

Estimation of the Impact of Sampling Errors in the VOS Observations on Air–Sea Fluxes. Part II: Impact on Trends and Interannual Variability

SERGEY GULEV

P. P. Shirshov Institute of Oceanology, Moscow, Russia, and IFM-GEOMAR, Kiel, Germany

THOMAS JUNG

ECMWF, Reading, United Kingdom

EBERHARD RUPRECHT

IFM-GEOMAR, Kiel, Germany

(Manuscript received 28 July 2005, in final form 2 June 2006)

ABSTRACT

Using the same approach as in Part I, here it is shown how sampling problems in voluntary observing ship (VOS) data affect conclusions about interannual variations and secular changes of surface heat fluxes. The largest uncertainties in linear trend estimates are found in relatively poorly sampled regions like the high-latitude North Atlantic and North Pacific as well as the Southern Ocean, where trends can locally show opposite signs when computed from the regularly sampled and undersampled data. Spatial patterns of shorter-period interannual variability, quantified through the EOF analysis, also show remarkable differences between the regularly sampled and undersampled flux datasets in the Labrador Sea and northwest Pacific. In particular, it is shown that in the Labrador Sea region, in contrast to regularly sampled NCEP–NCAR reanalysis fluxes, VOS-like sampled NCEP–NCAR reanalysis fluxes neither show significant interannual variability nor significant trends. These regions, although quite localized covering small parts of the globe, play a crucial role for the coupled atmosphere–ocean system. In the Labrador Sea, for instance, interannual and decadal-scale changes of the surface net heat fluxes are known to affect oceanic convection and, thus, the meridional overturning circulation of the Atlantic Ocean. From a discussion of current atmospheric data assimilation systems it is argued that in poorly sampled regions reanalysis products are superior to VOS-based products for studying interannual and interdecadal variations of atmosphere–ocean interaction. In well-sampled regions, on the other hand, conclusions about surface heat flux variations are relatively insensitive to the choice of the flux products used (VOS versus reanalysis data). The results are confirmed for two different datasets, that is, ECMWF 40-yr Re-Analysis (ERA-40) data and seasonal integrations with a recent version of the ECMWF model in which no actual data were assimilated.

1. Introduction

Understanding and predicting climate variations from interannual to interdecadal scales is one of the main current challenges in climate research. On the one hand, Bjerknes (1964) showed that atmospheric variations are capable of altering the ocean circulation. On the other hand, there is also evidence suggesting that

SST anomalies influence the atmospheric circulation (e.g., Rodwell et al. 1999), although this topic is more controversial, at least in the extratropics. The possibility of two-way atmosphere–ocean interaction is very attractive because it implies that climate variations may to some extent be predictable. In this context surface heat fluxes play a crucial role, since it is through the fluxes at the sea surface that the atmosphere and ocean communicate.

There are different ways to study the variability of air–sea interaction. Coupled models have been widely used since, among others, they can provide relatively long time series of the characteristics of the coupled

Corresponding author address: Sergey Gulev, P. P. Shirshov Institute of Oceanology, RAS, 36 Nakhimovsky Ave., 117851 Moscow, Russia.
E-mail: gul@sail.msk.ru

climate system without any gaps. Recently, the availability of reanalysis products (Kalnay et al. 1996; Kistler et al. 2001; Uppala et al. 2005) has also attracted widespread interest. Reanalyses are being carried out using sophisticated atmospheric data assimilation systems, which combine model estimates of the state of the atmosphere (so-called first guesses) with observational data in some optimal way (e.g., Kalnay 2003). The extent to which reanalysis data are constrained by observations depends on the parameter being considered. For example, geopotential height fields are largely determined by the observations. On the other hand, surface fluxes are determined from short-range forecasts and, hence, are more model dependent. Therefore, observational data remain crucial for the validation of both coupled and uncoupled model runs as well as reanalysis products. Surface fluxes from reanalyses data were extensively used for the model experiments targeted on the diagnostics of climate variability of the ocean circulation (Eden and Willebrand 2001; Eden and Jung 2001; Gulev et al. 2003; Beismann and Barnier 2004).

In many studies, based on voluntary observing ship (VOS) data, analysis of climate signals associated with the ocean was quantified in terms of SST anomalies (e.g., Deser and Blackmon 1993; Kushnir 1994). Diagnosis of the long-term variability using surface sea–air fluxes may be even more informative for the description of changes at the sea–air interface. However, any attempts to investigate climate variability in the VOS-derived surface ocean–atmosphere fluxes (e.g., Cayan 1992a,b; Gulev 1995; Iwasaka and Wallace 1995; Josey and Marsh 2005) are always influenced by uncertainties inherent to long sea–air flux time series, many of which are of time-dependent nature. Differences in results of ocean model experiments driven by reanalyses and VOS surface flux anomalies (e.g., Häkkinen 1999; Eden and Willebrand 2001; Gulev et al. 2003) may originate not only from differences in the model formulation used, but also from the differences in the forcing functions. Uncertainties associated with bulk parameterizations are not time dependent and the use of different turbulent or radiation schemes usually has rather little impact on the qualitative spatial structure of major variability patterns of surface heat flux fields. On the other hand, the uncertainties associated with variable corrections and uncertainties associated with inhomogeneous sampling may influence estimates of climate variability. For instance, the time dependence of the ratio between anemometer measurements and Beaufort estimates or growing ship size may result in artificial secular tendencies in the wind speed and therefore in surface turbulent fluxes (e.g., Peterson and Hasse 1987; Cardone et

al. 1990; Lindau et al. 1990; Lindau 2003). Sampling inhomogeneity, whose impact on climatological flux estimates was addressed in Gulev et al. (2007, hereafter Part I), may also be a crucial factor when air–sea flux variability is investigated using VOS data. This issue will be investigated in this study.

In Part I it has been shown that in some areas—for example, the Labrador Sea, the northern North Atlantic, and the Southern Ocean—the relatively poor sampling in VOS products is a significant source of uncertainties of monthly mean surface flux estimates (Gulev et al. 2007). Uncertainty estimates were based on the subsampling of 6-hourly surface variables from the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis and the fluxes computed from these variables using bulk formulas. For this purpose for each month during the period 1948–2002 and for every grid box the actual sampling was determined from International Comprehensive Ocean–Atmosphere Data Set (ICOADS) archives (Worley et al. 2005). Then, monthly mean values of surface variables and fluxes have been derived from 6-hourly NCEP–NCAR data using the *actual* VOS-like sampling (the same number of samples was used at exactly the same times when VOS observations were available) and *random* VOS-like sampling (the same number of samples as available from VOS data was used, but the times were picked randomly within the particular month). Hereafter, we shall use the abbreviation FULL for the full 6-hourly NCEP–NCAR surface fluxes (our “truth”), REAL for the actual VOS-like sampling product, and RAND for the random VOS-like sampling fields. Additionally we estimate uncertainties associated with the spatial interpolation of fluxes into fully unsampled grid cells.

In Part I it was found that the uncertainty of monthly mean fluxes based on REAL was generally larger than that of RAND. This can largely be explained by the fact that VOS observations are usually taken for limited periods of time only (e.g., the passage of one ship through one grid box), which leads to the availability of relatively fewer independent observations in REAL compared to RAND. For the period 1992–2001 the same sampling procedure was applied to two additional datasets, that is, European Centre for Medium-Range Weather Forecasts (ECMWF) 40-yr Re-Analysis (ERA-40) data and seasonal integration with the ECMWF model (using the same resolution as for ERA-40) in which no data were assimilated (ECF, hereafter). Conclusions from these computations were consistent with those drawn from the NCEP–NCAR-based flux fields. Spatial patterns of sampling uncertainties qualitatively agree well in all three products, but may exhibit

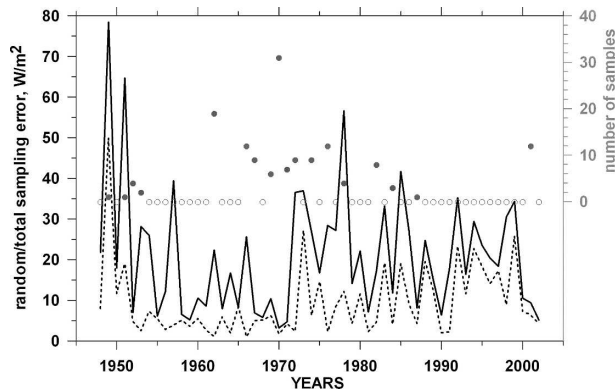


FIG. 1. Time variability of the January random (dashed line) and total (solid line) sampling errors in the latent heat flux and the number of VOS observations for one 2° box in the Labrador Sea (filled circles stand for the existence of at least one observation per month; open circles stand for the absence of samples).

quantitative differences, due to somewhat different magnitudes of intramonthly synoptic variability in different NWP products.

Figure 2 of Part I shows a striking example of the time-dependent nature of sampling uncertainties in surface fluxes [see also Chang (2005) for a discussion of mean sea level pressure from ICOADS]. First, time-dependent sampling uncertainties arise from changes in sampling density. Second, for some grid cells there are periods when observations were available (they are influenced by sampling errors), and there are periods for which there were no observations at all and the values were provided by spatial interpolation (they are influenced by interpolation or extrapolation errors). Figure 1 shows the time series for January of random [$\delta_f(Q_e)$] and total [$\chi_f(Q_e)$] sampling errors (see Part I for definitions) for latent heat fluxes along with the number of VOS observations for the 2° box centered at 59°N , 53°E in a poorly sampled area of the Labrador Sea. Clearly, there is strong interannual variability in the magnitude of both $\chi_f(Q_e)$ and $\delta_f(Q_e)$ and this magnitude is comparable to the magnitude of interannual variability of the fluxes themselves. Also, there are periods (e.g., from the mid-1960s to mid-1970s) during which both random and total sampling errors originate from the inadequate sampling and there are other periods (e.g., during the 1990s) characterized by the absence of observations, when sampling uncertainty results from the spatial interpolation and extrapolation procedures.

By using the method outlined in Part I, here we investigate how the temporally and spatially inhomogeneous sampling of VOS data affects estimates of climate variability of surface heat fluxes. As will become clear, the results of this study have important implications for future studies on the nature of climate vari-

ability of surface heat fluxes. The paper is organized as follows. In the next two sections the results are presented. We start with a discussion of the sensitivity of secular changes (or trends) to sampling and continue with the analysis of sampling impact on interannual sea–air flux variability. Finally, the results are summarized and discussed.

2. Impact of sampling uncertainties on secular tendencies in air–sea fluxes

In this section we analyze the influence of sampling on linear trends of surface heat fluxes. To this end, first we estimated linear trends in different sea–air flux components from the regularly sampled (FULL) and undersampled (REAL and RAND) time series for the period from 1948 to 2002. Trends in variables taken from NCEP–NCAR reanalyses or in products derived from reanalyses variables are known to be influenced by the inhomogeneities of data assimilation input, although the data assimilation system itself has been frozen during the production period. Time-dependent uncertainties originate from the sharp increase of the assimilated information with the availability of satellite data and resulting strong temporal inhomogeneities of the assimilated information, especially in the Southern Hemisphere. White (2000) and Sterl (2004) found strong temporal inconsistencies of reanalyses of NCEP and ECMWF, associated with inhomogeneity of data assimilation input. Bengtsson et al. (2004b) showed that changes in data assimilation input can also affect trends in global quantities, such as global mean temperature, integrated water vapor, and kinetic energy. However, in this study, it is not our intention to make new inferences about secular changes in the fluxes computed from the reanalysis variables. Rather, the aim is to quantify the uncertainties inherent in estimates of secular changes due to sampling inhomogeneity.

Figure 2 shows estimates of linear trends together with their statistical significance according to a Student's t test for the winter [January–March (JFM)] and summer [July–September (JAS)] seasons in the sensible plus latent turbulent fluxes Q_{he} . Seasonal time series were derived from the monthly values of all three products (FULL, RAND, and REAL). Trends in the heat fluxes derived from the regularly sampled time series (Fig. 2a) are significantly positive in the subpolar latitudes of the North Atlantic, the northeast subtropical and midlatitudinal Pacific, as well as in tropical and subtropical areas of the Southern Hemisphere, with the highest values of about $3\text{--}6\text{ W m}^{-2}$ per decade. Significantly negative trends are observed in the North Atlantic subtropics, particularly over the Gulf Stream

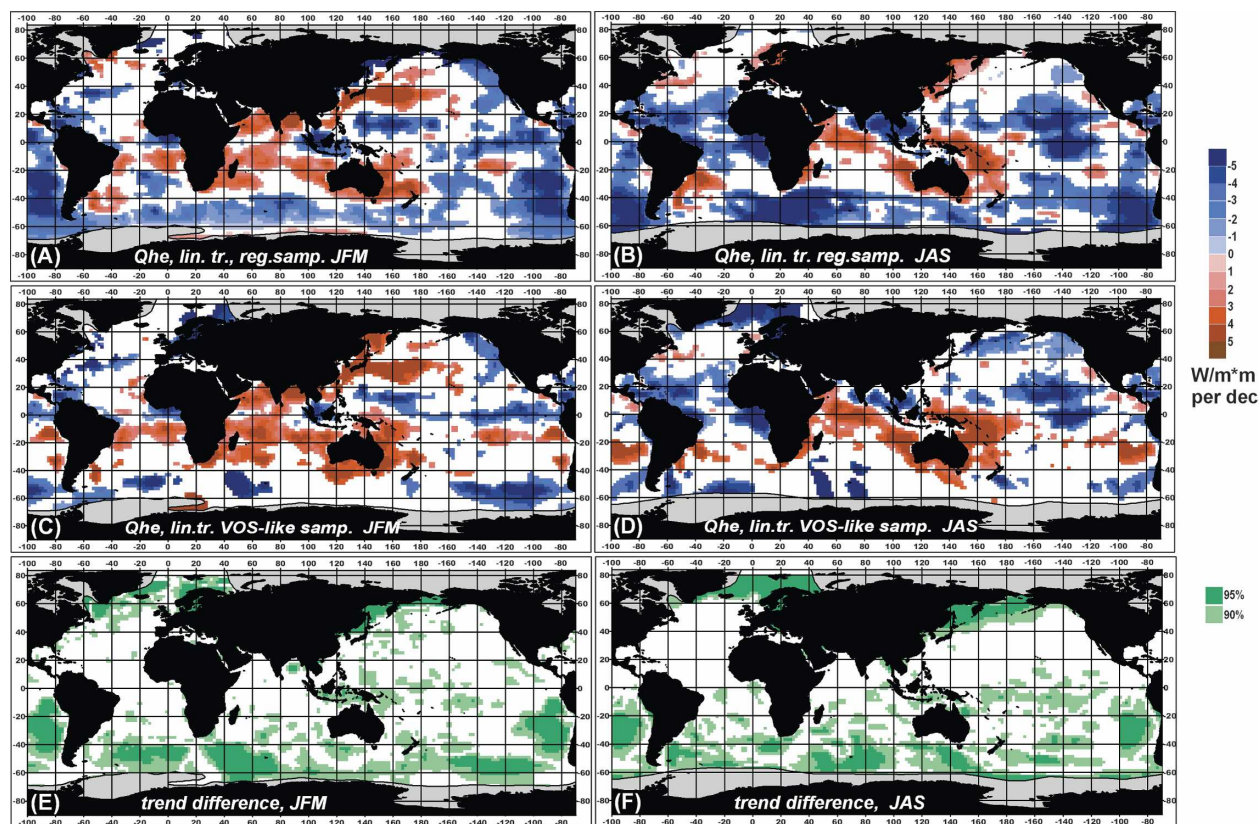


FIG. 2. Linear trends (W m^{-2} per decade) in the sensible plus latent heat flux, estimated from (a), (b) the FULL product and (c), (d) the REAL product along with (e), (f) the significance of the difference between the trends for (a), (c), (e) boreal winter (JFM) and (b), (d), (f) summer boreal (JAS). In (a)–(d) only trends significant at 95% level (t test) are shown. In (e), (f) only differences significant at the 90% and 95% level are shown.

area, in the northeast Pacific, and in the Southern Ocean, where they can be largely attributed to the changes in data assimilation input during the last two decades. During the July–September period (Fig. 2b) the situation in the Southern Hemisphere is quite comparable to that for boreal winter. Late summertime trends are quite different in the Northern Hemisphere, however, with weakly positive significant changes being found in the subpolar latitudes of the North Atlantic; the pattern of the strong positive trends in the northwest Pacific is absent.

Comparison of the spatial structure of the estimated trend patterns in Figs. 2a,b with those derived from the fluxes computed from the VOS-like sampled individual variables (Figs. 2c,d), despite general similarity, especially in the tropical regions, show local differences, first of all in the subpolar North Atlantic and in the Southern Ocean. In the Labrador Sea, for example, the fluxes computed from the regularly sampled variables give positive winter trends of up to 5 W m^{-2} per decade. However, when subsampled variables are used the trends become significantly negative and amount to

$-3\text{--}4 \text{ W m}^{-2}$ per decade. Remarkably different trend patterns are observed in summer in the marginal seas of the northwest Pacific, where regularly sampled time series show significantly positive trends and VOS-like sampled time series show no indication of any secular changes. Locally large differences in the trend estimates are also observed in the Southern Ocean, where the continuous area of negative trends disappears, being largely influenced by the interpolation in fully unsampled grid cells.

Figures 2e,f show the regions where the differences between the trend estimates obtained from the regularly sampled and VOS-like sampled time series (REAL) are significant at the 90% and 95% significance level. The strongest differences are observed in the subpolar latitudes of the Northern Hemisphere and in the Southern Ocean, reflecting a substantial influence of sampling on the estimated trends in poorly observed regions. Comparison of the trend patterns in the fluxes derived from RAND and REAL procedures of subsampling (not shown) shows much closer comparability of the RAND with FULL than in the case of

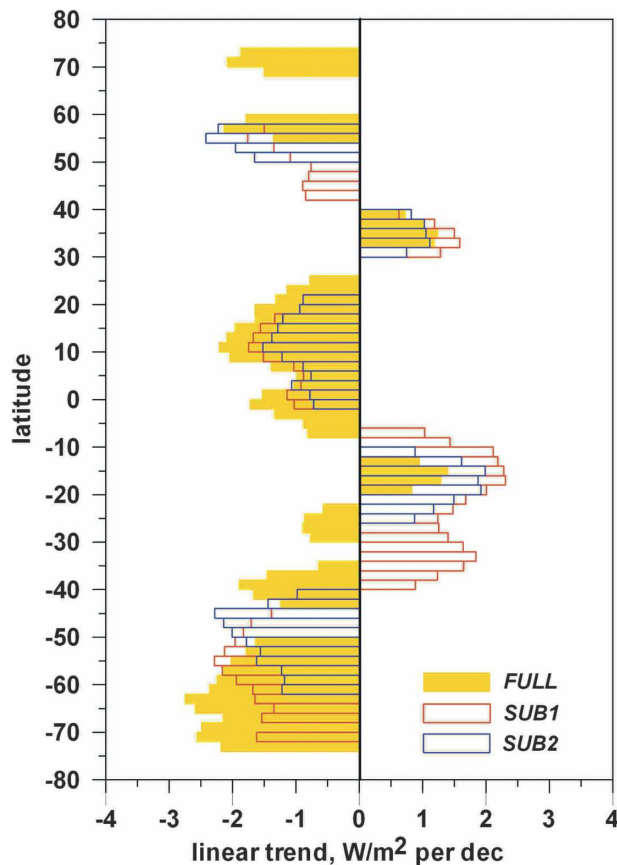


FIG. 3. Linear trends (W m^{-2} per decade) in the zonally averaged net heat flux, estimated from the fluxes computed from the regularly sampled parameters (FULL; yellow bars), the actual VOS-like sampled parameters (REAL; red bars), and from the randomly sampled parameters (RAND; blue bars).

REAL in the Northern Hemisphere mid- and subpolar latitudes during winter. Typically, trend estimates derived from RAND and FULL show qualitatively similar spatial patterns with a smaller magnitude of trends for RAND. Effects shown in Fig. 2 for sensible plus latent turbulent fluxes are also evident for the radiative fluxes (no figure shown). Relative differences between the trend estimates in LW and SW radiation, derived from FULL and REAL, are smaller than in the turbulent fluxes in agreement with the smaller level of sampling errors in radiative fluxes.

Figure 3 shows estimates of linear trends in zonally averaged annual mean net heat fluxes, derived from FULL, REAL, and RAND products. In the Northern Hemisphere the largest disagreement occurs in the mid-latitudes, where REAL and RAND show significantly negative changes and no significant trends were identified in FULL. In the Southern Hemisphere Tropics and subtropics both REAL and RAND indicate positive trends, while FULL shows weak positive trends be-

tween 10° and 20°S and the trends of the opposite sign south of 20°S . In the mid- and subpolar latitudes of the Southern Hemisphere the trends in REAL and RAND are in a qualitative agreement with those computed from FULL; however, the values are considerably smaller, especially in RAND. Thus, sampling may have impacts even on the secular tendencies in zonally averaged climatological mean flux estimates.

A similar analysis was performed with the subsampled data from ERA-40. Comparison of the linear trends for the period 1958–2001 (no figure shown) indicates that the trend estimates in FULL and REAL products derived from ERA-40 data disagree in the Labrador Sea and in the northwestern Pacific as they do in the case of the use of NCEP–NCAR data. The Labrador Sea trends in turbulent fluxes change sign from $+3$ to -2 W m^{-2} per decade between FULL and REAL, which is quite consistent with the estimates derived from NCEP–NCAR data for the same 44-yr period (1958–2001). Patterns of the linear trend differences between FULL and REAL, derived from ERA-40 in the Southern Ocean, are qualitatively consistent with those for NCEP–NCAR; however, they show about 30% larger differences in the South Pacific and somewhat weaker differences in the South Atlantic. Summarizing, we can conclude that the use of alternative reanalysis (ERA-40) for the period 1958–2001 qualitatively and quantitatively supported findings derived from the NCEP–NCAR reanalysis.

3. Sampling uncertainties influencing interannual variability of surface fluxes

Let us now consider the influence that sampling has on shorter-period interannual variations of surface heat flux fields. We computed correlations between the detrended anomalies of differently sampled net surface heat flux products. Figure 4 shows maps of the correlation coefficients between the net heat fluxes computed from the regularly sampled variables (FULL) and those derived from subsampled fields (REAL and RAND) for the winter and summer season. Similar results are obtained for all flux components (not shown). They are consistent with the results of Sterl (2001), who analyzed correlations between reanalyses and VOS-based flux products. The first thing to notice is that in both seasons highest correlations between subsampled (RAND and REAL) and regularly sampled (FULL) net surface heat fluxes are found in relatively well sampled regions like, for example, the central North Atlantic region. In both seasons relatively low correlations (below the significance level of 0.4) are found in

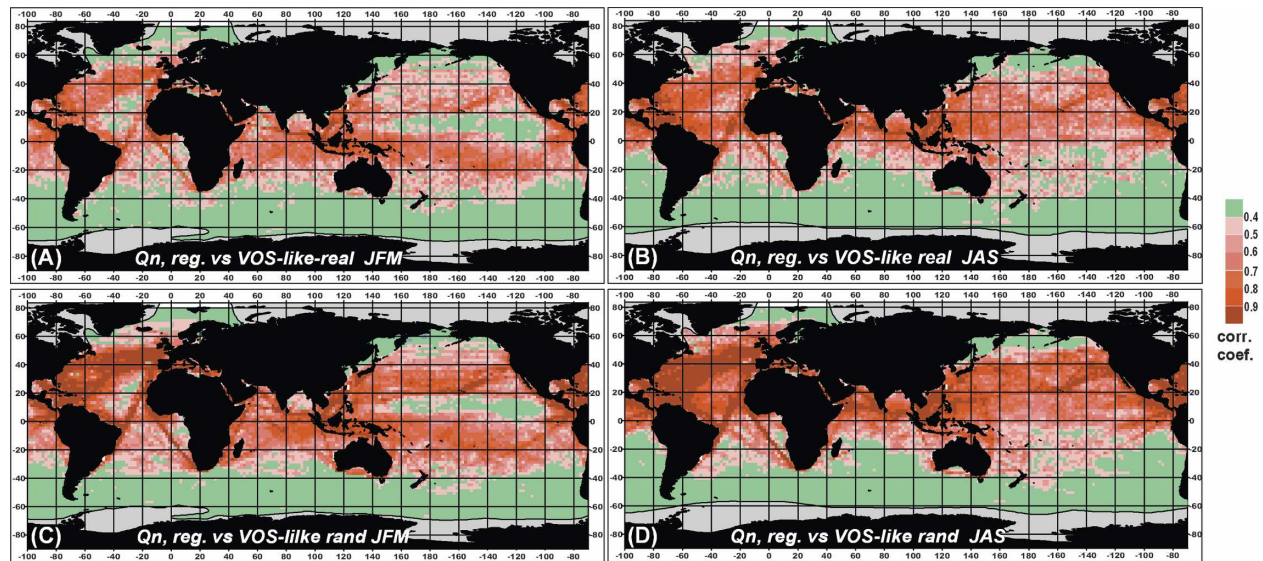


FIG. 4. Correlation coefficients between (a), (b) the net flux computed from the regularly sampled variables (FULL) and VOS-like sampled variables (REAL) as well as (c), (d) between the net flux computed from the regularly sampled variables (FULL) and randomly sampled variables (RAND) for (a), (c) boreal winter and (b), (d) boreal summer derived from the NCEP–NCAR reanalysis for the period 1948–2002.

the high-latitude North Atlantic and North Pacific as well as in the Southern Ocean where sampling is relatively poor. Furthermore, it turns out that subsampling in the tropical Pacific and the subtropical North Atlantic has a larger impact on interannual near-surface heat flux variations during wintertime compared to the summer season. Besides slightly higher sampling density in summer, this difference can also be attributed to the stronger synoptic variability during winter time in these areas. Moreover, it reveals that correlations between FULL and REAL (Figs. 4a,b) are generally lower than those between FULL and RAND (Figs. 4c,d). In the Labrador Sea in summer and in the Greenland–Iceland–Norwegian (GIN) Sea for both seasons random sampling results in correlations of about 0.5–0.6, whereas the actual sampling (REAL) leads to substantially lower correlation coefficients (0.1–0.3). In summary, it can be said that sampling is an issue in the context of temporal characteristics of interannual surface heat flux variations in some of the regions, which are known to play key roles in terms of interannual variability of the coupled atmosphere–ocean system. This is particularly true if the actual sampling is taken into account (REAL).

In Part I we compared the effects of undersampling in NCEP–NCAR reanalyses with those for ERA-40 reanalysis data and seasonal forecasts with the ECMWF model (ECF) for which no data were assimilated. The comparison, which was carried out for the period 1992–2001, is characterized by quite different data assimilated in the two reanalyses (primarily satellites). To

assess the robustness of sampling impact on variability patterns in different NWP products, for boreal winters of the period 1992–2001, we computed correlations between FULL and RAND products derived from NCEP–NCAR, ERA-40, and ECF (Fig. 5). Since seasonal forecasts are involved in the comparison, we only considered the effect of random sampling errors, because ECF cannot reproduce particular synoptic events, although it can quite adequately simulate the magnitudes of intramonthly synoptic variability (Part I). The correlations for NCEP–NCAR fluxes in this period (Fig. 5a) reveal some differences from that derived using the full 55-yr dataset (Fig. 4c), showing reasonably higher correlations (with, however, the 95% significance level being around 0.6 due to the shorter time series used) and reflecting differences in sampling for the two periods. Nevertheless, our major conclusions from the analysis of Fig. 4 are also valid for Fig. 5a—poorly sampled areas in the subpolar latitudes of the Northern Hemisphere and in the Southern Ocean are characterized by very low, nonsignificant correlation. Figure 5b, derived from the ERA-40 reanalysis, shows very good similarity with that for NCEP–NCAR fluxes. Some minor local differences can be noticed in the Southern Ocean where in some areas ERA-40 shows higher correlations. Interestingly, the correlation pattern derived from ECF shows very close comparability with that for ERA-40, implying that the ECMWF model is capable of producing realistic synoptic variability even if no data are assimilated (i.e., the model climate is realistic).

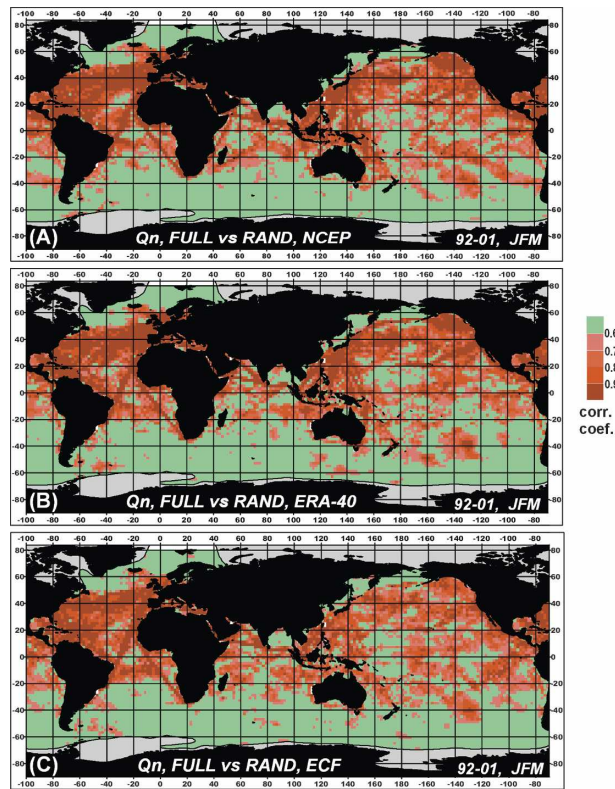


FIG. 5. Correlation coefficients between the net flux computed from the regularly sampled variables (FULL) and randomly sampled variables (RAND) for the boreal winter season during the period 1992–2001 from (a) the NCEP–NCAR reanalysis, (b) the ERA-40 reanalysis, and (c) the ECMWF seasonal forecasts.

A commonly used tool in climate research to describe dominant (in terms of explained variance), spatially covarying patterns in multivariate time series is empirical orthogonal function (EOF) analysis (von Storch and Zwiers 1999). In the following it is shown how sampling affects the leading two EOFs of surface net heat flux anomalies, both in the North Atlantic and North Pacific. We will focus on the winter season, which is characterized by the strongest synoptic and interannual variability and by the poorest sampling in the Northern Hemisphere. Since the ice cover shows interdecadal variations in the North Atlantic and North Pacific and, therefore, can affect the interannual variability, we applied the largest ice extent for the period 1948–2002 as ice mask throughout the whole period. This mask was derived from the Global Sea Ice and SST dataset GISST climatology (Parker et al. 1995) and the ice cover from the NCEP–NCAR reanalysis by setting the particular grid cell to the ice mask when the ice was identified in either dataset.

The leading two EOFs of North Atlantic net surface heat flux anomalies are shown in Fig. 6 for (a), (c) FULL and (b), (d) REAL. The main characteristics of the leading two EOFs are very similar for the two datasets; that is, EOF1 and EOF2 show tripole structures, which are in quadrature (see also Cayan 1992a,b). There are differences, however, which are particularly prominent for EOF1 in the Labrador Sea region, whereas EOF1 for FULL shows anomalously

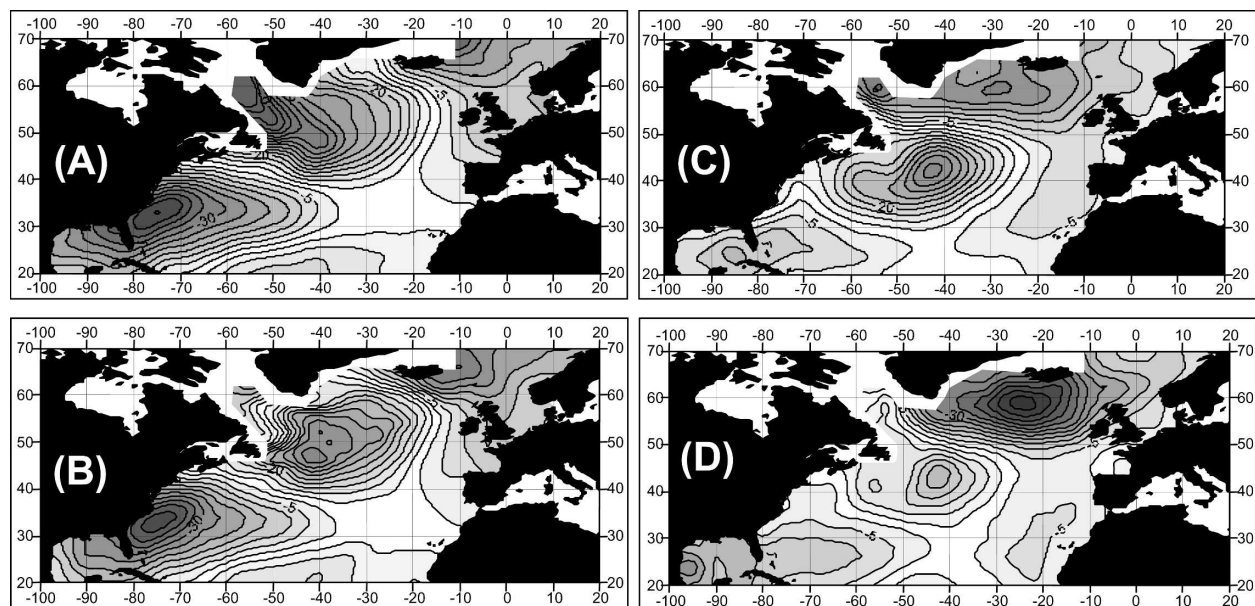


FIG. 6. The (a), (b) first and (c), (d) second EOFs of net heat flux anomalies in the North Atlantic computed from the (a), (c) regularly sampled variables (FULL) and (b), (d) VOS-like sampled variables (REAL).

strong net surface heat fluxes out of the ocean in excess of 70 W m^{-2} , and the net surface heat flux anomalies for REAL suggest that the ocean is subject to weak anomalous heating. These differences are most pronounced for the years after 1973, when the sampling in the Labrador Sea has decreased sharply (not shown). Albeit less pronounced, differences between EOF2 for FULL and REAL are also most prominent in the Labrador Sea region. In FULL the entire Labrador Sea belongs to the northern center of action demonstrating the same sign of anomalies as in the central subpolar Atlantic in contrast to the REAL fluxes, whose second EOF shows in the Labrador Sea the same sign of anomalies as in the southern center of action. Differences of EOF1 between FULL and REAL also show up in terms of the amount of variance explained by the first EOF (39% versus 27% for FULL and REAL, respectively). The EOFs for RAND (not shown) are intermediate between FULL and REAL, though more closely resembling those for REAL. The first EOF of RAND explains 31% of the total variance.

The first principal components (PC1, hereafter), which are associated with the first EOFs of North Atlantic net surface heat flux anomalies are shown in Fig. 7a for FULL and REAL. The correlation between the two time series amounts to 0.94; that is, about 89% of the variance is explained by a linear relationship. Evidently the subsampled dataset (REAL) captures most of the temporal characteristics in a larger-scale context. We also note in passing that the first EOFs and PCs for FULL, REAL, and RAND capture the influence of the North Atlantic Oscillation (NAO). Figure 6a also shows the NAO index as given by Hurrell (1995). The correlation coefficients of PC1 for FULL and REAL with the NAO index amount to 0.72 and 0.63, respectively.

EOF analysis of the anomalies of surface fluxes in the North Pacific (Fig. 8) reveals similar results to those found for the North Atlantic with respect to the impact of sampling on the variability patterns. The first EOFs in all flux products show the southwest–northeast pattern described by Cayan (1992a), Iwasaka and Wallace (1995), and Tanimoto et al. (2003). This pattern accounts for 23% of variance in FULL and for 16% and 17% of variance in REAL and RAND, respectively. As for the North Atlantic, the first normalized PCs (Fig. 7b) for the North Pacific are highly correlated with each other, with a correlation coefficient of 0.86. They are also linked to the North Pacific index (NPI; Trenberth and Hurrell 1994) (Fig. 7b), with correlations amounting to 0.61 and 0.59 for FULL and REAL, respectively. The major difference between the first EOF of FULL on the one hand and REAL and RAND on

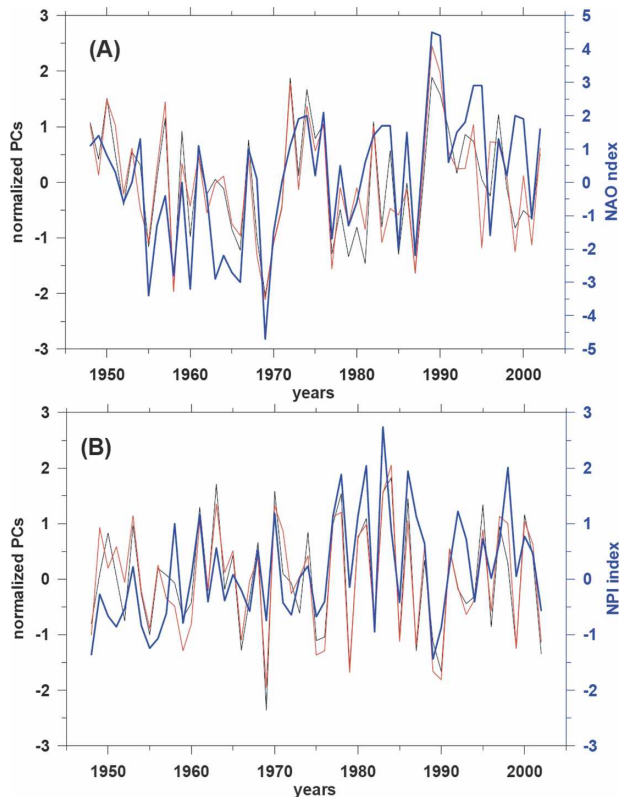


FIG. 7. The first normalized PCs of the net heat flux in the (a) North Atlantic and (b) North Pacific computed from the regularly sampled variables (FULL; black) and VOS-like sampled variables (REAL; red) along with the NAO and NPI indices (blue).

the other is observed in the subpolar northwest Pacific, where FULL shows the same sign of anomalies as in the northeastern center of action, while in the REAL and RAND products the sign of anomalies here is steered by the Kuroshio pattern. Comparison of the second EOFs (Figs. 8c,d) shows pronounced difference in the spatial patterns of fluxes computed from the regularly sampled and undersampled variables in the subpolar northwest Pacific. Correlation coefficients of the second PC of FULL with those of REAL and RAND amount to 0.54 and 0.63, respectively.

EOF analysis performed with the FULL, REAL, and RAND flux products derived from ERA-40 reanalysis data for 1958–2001 (no figure shown) reveals similar (with respect to the sampling impact on variability patterns) conclusions to those drawn from the analysis of NCEP–NCAR data. In the North Atlantic the first EOF of FULL derived from ERA-40 has a spatial structure, which is very similar to that obtained for NCEP–NCAR data. In the Labrador Sea region, however, the associated anomaly is slightly smaller, explaining a smaller amount of the variance. As in the case with NCEP–NCAR, the impact of sampling becomes

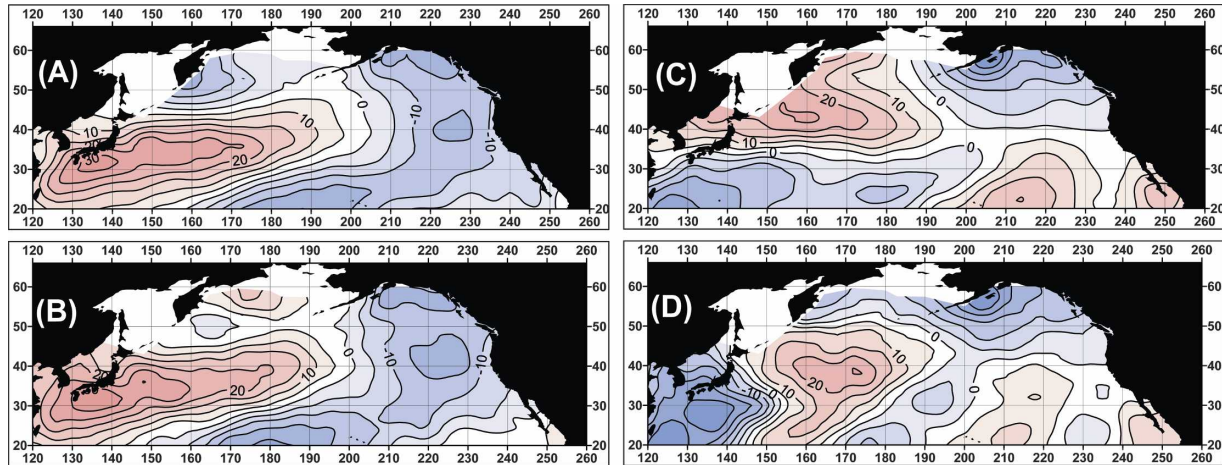


FIG. 8. The (a), (b) first and (c), (d) second EOFs of net heat flux anomalies in the North Pacific computed from (a), (c) the regularly sampled variables (FULL) and (b), (d) VOS-like sampled variables (REAL).

evident in the first EOF and especially noticeable in the second EOF being even more pronounced than for NCEP–NCAR data. In the Pacific, alternatively, sampling impact on the first and second EOFs, although quite evident, is a little bit less pronounced than in NCEP–NCAR. It is important to note that the difference of EOF patterns between NCEP–NCAR and ERA-40 is smaller compared to differences of the EOFs between FULL and REAL for each of the reanalysis products even in the poorly sampled Labrador Sea and northwestern subpolar Pacific.

To summarize the results of the intercomparison of EOFs and PCs obtained from different flux products we computed the correlation between the PCs of surface heat flux anomalies from FULL (reference field) and the projections of the anomalies of REAL and RAND onto the EOFs of the reference field (i.e., quasi-principal components). The correlations between the projections of RAND flux anomalies are somewhat higher than those for the REAL (Table 1). Squared correlation in the North Atlantic is higher than in the North Pacific. Correlations for the first PCs vary from 80% to 90%, implying quite close comparability of the leading modes. However, for the second PCs correlation between the projections decreases to about 60%, implying significant differences in the second modes. Table 1 also shows that standard deviations (std) of the PCs and projections using different flux products are quite different. Projections derived from RAND and REAL show variability, which is typically 10%–15% and 23% smaller than those of the first and second, respectively, PC of FULL anomalies. A corresponding analysis performed for the ERA-40 reanalysis (both RAND and REAL) and ECF (RAND) shows quite similar results, implying the robustness of our conclu-

sions to the use of different NWP products. Again, the first PCs of undersampled fluxes are highly correlated with the FULL product. Correlation coefficients for ECF are about 10% lower than those for ERA-40. Similar to the NCEP–NCAR, both ERA-40 and ECF exhibit noticeable drop of correlation for the second PCs to 54%–70%.

4. Summary and discussion

The influence that temporal and spatial inhomogeneities of the observed sampling density have on the characteristics of estimates of sea–air flux variations has been investigated. It has been found that sampling locally does have a substantial impact on the characteristics of interannual variability of surface fluxes. While the leading two EOFs of surface heat flux variability, derived from the regularly sampled and VOS-like subsampled fields, are comparable with each other over most parts of the North Atlantic and North Pacific, large differences are found in the Labrador Sea region and in the subpolar northwest Pacific. Moreover, a similar analysis for linear trends in surface flux components reveals statistically significant differences in the flux anomalies derived from regularly sampled and from VOS-like subsampled individual parameters in the Southern Ocean and in the subpolar latitudes of the Northern Hemisphere. The largest sampling influence is found during the winter season characterized by generally poorer sampling in comparison to the summer season. The results of this study suggest that in well-sampled regions (e.g., in the Northern Hemisphere midlatitudes and along the major ship routes of the Southern Hemisphere) VOS data can provide reliable estimates of climate variability in air–sea exchanges.

TABLE 1. Results of the projections of the net heat flux anomalies derived using randomly sampled (RAND) and VOS-like sampled (REAL) individual variables onto the EOFs of the “reference” (FULL) net heat flux anomalies for NCEP–NCAR and ERA-40 reanalyses as well as for ECMWF seasonal forecasts (ECF). Left numbers are squared correlations in percent with FULL (regularly sampled) projection, and right numbers are the ratios between the standard deviations of the projected fluxes and the reference field.

Flux product		First EOF		Second EOF	
		Squared correlation (%)	Ratio of standard deviations	Squared correlation (%)	Ratio of standard deviations
North Atlantic					
REAL	NCEP (48–02)	84.8	0.89	61.4	0.81
	ERA-40 (92–01)	92.8	0.90	64.5	0.87
RAND	NCEP (48–02)	89.1	0.93	67.6	0.84
	ERA-40 (92–01)	94.9	0.91	70.1	0.90
	ECF (92–01)	83.4	0.86	61.0	0.81
North Pacific					
REAL	NCEP (48–02)	81.2	0.85	58.7	0.77
	ERA-40 (92–01)	85.3	0.83	64.3	0.81
RAND	NCEP (48–02)	83.8	0.91	62.4	0.82
	ERA-40 (92–01)	88.0	0.89	65.9	0.88
	ECF (92–01)	79.5	0.84	54.6	0.72

This is especially true for the period before the 1950s when VOS data represent the only source of information about air–sea interaction. However, in poorly sampled regions sampling errors may seriously affect the conclusions drawn exclusively from VOS data. Despite many inhomogeneities also inherent in reanalyses data, for poorly sampled areas they represent the most reliable source of information about air–sea flux variability since the International Geophysical Year in 1957/58.

Artifacts in variability patterns associated with inadequate and temporally changing sampling can be higher than those resulting from the uncertainties associated with variable corrections, which represent another possible source of time-dependent uncertainties. For instance, the largest differences in linear trends derived from FULL and REAL (Fig. 2) amount to as much as $5\text{--}7\text{ W m}^{-2}$ per decade. Tentative estimates show that changes in the ratio of anemometer measurements and Beaufort estimates of the wind speed or growing ship size (i.e., increasing height of anemometer measurements) can result in artificial tendencies in turbulent fluxes of about $1\text{--}4\text{ W m}^{-2}$ during the period from 1973 to 2002, for which the WMO-47 metadata are available (Kent et al. 2007). Furthermore, the influence of sampling on interannual variability is highly localized, which is not the case for the effect of variable corrections (Kent et al. 1993; Josey et al. 1999). The impact of changes in observational practices for measuring wind speed is not localized and shows a somewhat stronger magnitude in the midlatitudinal regions, where the contribution of the wind speed to the turbulent fluxes is higher than in the Tropics.

Lindau (2003) quantified the contribution of the uncertainties in individual monthly mean latent heat fluxes to the interannual variability and found that errors in the monthly mean latent fluxes contribute up to 60%–80% of the magnitude of interannual variability. In Fig. 9 we show the interannual standard deviation (std) of the latent heat fluxes computed from the regularly sampled individual variables along with the ratio between the std of fluxes computed from FULL and REAL. In well-sampled areas, where sampling errors are relatively small, the ratio is slightly higher than 1. However, in regions of strong undersampling this ratio may increase to 3–4 with the highest values found in the Southern Ocean. This shows the strong impact that sampling errors have onto the estimated magnitude of interannual variability. Note that the same result is implied by Table 1, quantifying the covarying modes in FULL, REAL, and RAND products. Lindau (2003), by using VOS observations, has been dealing with all sources of uncertainties, including observational errors. In our case we exclusively estimated the effect of sampling uncertainties onto interannual variability. First, in most areas where large values of the ratio are found (Figs. 9c,d), many monthly values were produced by spatial interpolation. Whatever interpolation scheme is used, it tends to reduce the actual magnitude of interannual variability. However, even in the regions where there are samples, three–seven observations per month even from different platforms are not enough to capture the influence of extreme flux values on the monthly mean.

The closest to 1 and smaller than 1 values of the ratio in Fig. 9 were found in the areas of oversampling and

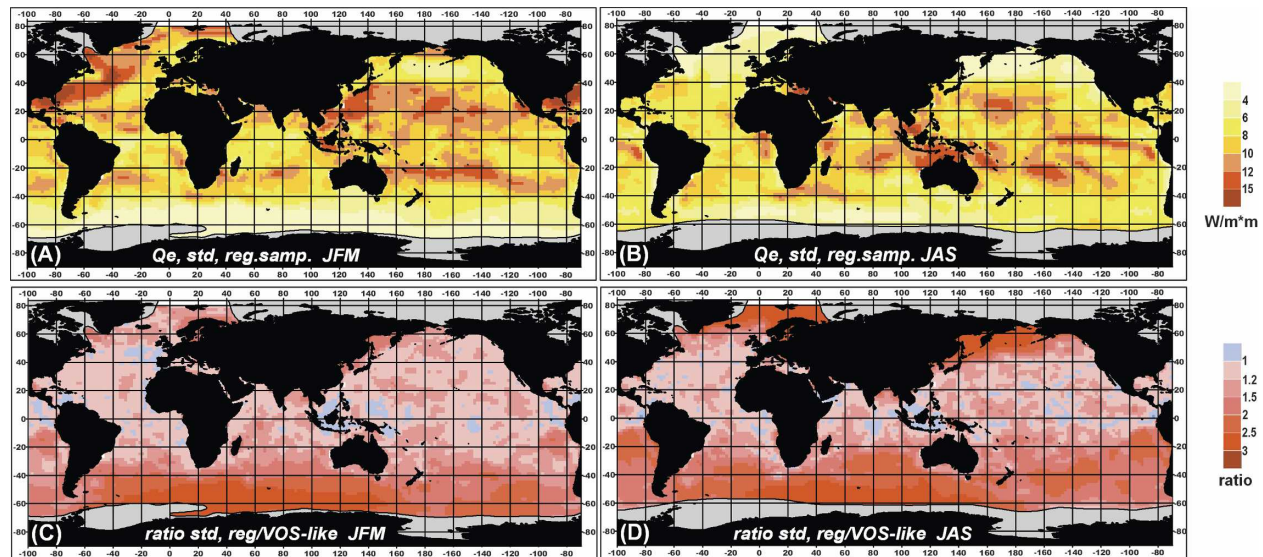


FIG. 9. Interannual standard deviation of the latent heat flux computed from the regularly sampled NCEP-NCAR variables (FULL) for (a) boreal winter and (b) boreal summer as well the ratio between the std computed from regularly sampled (FULL) and VOS-like sampled (REAL) variables for (c) boreal winter and (d) boreal summer.

not undersampling, implying that the true variability is better captured in more densely and evenly sampled regions. In the FULL flux product (Figs. 9a,b) the magnitudes of interannual variability in the Southern Ocean are 1.5–2.5 times smaller than in the Northern Hemisphere midlatitudes, which reflects the differences in synoptic activity and less pronounced temperature and humidity gradients, but can be partly attributed also to the smaller data assimilation input in the Southern Ocean. However, for the recognized VOS products (da Silva et al. 1994; Josey et al. 1999; Lindau 2003) this ratio is considerably smaller, ranging from 2 to 7–10 and implying an underestimation of the magnitude of interannual variability in VOS fluxes over data-sparse areas due to undersampling.

As was pointed out in Part I, the development of new algorithms for the reconstruction of flux anomalies is an important outstanding issue. Since methods of Kaplan et al. (1998, 2000, 2003) and Smith and Reynolds (2003, 2004) are quite skillful for the reconstruction of SLP and SST anomalies, their adaptation to the long-term series of air–sea fluxes would be desirable. The results of this study show, however, that care has to be taken when choosing the dataset used for computing the EOFs (used later on in the reconstruction). Any set of EOFs, which is based on VOS-data only, will suffer from large sampling uncertainties in some key areas as has been demonstrated in this study.

In Part I we have shown that the sampling density for individual VOS variables in ICOADS is typically higher than that of the full sets of variables, required for

the proper computation of surface fluxes (Fig. 1 of Part I). Given the large contribution of drifting buoys in ICOADS during the last decade (Worley et al. 2005), we can expect that sampling of SST, SLP and, to a lesser degree, wind speed is better than the other flux-related parameters. This implies that for some poorly sampled regions the so-called monthly summary trimmed groups (MSTG) products can be used for pilot estimates of variability. MSTG provides 1° and 2° averaged individual parameters derived from all available reports of this parameter and thus is less affected by undersampling. On the other hand, the use of MSTG requires application of the bulk formulas to the monthly means (the so-called classical method) and not to individual variables (the so-called sampling method). The classical method of the surface flux computation is known to affect climatological means (e.g., Esbensen and Reynolds 1981; Hanawa and Toba 1987; Gulev 1994), but it influences variability patterns to a lesser extent, changing the magnitudes of variability but not the tendencies (e.g., Gulev 1997). In this sense future improvement of MSTG products (Worley et al. 2005) could be very desirable. It is important to quantify the impact of the “standard” and “enhanced” MSTG on the variability patterns (Wolter 1997). Note that improvement of MSTG requires application of variable corrections. Their application, however, decreases the number of reports available, because some corrections need information about the other variables. Thus, the “fully corrected” MSTG product will have sampling density comparable to that typical for the flux sampling.

An alternative way is to apply some corrections a posteriori to the monthly values (Ward and Hoskins 1996; Rayner et al. 2003). The same problem appears when the so-called MSTG pseudofluxes (products $\langle \delta T \cdot V \rangle$, $\langle \delta e \cdot V \rangle$) are used (e.g., Cayan 1992a,b) to partly account for the averaging effects.

Given that the impact of poor sampling is confined to rather small regions, it is natural to ask whether these differences actually matter. At least from the point of view of the atmosphere-driving anomalies of the North Atlantic circulation, we think that these differences are of paramount importance. Both observational (Curry et al. 1998; LabSea Group 1998) and modeling studies (e.g., Eden and Jung 2001; Eden and Willebrand 2001; Gulev et al. 2003) suggest that changes in the wintertime convective activity in the Labrador Sea region lead to subsequent changes of the North Atlantic circulation; in this context turbulent surface heat flux anomalies in the Labrador Sea play a crucial role. For example, Eden and Jung (2001) tested the NAO impact on the *observed* interdecadal variations of the North Atlantic circulation by forcing an ocean GCM with an NAO-related surface heat flux forcing over the period 1865–1997. The NAO-related forcing function was obtained by regressing NCEP–NCAR surface heat flux anomalies for the period 1957–97 onto the observed NAO index and then further combining the regression pattern with the observed NAO index for the period 1865–1997. Their spatial pattern largely resembles EOF1 of FULL (Fig. 6), and their model run has captured key aspects of North Atlantic interdecadal variability. Similarly, Eden and Willebrand (2001) and Gulev et al. (2003) obtained a realistic response of the North Atlantic meridional overturning in their ocean GCM experiments, which were forced using surface heat fluxes taken from the NCEP–NCAR reanalysis during the last several decades.

Given the importance of NAO-related surface heat flux forcing, particularly in the Labrador Sea, it is likely that a much weaker response (if any) would have been found had the mentioned studies used VOS data (cf. Fig. 6b) for their runs. This was exactly what happened in the study of Häkkinen (1999). She used for her simulation 3-yr ECMWF operational flux climatology (Barnier et al. 1995) to which interannual anomalies for 1950–89 were added from the ICOADS-based University of Wisconsin—Milwaukee (UWM) flux climatology (da Silva et al. 1994). Thus, the forcing function was suffering from sampling effects in the Labrador Sea as any VOS-derived climatology. Being quite reasonable in many respects, especially in the subtropical and mid-latitude regions, the results by Häkkinen (1999) show strong differences with the above-mentioned ex-

periments in the subpolar latitudes. Apart from the different model being used by Häkkinen (1999), sampling errors in the Labrador Sea region could explain the different outcomes of the modeling studies.

Given that large differences occur in relatively poorly observed areas one might ask how reliable “regularly sampled” reanalysis products actually are. We think that in poorly observed regions, at least in the Northern Hemisphere, their quality is relatively high compared to interpolated fields obtained from VOS data only. Comparisons with alternative reanalysis (ERA-40) and ECF support indirectly our conjecture demonstrating very similar results to those obtained using NCEP–NCAR. Moreover, medium-range forecasts based on state-of-the-art forecasting systems show a remarkable skill in predicting the large-scale atmospheric flow well into the far-medium range (e.g., Simmons and Hollingsworth 2002; Kalnay 2003), suggesting that the analyses are of high quality. This is especially true for subpolar oceanic regions where analysis error growth is known to be potentially at its largest (Buizza and Palmer 1995).

Of course, one might argue that the analysis quality has been improved dramatically with the availability of satellite data around the late 1970s. While this is certainly the case for the Southern Hemisphere, forecast experiments with recent versions of the ECMWF forecasting system for the Northern Hemisphere and the presatellite era show a remarkable forecast skill as well (Jung et al. 2004; Uppala et al. 2005), which is not too dissimilar from forecasts carried out for recent years (satellite era). This suggests that conventional observations alone are sufficient to yield very reliable analysis fields for the Northern Hemisphere. Finally, it is worth pointing out that atmospheric data assimilation systems actually make use of a variety of observational data (not just VOS data). First, the atmospheric flow—and therefore indirectly surface heat fluxes—is constrained by both sea and land observations, which are critical for the reliable representation of state of the atmosphere (Bengtsson et al. 2004a). The additional use of land observations, for example, from the west coast of Greenland, might be crucial for areas such as the Labrador Sea. Second, atmospheric data assimilation systems make effective use of past information as well (so-called cycling). In fact, it has been estimated that the global observation influence per assimilation cycle amounts to 15% whereas the first guess (short-range forecast from a previous analysis) influence amounts to 85% (Cardinali et al. 2004). It is this effective use of all different kinds of available data that make reanalysis products so valuable for use in climate-related studies. We do believe that in the future long-term time series of blended flux products (Yu et al. 2004) and satellite-

based flux products (e.g., Chou et al. 2003; Bentamy et al. 2003) will become available at the reanalyses temporal resolution as the last years of the blended wind products (Zhang et al. 2006). This will allow the use of them for estimation of sampling errors impact on variability patterns.

Acknowledgments. We thank Bernard Barnier of LEGI (Grenoble), Simon Josey and Liz Kent of SOC (Southampton), Bill Large of NCAR (Boulder), Ralf Lindau of MIUB (Bonn), Andreas Sterl of KNMI (De Bilt), and Glenn White of NCEP (Camp Spring) for fruitful discussion on different aspects of this work. Reliable feedback on data provision during many years from Scott Woodruff of CDC/NOAA (Boulder) and Steve Worley of DSS/NCAR (Boulder) is greatly appreciated. We are grateful to the four anonymous reviewers and the editor for valuable suggestions that largely helped to improve the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft Sonderforschungsbereich SFB-460, Ministry of Science and Education of Russian Federation under the World Ocean National Programme, and by Russian Foundation for Basic Research (Grant 05-05-64882).

REFERENCES

- Barnier, B., L. Siefridt, and P. Marchesiello, 1995: Thermal forcing for a global ocean circulation model using a three-year climatology of ECMWF analyses. *J. Mar. Syst.*, **6**, 363–380.
- Beismann, J. O., and B. Barnier, 2004: Variability of the meridional overturning circulation of the North Atlantic: Sensitivity to the overflow of dense water masses. *Ocean Dyn.*, **54**, 92–106.
- Bengtsson, L., K. I. Hodges, and S. Hagemann, 2004a: Sensitivity of the ERA-40 reanalysis to the observing system: Determination of the global atmospheric circulation from reduced observations. *Tellus*, **56A**, 456–471.
- , S. Hagemann, and K. I. Hodges, 2004b: Can climate trends be calculated from reanalysis data? *J. Geophys. Res.*, **109**, D11111, doi:10.1029/2004JD004536.
- Bentamy, A., K. B. Katsaros, A. M. Mestas-Núñez, W. M. Drennan, E. B. Forde, and H. Roquet, 2003: Satellite estimates of wind speed and latent heat flux over the global oceans. *J. Climate*, **16**, 637–656.
- Bjerknes, J., 1964: Atlantic air sea interaction. *Advances in Geophysics*, Vol. 10, Academic Press, 1–82.
- Buizza, R., and T. N. Palmer, 1995: The singular-vector structure of atmospheric global circulation. *J. Atmos. Sci.*, **52**, 1434–1456.
- Cardinali, C., S. Pezzulli, and E. Andersson, 2004: Influence matrix diagnostic of a data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **130**, 2767–2786.
- Cardone, J. S., J. G. Greenwood, and M. A. Cane, 1990: On trends in historical marine wind data. *J. Climate*, **3**, 113–127.
- Cayan, D., 1992a: Variability of latent and sensible heat fluxes estimated using bulk formulae. *Atmos.–Ocean*, **30**, 1–42.
- , 1992b: Latent and sensible heat flux anomalies over the Northern Oceans: The connection to monthly atmospheric circulation. *J. Climate*, **5**, 354–369.
- Chang, E. K. M., 2005: Effects of secular changes in frequency of observations and observational errors on monthly mean MSLP summary statistics derived from ICOADS. *J. Climate*, **18**, 3623–3633.
- Chou, S.-H., E. Nelkin, J. Ardizzone, R. M. Atlas, and C.-L. Shie, 2003: Surface turbulent heat and momentum fluxes over Global Oceans based on the Goddard satellite retrievals, version 2 (GSSTF2). *J. Climate*, **16**, 3256–3273.
- Curry, R. G., M. S. McCartney, and T. M. Joyce, 1998: Oceanic transport of subpolar climate signals to mid-depth subtropical waters. *Nature*, **391**, 575–577.
- da Silva, A. M., C. C. Young, and S. Levitus, 1994: *Anomalies of Directly Observed Quantities*. Vol. 2, *Atlas of Surface Marine Data*. NOAA Atlas NESDIS 2, 419 pp.
- Deser, C., and M. L. Blackmon, 1993: Surface climate variations over the North Atlantic Ocean during winter: 1900–1989. *J. Climate*, **6**, 1743–1753.
- Eden, C., and T. Jung, 2001: North Atlantic interdecadal variability: Oceanic response to the North Atlantic Oscillation (1865–1997). *J. Climate*, **14**, 676–691.
- , and J. Willebrand, 2001: Mechanisms of interannual to decadal variability of the North Atlantic circulation. *J. Climate*, **14**, 676–691.
- Esbensen, S. K., and R. W. Reynolds, 1981: Estimating monthly averaged air-sea transfers of heat and momentum using the bulk aerodynamic method. *J. Phys. Oceanogr.*, **11**, 457–465.
- Gulev, S. K., 1994: Influence of space–time averaging on the ocean–atmosphere exchange estimates in the North Atlantic midlatitudes. *J. Phys. Oceanogr.*, **24**, 1236–1255.
- , 1995: Long-term variability of sea-air heat transfer in the North Atlantic Ocean. *Int. J. Climatol.*, **15**, 825–852.
- , 1997: Climatologically significant effects of space–time averaging in the North Atlantic sea–air heat flux fields. *J. Climate*, **10**, 2743–2763.
- , B. Barnier, H. Knochel, J.-M. Molines, and M. Cottet, 2003: Water mass transformation in the North Atlantic and its impact on the meridional circulation: Insights from an ocean model forced by NCEP–NCAR reanalysis surface fluxes. *J. Climate*, **16**, 3085–3110.
- , T. Jung, and 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.
- Häkkinen, S., 1999: Variability in the simulated meridional heat transport in the North Atlantic for the period 1951–1993. *J. Geophys. Res.*, **104**, 10 991–11 007.
- Hanawa, K., and Y. Toba, 1987: Critical examination of estimation methods of long-term mean air-sea heat and momentum transfers. *Ocean–Air Interactions*, **1**, 79–93.
- Hurrell, J. W., 1995: Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation. *Science*, **269**, 676–679.
- Iwasaka, N., and J. M. Wallace, 1995: Large-scale air-sea interaction in the Northern Hemisphere from a view point of variations of surface heat flux by SVD analysis. *J. Meteor. Soc. Japan*, **73**, 781–794.
- Josey, S., and R. Marsh, 2005: Surface freshwater flux variability and recent freshening of the North Atlantic in the eastern subpolar gyre. *J. Geophys. Res.*, **110**, C05008, doi:10.1029/2004JC2521.
- , E. C. Kent, and P. K. Taylor, 1999: New insights into the

- ocean heat budget closure problem from analysis of the SOC air–sea flux climatology. *J. Climate*, **12**, 2856–2880.
- Jung, T., E. Klinker, and S. Uppala, 2004: Reanalysis and reforecast of three major European storms of the twentieth century using the ECMWF forecasting system. Part I: Analyses and deterministic forecasts. *Meteor. Appl.*, **11**, 343–361.
- Kalnay, E., 2003: *Atmospheric Modelling, Data Assimilation and Predictability*. Cambridge University Press, 341 pp.
- , and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471.
- Kaplan, A., M. A. Cane, Y. Kushnir, A. C. Clement, M. B. Blumenthal, and B. Rajapalan, 1998: Analyses of global sea surface temperatures 1856–1991. *J. Geophys. Res.*, **103**, 18 567–18 589.
- , Y. Kushnir, and M. A. Cane, 2000: Reduced space optimal interpolation of historical marine sea level pressure: 1854–1992. *J. Climate*, **13**, 2987–3002.
- , M. A. Cane, and Y. Kushnir, 2003: Reduced space approach to the optimal analysis interpolation of historical marine observations: Accomplishments, difficulties, and prospects. *Advances in the Applications of Marine Climatology*, World Meteorological Organization, WMO Tech. Doc. 1081, 199–216.
- Kent, E. C., P. K. Taylor, B. S. Truscott, and J. S. Hopkins, 1993: The accuracy of voluntary observing ships' meteorological observations—Results of the VSOP-NA. *J. Atmos. Oceanic Technol.*, **10**, 591–608.
- , and D. I. Berry, 2007: Metadata from WMO publication no. 47 and an assessment of voluntary observing ship observation heights in ICOADS. *J. Atmos. Oceanic Technol.*, in press.
- Kistler, R., and Coauthors, 2001: The NCEP–NCAR 50-Year Reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteor. Soc.*, **82**, 247–267.
- Kushnir, Y., 1994: Interdecadal variations in North Atlantic sea surface temperature and associated atmospheric conditions. *J. Climate*, **7**, 142–157.
- LabSea Group, 1998: The Labrador Sea Deep Convection Experiment. *Bull. Amer. Meteor. Soc.*, **79**, 2033–2058.
- Lindau, R., 2003: Errors of Atlantic air–sea fluxes derived from ship observations. *J. Climate*, **16**, 783–788.
- , H.-J. Isemer, and L. Hasse, 1990: Towards time-dependent calibration of historical wind observations at sea. *Trop. Ocean-Atmos. Newsl.*, **54**, 7–12.
- Parker, D. E., C. K. Folland, and M. Jackson, 1995: Marine surface temperature: Observed variations and data requirements. *Climatic Change*, **31**, 559–600.
- Peterson, E. W., and L. Hasse, 1987: Did the Beaufort scale or the wind climate change? *J. Phys. Oceanogr.*, **17**, 1071–1074.
- Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, 2003: Globally complete analyses of sea surface temperature, sea ice and night marine air temperature, 1871–2000. *J. Geophys. Res.*, **108**, 4407, doi:10.1029/2002JD002670.
- Rodwell, M. J., D. P. Rowell, and C. K. Folland, 1999: Oceanic forcing of the wintertime North Atlantic Oscillation and European climate. *Nature*, **398**, 320–323.
- Simmons, A. J., and A. Hollingsworth, 2002: Some aspects of the improvement of skill of numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **128**, 647–677.
- Smith, T. M., and R. W. Reynolds, 2003: Extended reconstruction of global sea surface temperatures based on COADS data (1854–1997). *J. Climate*, **16**, 1495–1510.
- , and —, 2004: Improved extended reconstruction of SST (1854–1997). *J. Climate*, **17**, 2466–2477.
- Sterl, A., 2001: On the accuracy of gap-filling algorithms in global surface fields. *Geophys. Res. Lett.*, **28**, 2473–2476.
- , 2004: On the (in)homogeneity of reanalysis products. *J. Climate*, **17**, 3866–3873.
- Tanimoto, Y., H. Nakamura, T. Kagimoto, and S. Yamane, 2003: An active role of extratropical sea surface temperature anomalies in determining anomalous turbulent heat flux. *J. Geophys. Res.*, **108**, 3304, doi:10.1029/2002JC001750.
- Trenberth, K. E., and J. W. Hurrell, 1994: Decadal atmosphere–ocean variations in the Pacific. *Climate Dyn.*, **9**, 303–319.
- Uppala, S., and Coauthors, 2005: The ERA-40 reanalysis. *Quart. J. Roy. Meteor. Soc.*, **131**, 2961–3012.
- von Storch, H., and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press, 503 pp.
- Ward, N. M., and B. J. Hoskins, 1996: Near surface wind over the global ocean 1949–1988. *J. Climate*, **9**, 1877–1895.
- White, G., 2000: Long-term trends in the NCEP/NCAR reanalysis. *Proc. Second Int. Conf. on Reanalyses*, Reading, United Kingdom, ECMWF, WCRP-109, WMO Tech. Doc. 985, 54–57.
- Wolter, K., 1997: Trimming problems and remedies in COADS. *J. Climate*, **10**, 1980–1997.
- Worley, S. J., S. D. Woodruff, R. W. Reynolds, S. J. Lubker, and N. Lott, 2005: ICOADS release 2.1 data and products. *Int. J. Climatol.*, **25**, 823–842.
- Yu, L., R. A. Weller, and B. Sun, 2004: Mean and variability of the WHOI daily latent and sensible heat fluxes at in-situ flux measurement sites in the Atlantic Ocean. *J. Climate*, **17**, 2096–2118.
- Zhang, H.-M., J. J. Bates, and R. W. Reynolds, 2006: Advances in higher resolution Global Ocean Observing System: Sea surface wind speed perspective. Preprints, *14th Conf. on Interaction of the Sea and Atmosphere*, Atlanta, GA, Amer. Meteor. Soc., CD-ROM, 6.5.