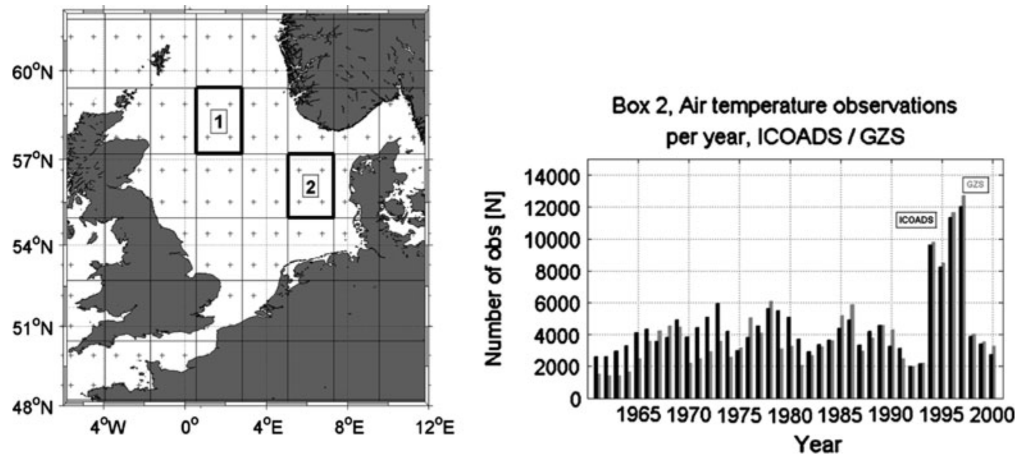

Figure 2: North Sea area showing the two investigated grid boxes (left). Both boxes consist of four sub grid boxes, each referring to one ERA-40 grid point (grey plus). Also shown: Histogram of air temperature observations by ICOADS (black) and GZS (grey) in Box 2 (right). Base period: 1961–2000.

of yearly mean air temperatures derived from monthly means show a good agreement between GZS and ICOADS in both boxes, both data sets should be comparable here. The same arguments apply for sea level pressure observations which are in even better agreement.

To address biases in marine observational data, information about the metadata are needed. Kent et al. (2007) used the WMO *Publication No .47* (e.g. WMO, 1994), which contain information about the contributing VOS ships and the instruments used. Until the late 1980s, the thermometer type was mostly unknown, also was the method of exposure. Therefore, no correction has been applied. Biases related to changing measurement heights (about 0.01 K/m) due to the fact that ships became larger would lead to a cold bias in global air temperatures of about 0.07 K in 2000, relative to 1970 values. In the North Sea area, the mean measurement height changed from the 1970–1979 to the 1995–2004 period from 10 m to 20 m, which result in a 0.1 K cold bias (or 0.18 K, if corrected at 2 m). A height correction would therefore reduce the random measurement error by Kent and Berry (2005) to 1.1 (1.0) ± 0.3 K.

Further, there is a variety of nations contributing to GZS, associated with specific biases. Some Russian ships are known to deviate about 2–3 K in air temperature observations (personal communication). Most of these substantial biases should have been sorted out by the "vicinity test", unless there are no neighbouring observations. For the investigation only those observations were used that appear to be inconspicuous according to the HQC. Furthermore, the GZS data in both boxes were manually evaluated concerning possible biases caused by single vessels and fixed positions, but showed no noticeable inhomogeneities as to space and time, in spite of the fact that some vessels and platforms (prominently the German fishing trawler "Walther Herwig III" and the Dutch oil rig "Maersk Endeavour") in Box 2 have submitted a relatively high number of observations. Since only time averaged values have been further used, and the exclusion of those vessels shows no substantial differences in the results, this fact can be considered negligible for our investigations (which of course does not mean it should be assumed to apply for other boxes and areas as well without prior investigation).

Table 1: Total Sampling Error (TSU) for air temperatures (AT) and sea level pressure (SLP) for both North Sea boxes for the whole period 1961–2000 and the respective summer (August) and winter (December) month.

	Box 1	Box 2
TSU AT	±0.2 K	±0.36 K
TSU AT (summer)	±0.32 K	±0.25 K
TSU AT (winter)	±0.15 K	±0.56 K
TSU SLP	±0.3 hPa	±0.47 hPa
TSU SLP (summer)	±0.18 hPa	±0.34 hPa
TSU SLP (winter)	±0.51 hPa	±0.77 hPa

Concerning sampling errors we followed the method of Gulev et al. (2007) and compared monthly means of our GZS-like sampled ERA-40 data with regularly sampled ERA-40 data to obtain the total sampling uncertainties (TSU). These would be inherent in both datasets compared in this study. Results are shown in Table 1: TSU of air temperatures for the whole period 1961–2000 are within ± 0.2 K for Box 1 and ± 0.36 K for Box 2 with higher values in the summer (winter) in Box 1(2). Furthermore, the TSU winter values for Box 2 are slightly increasing towards the end of the period, whereas they remain constant in summer and throughout the whole period in Box 1. Sea level pressure data show TSUs of ± 0.3 hPa for Box 1 and ± 0.47 hPa for Box 2. Here, TSU is close to zero since the late 80s in Box 1 but shows an increase up to ± 1 hPa in winter in Box 2. Since we investigate GZS-like sampled ERA-40 data with GZS data, these sampling errors are not accountable for any differences between the two data sets presented in the following. Nevertheless, they surely point out the need to correct GZS data for climate research studies accordingly, if they want to be compared with the regularly sampled ERA-40 (or any reanalyses) data for that matter.

Finally, the uncertain fraction variability has to be addressed. Since we have to assume, that observational errors occur, we have to account for the representativeness of GZS. STOFFELEN (1998) used a triple-collocation method to validate anemometer, scatterometer and NCEP winds. This can not be done with only two datasets available, unless one is divided into subsamples, which would not be practical for our analyses. Therefore, we have to assume the GZS temperature und sea level pressure errors by what we know. This is, as stated above, the random measurement error (KENT and BERRY, 2005), which is used as "expected variance" σ^2. Secondly, the "error variances" ε^2 are computed as variance of the GZS values of each time step per box. Now, an error estimate, or pseudobias, can be obtained that is the difference between the GZS value x and the biased mean $< y > = x\sigma^2 / (\sigma^2 + \varepsilon^2)$. This pseudobias is small if ε^2 is small, which means the variance of measured GZS values is small. Mean pseudobiases of 1.4 ± 1.9 K (Box 1) and 1.7 ± 2.5 K (Box 2) for air temperatures can be found, with minimum values of about 1 ± 0.5 K in winter (December) and maximum values of about 3 ± 1.2 K in summer (August). Sea level pressure biases are found to be 1.9 ± 2.1 hPa (Box 1) and 2.4 ± 2.9 hPa (Box 2).

Nevertheless, to correct the GZS data this way would imply to do so for every ship, variable and measurement device separately, which is not possible in the context of this work. Additional information about e.g. the shielding of thermometers, the exposure times, etc. would have to be investigated for every single measurement for the whole time period. Often this data not available at all and to correct only a part of the data and assume that this correction applies to the rest as well would introduce biases also. Instead we decided to use the "raw" GZS data after they passed HQC and the manual inspection for box 1 und 2 and to keep in mind that measurement errors and pseudobiases exist and to account for them in the discussion. As mentioned above, the sampling bias does not effect the differences between both datasets, if we compare GZS-like sampled ERA-40 with GZS.

3 Results

The distributions of air temperatures at 2 m height above ground (AT) are presented in Fig. 3 through box plots. They show the percentile values of both ERA-40 and observed ATs for all months of the reference period 1961–2000, and for the winter season (DJF) only. The upper row includes both grid boxes, the lower row shows the four subgrid boxes of Box 2 (see Fig. 2). For the overall data, ERA-40 (grey boxes) and observations (black boxes) are very close together and the differences in the respective percentiles are only marginal. For DJF however, differences in Box 2 are obvious. The median value of the ERA-40 data is 0.6 K below the respective one of the observations and the 99% percentile differs by 1 K. The percentiles 1 and 5 are in better agreement. A closer look at the four sub grid boxes of Box 2 reveals an interesting feature: Both eastern boxes facing to the Danish coast (even numbers) indicate large differences in the ERA-40 temperatures in the sea side part, compared to the GSZ, as well as to ERA-40 itself. Differences are about 1–2 K in the respective higher (warm) percentiles and the median, whereas the lower (cold) percentiles differ less. This reflects the characteristic of the reanalysis model not to differentiate exactly between land and sea and a land-influence reaching far into the open sea areas. The western, sea-side boxes (odd numbers) are much closer together, but show also colder maximum temperatures compared to GZS. Again, this can be observed only in DJF, all other periods are in good agreement. Even, if measurement errors and biases are accounted for, this would not explain the differences between the ERA-40 values in the eastern boxes to those in western boxes.

However, these differences are only apparent for the AT investigations. Comparisons of the SLP data show no such differences between ERA-40 and the observations (not displayed), neither in the median nor in the percentile values. This leads to the conclusion that the ERA-40 land-facing grid points are influenced by colder temperatures occurring over land in the winter months DJF, whereas the pressure fields are independent of the subsurface and land-sea effects can be disregarded. This independency of the land-sea distribution leads to a better consistency between observations and reanalysis. The assumption is supported by the results presented in Fig. 4: All ERA-40 monthly means of August and December during the period 1961–2000 for Box 2 are plotted against the respective means basing on observations of AT (left graph) and SLP (right graph). These months show the best and worst fitting data pairs. While in August all AT pairs are relatively close around the optimum accordance line, despite the high uncertain fraction variability, all ERA-40 mean ATs in December are systematically smaller than the respective means basing on observations, with a maximum difference of 2.2 ± 0.2 K. This is about twice the size of the monthly mean random measurement error of 1.2 ± 0.3 K found by KENT and BERRY (2005). Therefore, the colder ERA-40 temperatures can not result from measurement errors in the observations alone, especially in December where the pseudobias due to the uncertain fraction variability is low. In fact, measurement errors resulting from radiative heating should have a greater effect in the summer months. Here, ERA-40 means are smaller during the whole span from October to April (not shown), until in May the mean values become closer to the optimum accordance line and begin to differ again in September.

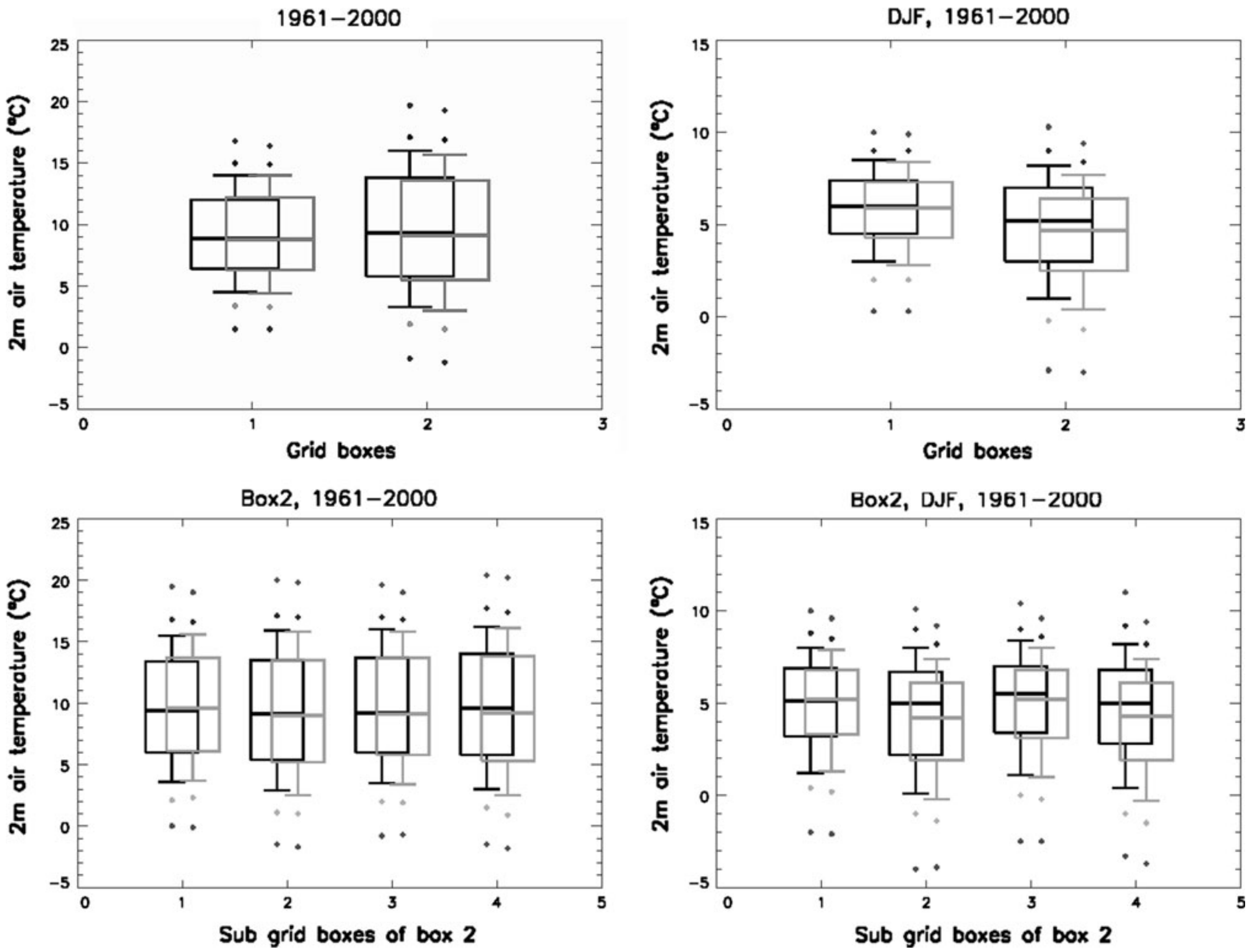

Figure 3: Box plots of 2 m air temperatures from measurements (black) and ERA-40 (grey) for the period 1961–2000. Above left: Annual values, above right: Winter month DJF for Box 1 and 2. Median values and the 25 and 75 percentiles are marked by the box, the whiskers show the 10 and 90 percentiles, the plus the percentiles 1, 5, 95 and 99. Below: Same as above but for the four subgrid boxes of Box 2 (see Fig.1). Boxes facing to the open sea have odd, land side boxes even numbers.

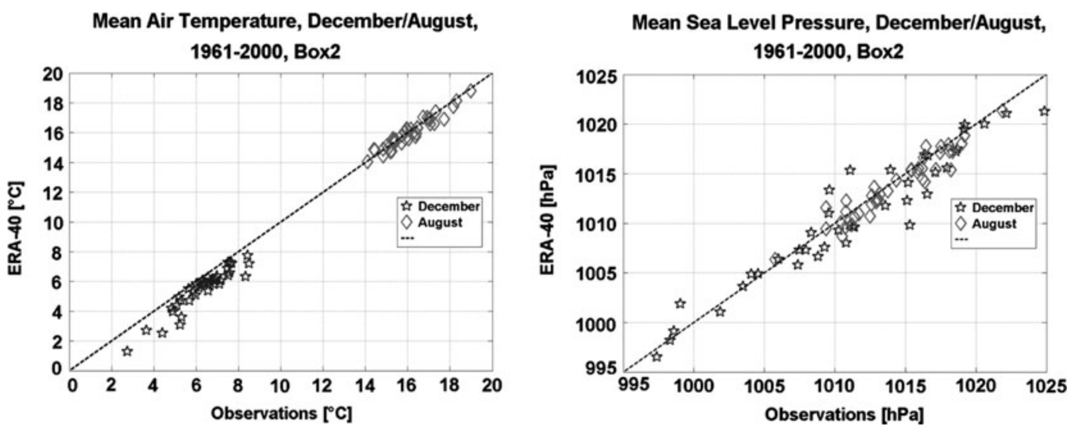

Figure 4: Scatter plot of monthly mean 2 m air temperatures (left) and sea level pressures (right) of measurement and ERA-40 data pairs in Box 2 in December (black stars) and August (grey diamonds). Base period: 1961–2000.

The values of mean SLP in contrast show no systematic differences between December and August. The December values differ more from the optimum accordance line and show a larger spread than the August values, but are more or less uniformly distributed. The larger spread in December can be explained by an increased low pressure activity and pronounced pressure differences. More low pressure system paths lead through

Northern Europe, especially during NAO + phases in winter. It also should be noted that ERA-40 and the observations differ most at SLP values between 1010 hPa and 1017 hPa in December.

At last the ATs are evaluated for two different reference periods 1961 to 1980 and 1981 to 2000 (see Fig. 5) to investigate possible temporal changes. Box 2, which is closer to the coast, shows an increasing spread in the whole frequency distribution as well for the observations as for the ERA-40 data. However, there is hardly any difference of the median between the two periods for the ERA-40 data. Over the open sea (Box 1), only minor changes in the percentile values of both data sets can be detected. A closer look into the summer (JJA) and the winter (DJF) seasons reveals temperature increases in both boxes for both data sets and both seasons for the 1981–2000 period compared to the 1961–1980 period: Slightly higher temperatures in all median values and an increased spread, but smaller minimum temperatures for DJF and higher maximum temperatures for JJA can be found in the southern Box 2, and a small overall increase in the higher percentiles and almost no changes in the median over the northern Box 1. The ERA-40 data show higher JJA maximum temperatures and lower DJF minimum temperatures than the observations, and also a smaller increase in the median. The differences between ERA-40 and GZS JJA maximum temperatures should be even higher when taking potential radiative heating errors and the uncertain fraction variability in GZS into account.

The SLP data have no such differences between the two reference periods in neither of the data sets. The frequency distributions of the percentile values suggest on one hand that the large scale circulation patterns have not or at least not noticeably changed. On the other hand it suggests that the SLP is reproduced well by the ERA-40 reanalysis system in the North Sea area in the investigated time periods.

4 Summary and conclusions

For the validation of regional climate models, high quality observations and/or reanalyses (i.e. ERA-40) data are needed as reference of the actual climate state. Comparisons of marine atmospheric in-situ observations with ERA-40 show a systematic cold bias for the air temperature near the Danish coast in the reanalyses data for the winter months (DJF). This apparent land-induced bias does not appear in the analysis for sea level pressure, probably due to its larger scale and land-sea insensitivity. The bias may be caused by the interpolation process of ERA-40 from the spherical calculations to the final grid. Overall, ERA-40 agrees well with the observations of sea level pressure and, off the coast, also with the 2 m air temperature results. Near the coasts, one should be careful using ERA-40 air temperature data. The cold bias compared to GZS temperatures could be partly related to measurement errors, prominently by radiative heating, and the uncertain fraction variability. It still should be accounted for, since ERA-40 itself is biased to the same degree near the coasts compared to sea-facing boxes. It should be further noted, that these results refer only to the two parameters air temperature and sea level pressure and base on two boxes in the North Sea area only. If a regularly sampled reanalysis product is compared with in-situ observations, the sampling bias has to be taken into consideration also.

A higher resolution of the reanalyses data set might reduce the land-influenced bias and improve the quality of air temperature results near the coast. This would provide a better reference for subsequent high resolution modelling, coastal engineering, etc., which will become more and more important in the future. Further investigations are currently under way supporting our thesis: An evaluation paper, comparing amongst others ERA-40 2 m air-temperatures with the ENSEMBLES RCM hindcast runs driven by ERA-40, is in preparation for KLIWAS. It indicates that the higher resolution of the regional models improves the 2 m air temperature results near the coasts. I.e., ERA-40 shows colder values in wintertime compared to the RCMs, too. Further steps have been undertaken by Saha et al. (2010) with the new NCEP *Climate Forecast System Reanalysis* (CFSR), using a global coupled atmosphere-ocean-land surface-sea ice system with an output of an hourly time resolution and a horizontal resolution of 0.5°. Validations in the same way as done here may hint on a better way of the development of a regional product.

Also, the development of a regional coupled Atmosphere-Ocean reanalysis has been discussed and might be implemented in future studies. At the moment, the *Hans-Ertel Centre for Weather Research* (HErZ) is conducting a retrospective analysis of regional climate at the *Meteorological Institute, University of Bonn* (MIUB) and the *Institute for Geophysics and Meteorology, University of Cologne* (IGMK). This regional reanalysis is based on the COSMO-EU(DE) model with a horizontal grid spacing of 7(2.8) km resolution, as it is in operational use at DWD in the forecasting model suite. A two step nesting, using ERA-Interim (Simmons et al., 2006) as boundary conditions, is performed. Two main periods of the reanalysis will focus on (1) a comparably short time frame of 5 years (2007–2011) with the maximum amount of observational data, and (2) the past decades (1982–2011) with a reduced data basis, in order to aim at more homogeneous time series than typically available in long-term reanalyses (Keller et al., 2012). First results will be available soon.

Our analysis clearly indicates the importance of in-situ observations over the sea to serve as validation set not only for reanalyses, but also for global and regional climate models. Even a sparse observational data base can help to the discovery and definition of problems in the models, if one accounts for the sampling bias. Nevertheless, pseudobiases and measurement errors are still

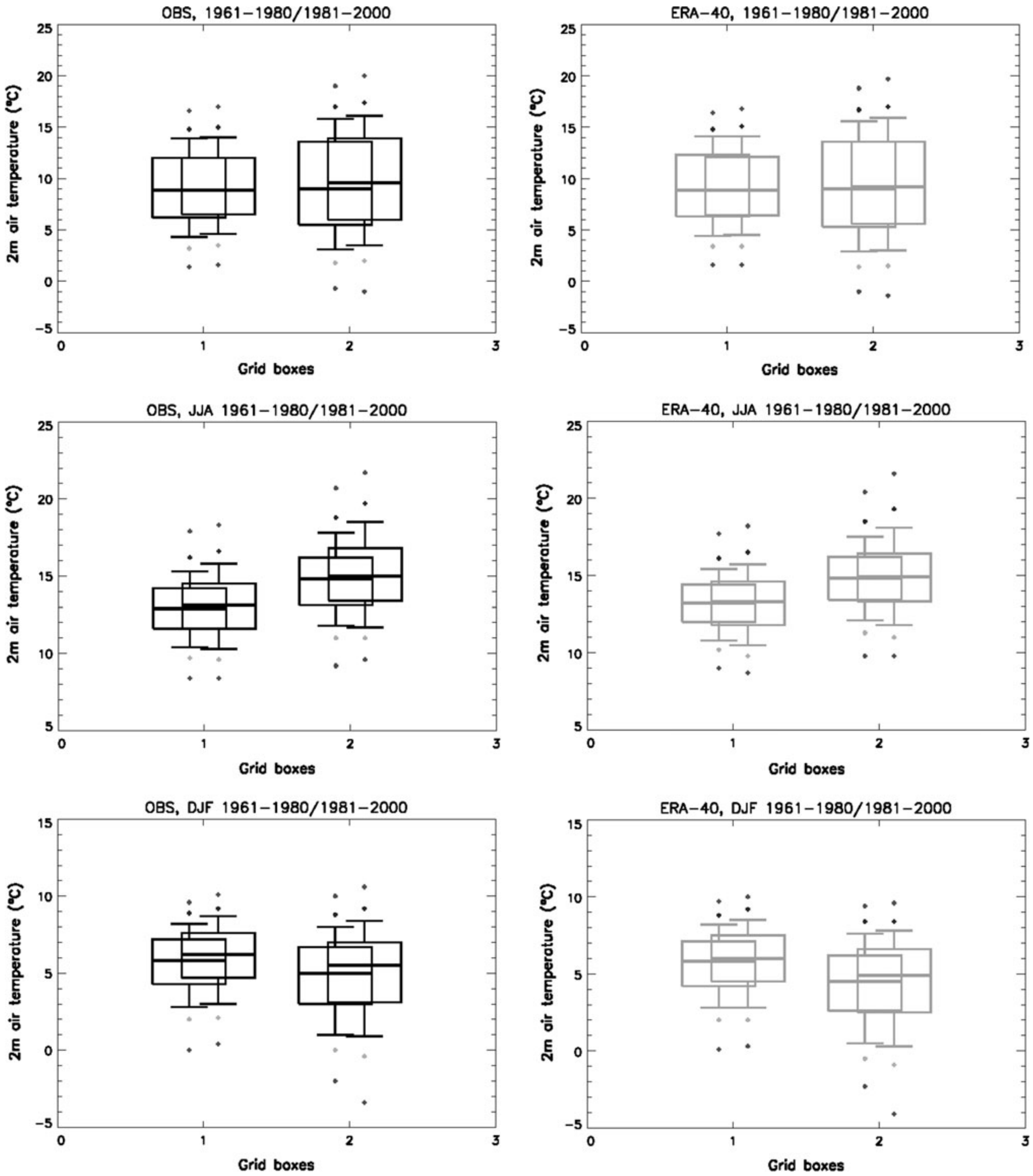

Figure 5: Box plots of 2 m air temperatures from measurements (left) and ERA-40 (right) for the time periods 1961–1980 and 1981–2000 (first row), and the respective summer (second row) and winter seasons (third row) for the North Sea Boxes 1 & 2. Median values and percentiles as described in Fig. 3.

pronounced and have to be accounted for. Improvements could be made by further analyses including an independent data set for triple collocation error modelling. Finally, measurements have to be continued on a high quality basis to ensure the availability of a reference data base in the future.

Acknowledgments

The BSH/DWD KLIWAS team, and a special thanks to Wolfgang Gloeden and his staff at Seewetteramt Hamburg for the help with the preparations of the GZS database and its system.

References

ALEXANDER, L.V., P. UOTILA, N. NICHOLLS, A. LYNCH, 2009: A new daily pressure dataset for australia and its application to the assessment of changes in synoptic patterns during the last century. – J. Climate **23**, 1111–1126. doi: 10.1175/2009/JCLI2972.1.

BERRY, D.I., E.C. KENT, P.K. TAYLOR, 2004: An analytical model of heating errors in marine air temperatures from ships. – J. Atmos. Ocean. Technol. **21**, 1198–1215.

CARDONE, V.J., J.G. GREENWOOD, M.A. CANE, 1990: On trends in historical marine wind data. – J. Climate **3**, 113–127.

GANDLIN, L.S., 1987: Complex quality control of meteorological observations. – Mon. Wea. Rev. **116**, 1137–1156.

GLOEDEN, W., 2011: HISTOR: DWD's project to make historical data available. – Deutscher Wetterdienst, Seewetteramt, Hamburg. ftp://ftp.wmo.int/Documents/PublicWeb/amp/mmop/documents/JCOMM-TR/J-TR-59-MARCDAT-III/posters/Pos12-Gloeden-Histor.pdf.

GULEV, S.K., V. GRIGORIEVA, A. STERL, D. WOLF, 2003: Assessment of the reliability of wave observations from voluntary observing ships: Insights from the validation of a global wind wave climatology based on voluntary observing ship data. – J. Geophys. Res. **180**, 3236. doi: 10.1029/2002JC001437.

GULEV, S.K., T. JUNG, E. RUPRECHT, 2007: Estimation of the impact of sampling errors in the vos observations on air-sea fluxes. part i: Uncertainties in climate means. – J. Climate **20**, 279–301. doi: 10.1175/JCLI4010.1.

HAYASHI, S., 1974: Some problems in marine meteorological observations, particularly of pressure and air temperature. – J. Meteor. Res. **26**, 84–87.

HAYLOCK, M.R., N. HOFSTRA, A.M.G. KLEINTANK, E.J. KLOK, P.D. JONES, M. NEW, 2008: A european daily high-resolution gridded dataset of surface temperature and precipitation for 1950–2006. – J. Geophys. Res. **113**, D20119. doi: 10.1029/2008JDO010101.

HOEFLICH, O., H.-H. MEISSNER, L. HOFFMANN, 1975: Beschreibung eines EDV-Programmes zur Überprüfung maritim-meteorologischer Beobachtungen von Handelsschiffen. – Deutscher Wetterdienst, Seewetteramt, Einzelveröffentlichungen Sonderheft 2.

ISHII, M., M. KIMOTO, M. KACHI, 2003: Historical ocean subsurface temperature analysis with error estimates. – Mon. Wea. Rev. **131**, 51–73.

ISHII, M., A. SHOUJI, S. SUGIMOTO, T. MATSUMOTO, 2005: Objective analyses of sea-surface temperature and marine meteorological variables for the 20th century using icoads and the kobe collection. – Int. J. Climatol. **25**, 865–879. doi: 10.1002/joc.1169.

JONES, P.D., T.D. DAVIES, D.H. LISTER, V. SLONOSKY, T. JONSSON, L. BARRING, P. JONSSON, P. MAHARAS, F. KOLYVA-MACHERA, M. BARRIENDOS, J. MARTIN-VIDE, R. RODRIGUEZ, M.J. ALCOFORADO, H. WANNER, C. PFISTER, J. LUTERBACHER, R. RICKLI, E. SCHUEPBACH, E. KAAS, T. SCHMITH, J. JACOBEIT, C. BECK, 1999: Monthly mean pressure reconstructions for Europe for the 1780–1995 period. – Int. J. Climatol. **19**, 347–364.

KELLER, J.D., C. OHLWEIN, P. FRIEDERICHS, A. HENSE, S. CREWELL, C. WOSNITZA, I. PSCHEIDT, S. KNEIFEL, S. REDL, S. STENKE, 2012: High Resolution Regional Reanalysis for Europe and Germany. 4th World Climate Research Programme. – International Conference on Reanalysis, 7–11 May 2012, Silver Spring, Maryland, USA.

KENT, E.C., D.I. BERRY, 2005: Quantifying random measurement errors in voluntary observing ship's meteorological observations. – Int. J. Climatol. **25**, 843–856. doi: 10.1002/joc.1167.

KENT, E.C., S.D. WOODRUFF, D.I. BERRY, 2007: Metadata from WMO publication no. 47 and an assessment of voluntary observing ship observation heights in ICOADS. – J. Atmos. Ocean. Technol. **24**, 214–234. doi: 10.1175/JTECH1949.1.

KLOK, P.D., A.M.G. KLEINTANK, 2009: Updatet and extended european dataset of daily climate observations. – Int. J. Climatol. **29**, 1182–1191.

KUETTEL, M., E. XOPLAKI, D. GALLEGO, J. LUTERBACHER, R. GARCIA-HERRERA, R. ALLEN, M. BARRIENDOS, P.D. JONES, D. WHEELER, H. WANNER, 2009: The importance of ship log data: Reconstructing north atlantic, euorpean and mediterranean sea level pressure fields back to 1750. – Climate Dynam. **34**, 1115–1128. doi: 10.1007/s00382-009-0577-9.

LORENZ, P., D. JACOB, 2010: Validation of temperature trends in the ensembles regional climate model runs driven by era-40. – Climate Res. **44**, 146–177. doi: 10.1007/s00382-003-0322-8.

SAHA, S., S. MOORTHI, H.-L. PAN, X. WU, J. WANG, S. NADIGA, P. TRIPP, R. KISTLER, J. WOOLLEN, D. BEHRINGER, H. LIU, D. STOKES, R. GRUMBINE, G. GAYNO, J. WANG, Y.-T. HOU, H.-T. CHUANG, H.-M.H. JUANG, J. SELA, M. IREDELL, R. TREADON, D. KLEIST, VAN P. DELST, D. KEYSER, J. DERBER, M. EK, J. MENG, H. WEI, R. YANG, S. LORD, VAN DEN H. DOOL, A. KUMAR, W. WANG, C. LONG, M. CHELLIAH, Y. XUE, B. HUANG, J.-K. SCHEMM, W. EBISUZAKI, R. LIN, P. XIE, M. CHEN, S. ZHOU, W. HIGGINS, C.-Z. ZOU, Q. LIU, Y. CHEN, Y. HAN, L. CUCURULL, R.W. REYNOLDS, G. RUTLEDGE, M. GOLDBERG, 2010: The NCEP climate forecast system reanalysis. – Bull. Amer. Meteor. Soc. **91**, 1015–1057. doi: 10.1175/2010BAMS3001.1.

SCHRUM, C., U. HUEBNER, D. JACOB, R. PODZUM, 2003: A coupled atmosphere/ice/ocean model for the North Sea and the Baltic Sea. – Climate Dynam. **21**, 133–151. doi: 10.1007/s00382-003-0322-8.

SIMMONS, A.J., S.M. UPPALA, D. DEE, S. KOBAYASHI, 2006: ERA-Interim: New ECMWF reanalysis products from 1989 onwards. – ECMWF Newsletter **110**, 26–35.

STOFFELEN, A., 1998: Toward the true near-surface wind speed: Error modeling and calibration using triple collocation. – J. Geophys. Res. **103**, 7755–7766.

UPPALA, S.M., P.W. KALLBERG, A.J. SIMMONS, U. ANDRAE, V.D.C. BECHTOLD, M. FIORINO, J.K. GIBSON, J. HASELER, A. HERNANDEZ, G.A. KELLY, X. LI, K. ONOGI, S. SAARINEN, R.P. ALLEN, E. ANDERSSON, K. ARPE, M.A. BALMASEDA, A.C.M. BELJAARS, L. VANDEBERG,

J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethoff, M. Dragosavac, M. Fisher, M. Fuentes, S. Hagemann, E. Holm, B.J. Hoskins, L. Isaksen, P.A.E.M. Janssen, R. Jenne, A.P. McNally, A.-F. Mahfouf, J.-J. Morcrette, N.A. Rayner, R.W. Saunders, P. Simon, A. Sterl, K.E. Trenberth, A. Untch, D. Vasiljevic, P. Viterbo, J. Woollen, 2005: The ERA-40 re-analysis. – Quart. J. Roy. Meteor. Soc. **131**, 2961–3012. doi: 10.1256/qj.04.176.

van den Linden, P., J.F.B. Mitchell, 2009: ENSEMBLES: Climate Change and its Impacts: Summary of research and results from the ENSEMBLES project. – MET Office Hadley Centre, FritzRoy Road, Exeter EX1 3 PB, UK, 160 pp.

WMO, 1994: International list of selected, supplementary and auxiliary ships. – WMO Report No. 47, WMO, Geneva.

Woodruff, S.D., S.J. Worley, S.J. Lubker, Z. Ji, J.E. Freeman, D.I. Berry, P. Brohan, E.C. Kent, R.W. Reynolds, S.R. Smith, C. Wilkinson, 2011: ICOADS release 2.5: extensions and enhancements to the surface marine meteorological archive. – Int. J. Climatol. **31**, 951–967. doi: 10.1002/joc.2103.

Worley, S.J., S.D. Woodruff, R.W. Reynolds, S.J. Lubker, N. Lott, 2005: ICOADS release 2.1 data and products. – Int. J. Climatol. **25**, 823–842.