This is a PDF of a manuscript that has been peer-reviewed and accepted for publication. As the article has not yet been formatted, copy edited or proofread, the final published version may be different from the early online release.

This pre-publication manuscript may be downloaded, distributed and used under the provisions of the Creative Commons Attribution 4.0 International (CC BY 4.0) license. It may be cited using the DOI below.

The DOI for this manuscript is
DOI:10.2151/jmsj.2019-050

J-STAGE Advance published date: May 24th, 2019

The final manuscript after publication will replace the preliminary version at the above DOI once it is available.
Turbulent heat flux reconstruction in the North Pacific from 1921 to 2014

Bin SONG, Xiefei ZHI, Mengting PAN, Meiyi HOU, Chengfei HE

Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD), Nanjing University of Information Science & Technology/Key Laboratory of Meteorological Disasters, Ministry of Education (KLME), Nanjing University of Information Science & Technology, Nanjing 210044, China;

Nanjing Joint Center for Atmospheric Research, Nanjing 210008, China;

and

Klaus FRAEDRICH

Max Planck Institute for Meteorology, Hamburg 20146, Germany;

Corresponding author: Xiefei Zhi, College of Atmospheric Sciences, Nanjing University of Information Sciences & Technology, Nanjing 210044, China.

E-mail: zhi@nuist.edu.cn
ABSTRACT

The turbulent heat flux is the main passageway for air–sea interactions. However, due to lack of long-term observations for the turbulent heat flux, it is difficult to investigate the mechanisms of coupled ocean-atmosphere variabilities, such as the Pacific Decadal Oscillation. We here reconstructed the long-term turbulent heat flux in the North Pacific for the period 1921–2014 based on observations in the International Comprehensive Ocean-Atmosphere Data Set–International Maritime Meteorological Archive. The sea surface temperature, air temperature, wind and humidity were used to reconstruct the turbulent heat flux using the Coupled Ocean–Atmosphere Response Experiment (COARE3.5) algorithm. The modified Fisher–Tippett distribution was employed to calculate the turbulent heat flux at each grid square, and then the missing values were further derived based on data interpolating empirical orthogonal functions (DINEOF). The reconstructed turbulent heat flux was shown to be in accordance with the commonly used short-term heat flux datasets. This reconstruction is further examined by comparison with the long-term data of twentieth century reanalysis from European Center for Medium-range Weather Forecasts twentieth century reanalysis (ERA-20C) and the 20th Century Reanalysis (20CR) dataset from National Oceanic and Atmospheric Administration. It displays a good agreement with the ERA-20C both in spatial and temporal scales, but some differences from the 20CR. By these examinations, the reconstructed turbulent heat flux can well reproduce the main features of the air-sea interaction in the North Pacific, which can be used in studies of the air-sea interaction in the North Pacific on multidecadal timescales.
Keywords: reconstruction, turbulent heat flux, North Pacific, multidecadal timescales
1. Introduction

Research on decadal changes in climate has attracted much interest during the past two decades (e.g., Kushnir 1994; Zhang et al. 1997; Mantua et al. 1997; Minobe 1997; Latif et al. 2007; Newman et al. 2016). There are two main modes in the multidecadal variability of the Northern Hemisphere—the Atlantic Multidecadal Oscillation (AMO; Delworth et al. 2007) and the Pacific Decadal Oscillation (PDO; Mantua et al. 1997)—and there have been many investigations of their mechanisms (e.g., Delworth et al. 1993; Delworth and Greatbatch 2000; Newman et al. 2003; Miller and Schneider 2000). The mechanism of the PDO is less understood than that of the AMO (Liu 2012; Newman et al. 2016). Air–sea interactions are a fundamental problem in discussing the mechanisms of the decadal-scale variability of the climate. As the main air–sea passageway, the turbulent heat flux (latent and sensible heat fluxes) is useful in understanding the possible physical processes of air–sea interactions (e.g., Cayan 1992; Tanimoto et al. 2003; Taguchi et al. 2012; Liang and Yu 2016). However, long-term records of the turbulent heat flux in observations in the North Pacific are not available and therefore it cannot be used to investigate air–sea interactions in multidecadal time series such as the PDO. Long-term datasets of the turbulent heat flux in observations are therefore required for analyses of the multidecadal variability of the Earth’s climate.

Several different datasets of the turbulent heat flux are available, including the Objectively Analyzed Air–Sea Flux (OAFlux) dataset (Yu et al. 2008) from Woods Hole Oceanographic Institution, the Reanalysis Flux dataset from the National Centers for Environmental Prediction (NCEP, Kalnay et al. 1996), the National Oceanography
Centre Southampton (NOCS2.0) Surface Flux dataset (Berry and Kent, 2011), the Japanese 55-year Reanalysis (JRA-55) (Kobayashi et al., 2015; Harada et al., 2016), CORE2 (the Geophysical Fluid Dynamics Laboratory version 2 forcing for common ocean–ice reference experiments) (Large and Yeager, 2009), ECMWF (European Center for Medium-range Weather Forecasts) twentieth century reanalysis (ERA-20C) (Poli et al., 2013), and the 20th Century Reanalysis (20CR) dataset from NOAA (National Oceanic and Atmospheric Administration) (Compo et al. 2011). The OAFlux, NCEP, JRA-55 and CORE2 datasets start from the 1950s and the NOCS2.0 dataset is available from 1973. These datasets are too short for the use in multidecadal climate investigations. The ERA-20C and 20CR heat flux dataset are sufficiently long and have been used in some studies (Gao et al. 2016). Nevertheless, these two reanalysis may contain biases due to the model uncertainties, especially some difference exists between the two. It is necessary to reconstruct long time series of turbulent heat flux based on observations. At present, the marine observations obtained by ship have long time series of variables related to the turbulent heat flux. It is possible to construct long time series for the turbulent heat flux by observations from ships (Bunker 1976; Isemer and Hasse 1985; Josey et al. 1999). It is notable that the observations from ships contain biases, and many works have tried to reduce the biases (Berry et al. 2004; Berry and Kent, 2011).

The available long-term marine observations are collated in the International Comprehensive Ocean-Atmosphere Data Set–International Maritime Meteorological Archive (ICOADS–IMMA) 3.0 (Freeman et al. 2017). The ICOADS 3.0 dataset provides some air–sea variables, such as the sea surface temperature (SST), the air temperature,
wind and humidity, which means that it is possible to reconstruct the turbulent heat flux. A traditional Coupled Ocean–Atmosphere Response Experiment (COARE3.5) algorithm (Edson et al. 2013) is often used to construct the turbulent heat flux. Based on this algorithm, there are two methods for calculating the turbulent heat flux. The first method is the classical approach, in which flux-related variables are gridded and the heat flux is calculated from these gridded values (Oberhuber 1988; Berry and Kent, 2009). The second method is the sampling approach, in which the heat flux is calculated at each observation site (Gulev 1997). Gulev et al. (2013) employed the later to construct the turbulent heat flux in the North Atlantic. In this study, we used the sampling approach to reconstruct the turbulent heat flux in the North Pacific in the domain (120° E–130° W, 25° N–50° N) from 1921 to 2014. The data from 2015 are also provided in ICOADS, and the version since 2015 is version 3.0.1, some difference from version 3.0.0, so the data after 2014 were not considered in this research.

This paper is structured as follows: the source data and bias correction are introduced in Section 2. The number of observations and reconstruction of the relative humidity are described in Section 3. The reconstruction methods are presented in Section 4. The quality of reconstructed data is tested through comparison with multiple data in Section 5. The reconstructed heat flux was further examined by the air-sea interactions in the North Pacific in Section 6, followed by the discussion and conclusions in Section 7.

2. Source data and bias correction

The ICOADS 3.0 dataset covers the most widely observed surface marine
meteorological data collected from 1662 to 2014. To reduce possible inhomogeneities from different types of observations, only the reports from ships were selected to construct the turbulent heat flux. Several variables are involved in each marine observation. Among these, the SST, air temperature, wind and atmospheric humidity are required in the turbulent heat flux reconstruction. However, these four variables are not always included simultaneously at each observation report in the ICOADS 3.0 dataset. For example, some observation reports include the SST and air temperature, but not the wind or humidity. Compared with other variables, the number of observations of atmospheric humidity is limited, especially in the earlier decades, where the number of observations was too small for sampling. To supplement these data, the relative humidity was first reconstructed (specific method introduced in next section), and then the observations (including the reconstructed relative humidity) containing all four variables simultaneously were selected to reconstruct the turbulent heat flux.

It has been reported that there are systematic biases in these variables obtained from marine observations (Berry and Kent 2011). Researchers have tried to reduce the errors (Josey et al. 1999) and we used the following processes to minimize the bias.

(1) For the air temperature, there are biases due to the solar heating of the instruments and their surroundings. Berry et al. (2004) tried to reduce the heating biases by the heat budget, and their method was employed in our research. Several variables, such as wind speed, cloud cover, and dew-point are required in this method. However, there are not so many observations for some variables (e.g., cloud cover and dew-point) before the 1970s, making it difficult for the solar heating bias estimates. In this research,
we employed the ERA-20CM cloud cover and dew-point reanalysis. The results based on ERA-20CM and observations since 1970s were compared, and it shows that the difference is negligible, with the RMSE less than 0.07°C (Figure not shown). So ERA-20CM cloud cover and dew-point reanalysis were adopted to reduce the bias before the 1970s. It was also pointed out that there are bias in the night marine air temperature (Kent et al., 2013), especially during the Second World War. The day-night difference (Kent et al., 2013) was employed to correct the night marine air temperature from 1942 to 1945. Moreover, as the heights of the air temperature observations are different, temperature values from different heights were adjusted to a common reference height (10m) (Kent et al., 2013).

(2) It was believed that there was a spurious upward trend in wind speed (Thomas et al., 2008), and scientists have made efforts to reduce that. We followed the method provided by Tokinaga and Xie (2011) to remove the biases. The measured wind speed was adjusted to a common reference height (10m). As the anemometer heights cannot be available before 1970, the measured wind was employed after 1970. There are some missing anemometer height records after 1970, a monthly mean of 5° area gridded dataset of anemometer heights was considered to be the defaults (Berry and Kent, 2009). For the estimated wind speed, Lindau (1995) estimates were employed (Kent and Taylor 1997) to reduce the biases. It was also reported that there was an underestimate of the night wind speed because of the poor sight visibility, a day-night difference was added on the night wind speed to make a correction (Thomas et al., 2008; Tokinaga and Xie, 2011). The estimated wind speed after 1980 were not used in our research because some reports of
the estimated wind were actually measured by anemometers after 1980 (Tokinaga and Xie, 2011). That is, the estimated wind speed employed in this paper was from 1921 to 1980.

(3) There are also some biases in SST. Different observation method may result different bias. There are three methods for ship SST measurement. Some use a bucket to haul a sample of water to measure the SST, including the uninsulated (canvas) bucket and insulated (wooden or rubber) bucket. The water sample in the bucket may be cooled in most cases while hauling. The heat lose in canvas bucket is larger than that of wooden or rubber bucket. The second ship measurement is known as Engine Room Intake (ERI) measurement. This measurement may be influenced by the depth of the observation, circumstances peculiar to each vessel and heating of the water sample by the superstructure of the ship. The third is measured by the dedicated hull contact sensors. Hull sensors can provide more consistent SST measurements than ERI but the bias still needs to be evaluated (Kent et al. 2010). We followed the way of HadSST3 (Kennedy et al. 2011) to adjust these biases.

It should be noted that the SST required in the COARE3.5 is the skin temperature. But the SST we obtained is the bulk temperature. In order to transfer the bulk temperature to the skin temperature, shortwave and longwave radiation are required. The COARE3.5 provides the code for transferring the bulk SST into skin temperature. The shortwave and longwave radiation were set at default values in ICOADS3.5. But these values could vary with region and time, which could influence the results. We further evaluated whether this influence was sensitive. Because of limited observations for the solar radiation and
thermal radiation in earlier decades, the ERA-20CM shortwave and longwave radiation
were employed to calculate the heat flux. The heat flux based on default and varying
values are almost the same (the RMSE of the latent and sensible heat flux is only 0.84
and 0.41 W m\(^{-2}\), respectively). It is shown that the latent and sensible heat fluxes are not
sensitive to the choice of the radiation values.

(4) The method of Berry and Kent (2011) was employed to adjust the specific
humidity. The relative humidity was firstly calculated by air temperature, dew-point and
sea level pressure (SLP). Due to the rare observations for the relative humidity in the
early decades, the regression was conducted to reconstruct the relative humidity. Then,
the specific humidity was calculated based on relative humidity, air temperature and sea
level pressure. To reduce the bias from different measurements (i.e., marine screen and
sling psychrometers), the specific humidity was multiplied by 0.966 for the measurement
of the marine screen, based on the previous work (Berry and Kent 2011). For the
unknown measurement methods, 30-70% of the unknown methods measured specific
humidity was randomly selected to be multiplied by 0.966. To further reduce the possible
random errors, it was repeated for 30 times and the ensemble mean was considered to be
the adjusted specific humidity. After the adjustment, the residual bias uncertainty is 0.25
g kg\(^{-1}\) based on the differences between observed humidity from the two different
observation ways (marine screens and sling psychrometers), similar to the result of Berry

3. Number of observations and reconstruction of relative humidity
The mean number of reports (the reports containing the four variables simultaneously) for each 5°×5° grid square in each season is shown in Figure 1a. There are more than 100 reports in each grid square after 1940 (Fig. 1a), less than 45 reports per grid square in many places and even less than 20 reports in some areas (Fig. 1a) before 1940. Because there were few observations of humidity in earlier decades, the relative humidity was reconstructed using multivariable linear regression of the wind speed, air temperature and SLP (Gulev et al. 2013) (the robustness of the relative humidity obtained through the linear regression being introduced in the next second paragraph). To obtain a better reconstruction, we tested the regression results for different time periods and the period 1991–2010 was selected to identify the regression formula for the reconstruction of the relative humidity.

After the regression, the number of observations for the reconstructed heat flux in each grid square was much larger than before, especially during the period 1921–1940. The number of observations after regression was more than 20 in most locations before 1940, much larger than the original number (Fig. 1a). This is illustrated by the time series shown in Figure 1b. For the original data (blue line), the number of observations was very small before 1950 and increased rapidly after 1950, reaching a peak in the mid-1960s when the number of observations from ships was at a maximum. After that, the number decreased and many other kinds of observations became available—for example, from buoys. By regressing the relative humidity, the number of observations increased in most decades, especially in the 1920s and 1930s. The reason for this is that there were many reports of the SST, air temperature and wind speed, but only a limited
number of reports containing humidity in those decades. After regression, the number in
1940s is the smallest compared to other decades. Regressing the relative humidity
increased the number of marine observations and made it possible to calculate the
turbulent heat flux.

Wind speed, air temperature and SLP were employed to reconstruct the relative
humidity. The regression model was performed for each grid and all seasons. The result
of reconstructed relative humidity was examined. Figure 2a and 2b show the climatology
of the reconstructed humidity and the difference in the climatology between the original
and the reconstructed data. The distribution of the humidity showed a south–north pattern,
with a larger magnitude in the north and a smaller magnitude in the south (Fig. 2a), the
same as for the original data (figure not shown). In most locations the differences in
climatology between the reconstructed and the original relative humidity was less than
±2.0% (Fig. 2b), very small relative to the climatology. The latent heat fluxes based on
the observed and reconstructed relative humidity are shown in Figure 2c. The RMSE is
8.76 W m$^{-2}$ with a slope of 0.94, showing that the reconstruction of relative humidity
introduced small errors. The sensible heat fluxes based on the original and reconstructed
relative humidity are shown in Fig 2d, with the RMSE of 0.94 W m$^{-2}$. The changes of
relative humidity do not impact the result of sensible heat flux obviously, because the
relative humidity is the main factor in latent heat flux calculation, but not in sensible heat
flux. Figure 2e-f shows the scatter plot of the heat flux based on the original and
climatological mean relative humidity. The RMSE is 11.98 W m$^{-2}$ for latent heat flux by
the two ways (Fig 2e), and 1.22 W m$^{-2}$ for sensible heat flux (Fig 2f), larger than the
values shown in Figure 2c-d, indicating that humidity reconstruction by the method of regressing is better than the method of using a climatological value.

4. Methods for reconstruction of heat flux

4.1 COARE3.5 algorithm

The COARE3.5 algorithm is widely used to calculate the turbulent heat flux. Liu et al. (1979) developed bulk aerodynamic formulas based on Monin–Obukhov similarity and the COARE algorithm is an example of this approach. There are several versions of the algorithm (Webster and Lukas 1992; Fairall et al. 1996; Brunke et al. 2002; Fairall et al. 2003) and Edson et al. (2013) updated the bulk flux algorithm to the COARE3.5 algorithm by improving the parameterization of the drag coefficients.

The formulas for the latent heat flux $Q_{LH}$ and the sensible heat flux $Q_{SH}$ are:

$$Q_{LH} = \rho L_e C_e U (q_s - q_a),$$
$$Q_{SH} = \rho C_p C_h U (T_s - T_a),$$

where $\rho$ is the air mass density with the unit of [kg m$^{-3}$], and $L_e$ is the latent heat of evaporation with the unit of [J kg$^{-1}$]. $C_e$ and $C_h$ are the turbulent exchange coefficients for the latent and sensible heat fluxes, respectively, and are functions of the wind speed, detection height and atmospheric stability. $C_p$ is the specific heat capacity of air at constant pressure with the unit of [J kg$^{-1}$C$^{-1}$], and $U$ is the scalar wind speed with the unit of [m s$^{-1}$]. $q_s$ and $q_a$ are the specific humidity over sea water and the atmospheric specific humidity near the sea surface, respectively, with the unit of [g kg$^{-1}$], and the unit being transferred to [kg kg$^{-1}$] while calculating. $T_s$ is the SST and $T_a$ is the air
temperature with the unit of [°C]. Detailed calculations of the coefficients ρ, L_e, C_e, C_h, and C_p have been reported by Edson et al. (2013).

4.2 Modified Fisher–Tippett distribution

The heat flux was reconstructed by the COARE3.5 algorithm at each report and processed to obtain values for each 5°×5° grid square. Figure 1a shows the mean number of observations for each grid square in each season. There were large regional inhomogeneities, with fewer observations available in the central North Pacific and more observations near the coastline. The number of observations also varied greatly over time and was much smaller in the early decades than in recent decades (Fig. 1b). To minimize these inhomogeneities in both space and time, we randomly selected the same number of observations in each 5°×5° grid square—for example, 45 (or 20) reports for each season.

We calculated the heat flux for every season (DJF as winter, MAM as spring, JJA as winter, SON as autumn) for each grid square instead of for each month as a result of the limited number of observations in the early decades. Traditionally, the mean of the 45 (20) reports would be used as the grid value, but the probability density distribution (PDF) of the turbulent heat flux does not follow a normal distribution (Gulev and Belyaev, 2012) and thus cannot be used. Gulev and Belyaev (2012) studied the probability distributions of surface turbulent heat fluxes. They found that a modified Fisher–Tippett (MFT) distribution gives a good description of the statistical properties of turbulent heat fluxes. Thus, MFT distribution was employed in this research.

a. Formulas of the MFT distribution
The MFT distribution was applied to the heat flux calculation for each 5°×5° grid square. In each grid square, 20 (or 45) samplings were selected to calculate the grid value.

The PDF of the turbulent heat flux is:

\[ P(x) = \alpha \beta \exp(-\beta x) \exp[-\alpha \exp(-\beta x)], \quad (2) \]

where \( \alpha \) and \( \beta \) are two parameters; \( \alpha \) is non-dimensional and \( \beta \) has the dimensions of the inverse units of a variable. The two parameters can be calculated as:

\[
\frac{n}{\alpha} = \sum_{i=1}^{n} \exp(-\beta x_i)
\]

\[
\frac{n}{-\beta} + \sum_{i=1}^{n} x_i = \alpha \sum_{i=1}^{n} x_i \exp(-\beta x_i),
\quad (3)
\]

where \( n \) is the sampling number of the heat flux data (20 or 45 in each grid square) and \( x \) is the turbulent heat flux value for each sample. As the analytical solution cannot be obtained using Eq. (3), we calculated the numerical solution instead. To avoid a floating overflow, the turbulent heat flux was scaled by \( 10^3 \) and the value of \( \beta \) was multiplied by \( 10^3 \). Using the values of the parameters \( \alpha \) and \( \beta \), the value of the latent and sensible heat flux for each grid can be computed as:

\[ \bar{x} = \ln \alpha / \beta. \quad (4) \]

For each 5°×5° grid square in each season, 20 (or 45) samplings were selected. The parameters \( \alpha \) and \( \beta \) were calculated by Eq. (3). The turbulent heat flux for each grid then was obtained using Eq. (4). By the process above, the turbulent heat flux for each 5°×5° grid square was acquired.

**b. Results of the MFT distribution**

Here we selected 20 samples for each 5°×5° grid square. Figure 3 plots the result of the parameters \( \alpha \) and \( \beta \). The \( \alpha \) parameter for the latent heat flux was between 2 and 3.5
in most locations in the North Pacific, larger at lower latitudes and smaller at high latitudes. The α parameter for the sensible heat flux was between 0 and 2 and was slightly larger in the western North Pacific. The magnitude and distribution of α for both latent and sensible heat fluxes are in good agreement with previously reported work (Gulev and Belyaev 2012). The magnitudes of β for both latent and sensible heat flux were between 0 and 250. The distribution of β for the latent heat flux was smaller in the southwest North Pacific and larger in the north, whereas β for the sensible heat flux was smaller in the northwest North Pacific and larger in the south. The magnitudes and distribution of β are in good agreement with previously reported work (Gulev and Belyaev 2012), showing our MFT distribution being reasonable.

We also tried to randomly select 45 samples for each 5°×5° grid square. Figure 4a and 4b show the difference in climatology between the methods with 20 and 45 samples for both the latent and sensible heat fluxes from 1950 to 2010. The difference between the two sampling methods is less than 2 W m⁻² in most areas, very small relative to the climatology. The largest difference locates in the low latitude and the central of the North Pacific, where the numbers of observation are relatively small (Fig. 1a). A point-to-point comparison is shown as a scatter plot (Fig. 4c and 4d). The slope in Figure 4c is 0.99, with an RMSE of 5.5 W m⁻² (latent heat flux) and the sensible heat flux has a slope of 1.01 and an RMSE of 2.8 W m⁻², only a slight difference between the two samplings. It seems that the differences are negligible for both reconstructions based on 20 and 45 samplings. Figure 4c,d show the uncertainties in sampling (5.5 W m⁻² for the latent heat flux and 2.8 W m⁻² for the sensible heat flux).
As the number of reports was less than 45 in many areas in the early decades, especially before 1950, we selected 20 reports in the entire period. There may be random errors due to the random selection of the 20 reports. To minimize these errors, we repeated the random selection 20 times and the ensemble mean of the 20 experiments was taken as the heat flux value over the 5°×5° grid square.

4.3 Data Interpolating Empirical Orthogonal Functions (DINEOF)

Using the processes described in the preceding sections, values were obtained for each 5°×5° grid square. But there are some missing data in some areas and periods. Figure 5 shows the missing number of the grid square in temporal and spatial scale. It shows that the missing number is small after the 1950s, and largest in the 1940s (Fig. 5a). The total number of grids without missing is 146, with about 100 grids missing in the 1940s. The reason for the large missing number in that decade may be many ships stopping sailing during the Second World War (Gulev et al. 2013), especially in the North Pacific, leading to an absence of observations. Figure 5b shows the distribution of the amount of missing data from 1921 to 2014. The dark blue color in the east and west North Pacific indicates small amount of missing dataset, less than 10. The missing number is the largest near the west seashore of the North Pacific, and a little smaller in the low latitude of the North Pacific. Because of these missing values, it is necessary to interpolate the missing heat flux. The DINEOF method (Beckers and Rixen, 2003; Alvera-Azcarate et al., 2005) was used to obtain a full dataset.

DINEOF is a widely accepted method for reconstructing missing dataset (Huang et al., 2017; Hernández-Carrasco et al., 2018). It is a self-consistent, parameter-free
technique with the advantage of not needing of a priori information. Based on EOF, the missing data can be reconstructed by the EOF modes and their amplitude time series. Compared to some other interpolation method such as Optimal Interpolation (OI) reconstruction, the results conducted by both methods are similar, but the computational time of DINEOF can be reduced 30 times (Alvera-Azcarate et al., 2005). In this method, the climatology is subtracted to derive the anomalies and parts of the non-missing data are randomly removed for cross validation. Then the missing and removed data are set to zero. EOF analyses are conducted to decompose the heat flux and the optimal number of EOF modes is identified by the cross validation with randomly removed data. Finally, the missing data were replaced by the EOF reconstruction.

Figure 6 shows the root mean squares errors (RMSE) for different number of EOF modes. When the RMSE is the smallest, the corresponding EOF mode was selected to construct the heat flux. The RMSE are large for the first several modes, and reduce to a relative small value as the number of modes grows. After that, the RMSE grows again. For latent heat flux, the minimum of RMSE occurs at the 13th EOF mode, with the value of 15.51 W m\(^{-2}\). The minimum of RMSE for sensible heat flux also occurs at the 12th EOF mode, with the value of 7.48 W m\(^{-2}\). That is, it may bring an uncertainty of 15.51 and 7.48 W m\(^{-2}\) for latent and sensible heat flux respectively by DINEOF. By the processes above, the latent and sensible heat flux from 1921 to 2014 of the whole domain in the North Pacific can be acquired.

5. Comparison of the reconstructed dataset with other available heat flux datasets
5.1 Compared with the heat flux datasets for recent decades

To examine the result for the reconstructed heat flux, the reconstructions were compared with other commonly used heat flux datasets, including the OAFlux, NCEP reanalysis, NOCS2.0, JRA-55 and CORE2 turbulent heat fluxes. The time series of these datasets are different: OAFlux is from 1958, the NCEP reanalysis is from 1948, NOCS2.0 is from 1973 to 2014, JRA-55 is from 1958 and CORE2 is from 1949 to 2006. The common period 1974–2006 was selected to make the comparison, and here the year of 1973 is not considered since the value in December 1972 cannot be acquired in NOCS2.0. The resolutions for OAFlux, NOCS2.0, JRA-55 and CORE2 are $1 \times 1^\circ$, $1 \times 1^\circ$, $1.25 \times 1.25^\circ$ and $1 \times 1^\circ$ respectively, and NCEP T62 Gaussian grid with $192 \times 94$ points. In order to make a comparison with the reconstruction, all these datasets are degraded to the resolution of $5 \times 5^\circ$. It should be noted that the positive value indicates the heat releasing from the ocean to the atmosphere in the reconstruction, OAFlux, NCEP, NOCS2.0 and JRA-55, but the opposite is true in CORE2. In order to make a consistence, we defined the positive value when heat released from the ocean to the atmosphere.

The climatology of the five turbulent heat flux datasets is compared in Figure 7. For the latent heat flux, all six datasets presented a south–north pattern, smaller in the north and larger in the south, with the largest over the Kuroshio extension. For the sensible heat flux, an east–west pattern is shown, smaller in the east and larger in the west, with the largest locating over the Kuroshio extension as well. The reconstructed latent and sensible heat flux are therefore consistent with the available heat flux in terms of the climatology.
The correlation coefficients between the reconstructed data and the other datasets were also computed and the spatial distribution of the correlation was shown in Figure 8. For the latent heat flux, in most areas, the correlation coefficients between the reconstructed data and the OAFlux data were >0.4 (Fig. 8a). The largest correlation locates in the west of the North Pacific, with the value about 0.8. A similar correlation pattern can be seen for the relation between the reconstructed data and the NCEP, NOCS2.0, JRA-55 and CORE2 datasets (Fig. 8b, c, d and e). Among the figures in Fig. 8(a-e), the correlations between the reconstruction and NOCS2.0 (Fig. 8(c)) are the largest. The reason may be that both the reconstructed and NOCS2.0 heat flux are from marine observations, but others may be from others sources, models as an example. The correlations for sensible heat flux present similar results (Fig 8(f-j)). In order to understand the somewhat weak correlation between the reconstruction and other heat flux data, the flux-related variables (air temperature, SST, wind speed and specific humidity) of the reconstruction were correlated with other source data (Table 1). The correlations between the adjusted and other source air temperature are high—larger than 0.9, also high for SST—about 0.80. And it is a little smaller for wind speed, about 0.7. The correlations between the adjusted and other source data specific humidity varies for different data, 0.87 between the adjusted and NOCS, and 0.37 between the adjusted and NCEP. It seems the uncertainties in humidity from different data are relative large. We conclude the reason for the somewhat weak correlation between the reconstructed and other source heat flux in some areas may be the result of humidity and wind. On the whole, the reconstruction correlates well with other commonly used heat flux data.
The time series of these turbulent heat flux data during the period 1974-2006 were shown in Fig. 9. As the North Pacific is so large, we divided the region into four domains marked in Fig. 8e as domain A (120° E–170° W, 25° N–40° N), domain B (120° E–170° W, 40° N–50° N), domain C (170° W–130° W, 25° N–40° N) and domain D (170° W–130° W, 40° N–50° N). For both the latent and sensible heat fluxes, the reconstruction, OAFlux, NCEP, NOCS2.0, JRA-55 and CORE2 data vary consistently from year to year. The reconstructed data correlate well with the OAFlux, NCEP, JRA-55, NOCS2.0 and CORE2 datasets, especially for the latent heat flux (all >0.85). Similar results were found for the other three domains (Fig. 9). The time series of the reconstruction are relatively in good accordance with other commonly used turbulent heat flux data.

The trends of the reconstruction, OAFlux, NCEP, NOCS2.0, JRA-55 and CORE2 heat flux datasets were compared during the period 1974–2006 (Fig. 10). These datasets were anomalies from seasonal average. The trends of these latent heat flux datasets show similar pattern, with positive trends in most areas of the domain, especially in the southwest North Pacific. There are also some differences between these datasets. The largest positive trend locates more northwest for the OAFlux, NCEP and JRA-55, but east for the reconstruction and NOCS2.0. The trends of reconstruction are similar to the NOCS2.0. That may be the results of both the two are from marine observations, while others from reanalysis or combination. The sensible heat flux also shows similar result, with the similarity between the reconstructed and NOCS2.0. In a word, the trends of the reconstruction are similar to other datasets, especially NOCS2.0. From the analyses above, the reconstruction is in good agreement with other common used heat flux data for
recent decades.

5.2 Compared with the long-term heat flux datasets

The reconstruction was further compared with the long-term heat flux data (ERA-20C and 20CR) to test how the reconstruction performed on long-term period. As the ERA-20C is from 1900 to 2010 and 20CR since 1871, the common period 1921-2010 was selected to make the comparison. The resolution for ERA-20C is $1 \times 1^\circ$ and 20CR with T62 Gaussian grid with $192 \times 94$ points. Both the heat flux data were degraded to the resolution of $5 \times 5^\circ$ to make a comparison with the reconstruction. All these datasets were anomalies from the seasonal climatology.

Figure 11 shows the correlations between the reconstruction and the other two long-term heat flux data. For the latent heat flux, the correlations between the reconstruction and ERA-20C are $>0.4$ in most areas (Fig. 11a), larger in the south than north. The correlation between the reconstruction and 20CR is different, with the positive correlation locating in the west and east of the North Pacific, while negative in the central of the North Pacific (Fig. 11b). The positive areas are the areas with enough observations, while negative areas with insufficient observations (Figure 1a and 5b). That is, in the areas with sufficient observations, the reconstruction and 20CR latent heat flux show large consistency, while some difference existed in the areas with deficient observations. It is not the case in the correlation between the reconstruction and ERA-20C latent heat flux. In the latter, the correlations are high in the areas no matter with sufficient or insufficient observations.

The correlations between the reconstruction and ERA-20C sensible heat flux are
larger than 0.4 in most areas, a little smaller than that of the latent heat flux (Fig. 11c).

Similar pattern can be seen in the correlations between reconstruction and 20CR sensible
heat flux (Fig. 11d). The patterns in Figure 11b and Figure 11d are different. The
correlations in the west and east North Pacific are both positive, but are opposite in the
central of the North Pacific. According to Eq. (1), except for the parameters and wind
speed, latent heat flux is determined by humidity, and sensible heat flux determined by
air temperature and SST. It is concluded that the negative correlation in the central of the
North Pacific in Figure 11c may be the inconsistency of the reconstructed and 20CR
humidity.

6. Air-sea interaction in the North Pacific

The reconstructed heat flux was applied to the air-sea interaction in the North Pacific
to test whether it can reproduce the result of previous research. Former research shows
that the atmosphere drives the ocean in the North Pacific in boreal winter in most areas
(Cayan 1992; Iwasaka and Wallace 1995). When the Aleutian low is anomaly strong,
enhanced wind speed and reduced air temperature and humidity will cool the ocean
through turbulent heat flux (Newman et al., 2016), and Ekman transport would also
amplify this pattern (Alexander and Scott 2008). But in the subarctic frontal zone (SAFZ),
specifically in the Kuroshio Extension (KE) and Oyashio frontal zones, positive anomaly
SST may release heat flux to the atmosphere (Tanimoto et al. 2003; Taguchi et al. 2012).
We correlated the heat flux and SST in the North Pacific to see whether the
reconstruction can reproduce the phenomenon. The SST is HadISST with the resolution
of $1 \times 1^\circ$ from the Met Office Hadley Centre (Rayner et al., 2003). To be correlated with the turbulent heat flux, the resolution of SST was degraded to $5 \times 5^\circ$. Figure 12 shows the correlation between the two variables in the North Pacific in boreal winter. The correlations are negative in most areas when the atmosphere leads the SST for a month (DJF average for turbulent (latent plus sensible) heat flux and JFM average for SST), indicating that the strengthening of the surface westerlies may enhance the underlying SST (Fig. 12a), (similar to the Figure 1c in Tanimoto et al. (2003)). When SST leads the heat flux for a month (NDJ average for SST and DJF average for turbulent heat flux), the strong positive correlations can be seen in the SAFZ (Fig. 12b). Similar result can also be seen in the lead-lag correlation between the ERA-20 turbulent heat flux and SST (Fig. 12c,d). These patterns are similar to Figure 2c in Tanimoto et al. (2003), indicating the ocean driving the atmosphere in that area (Miller et al. 1998; Deser et al. 1999). The reproducing the air-sea interaction in the North Pacific by the reconstructed turbulent heat flux indicates that our reconstruction is reasonable.

7. Conclusions and discussion

The turbulent heat flux in the North Pacific from 1921 to 2014 was reconstructed based on the ICOADS 3.0 dataset using the COARE3.5 algorithm. The variables SST, air temperature, wind speed and humidity were required to calculate the heat flux data. Because insufficient humidity data were available in the earlier decades, we first constructed the relative humidity through regression based on the wind speed, air temperature and SLP. The regressed humidity was in good agreement with the actual
observed humidity. The humidity was then combined with other three variables (SST, air temperature and scalar wind) to reconstruct the turbulent heat flux according to the COARE3.5 algorithm. To minimize the inhomogeneities in the sampling, 20 reports were randomly selected over each $5^\circ \times 5^\circ$ grid square for each season to calculate the grid value. The MFT distribution was used to calculate each grid value instead of the average value. To reduce the sampling errors, we repeated the experiment 20 times and the ensemble mean was taken as the value for the grid square. However, even when using this method, there were still some missing data in some locations during some periods, especially in the central North Pacific and in the 1940s. DINEOF was used to interpolate the missing values. An uncertainty of $17.22 \text{ W m}^{-2}$ for the latent heat flux and $7.51 \text{ W m}^{-2}$ for the sensible heat flux was introduced by the DINEOF reconstruction. These processes above were used to obtain the latent and sensible heat fluxes in each grid square in the North Pacific from 1921 to 2014.

The reconstructed heat flux was shown to be in accordance with the commonly used heat flux datasets (OAFlux, NCEP, NOCS2.0, JRA-55 and CORE2) for recent decades. The climatology patterns of the reconstruction coincide with other turbulent heat flux data, with the largest over the Kuroshio extension for both latent and sensible heat flux. The correlations between the reconstruction and other heat flux data are relatively high in most areas, except some minor correlations in some areas, which may be the result of the uncertainty of the relative humidity and wind speed. The trends of the reconstruction show similar patterns to other commonly used turbulent heat flux. The reconstructed heat flux was further compared with the long-term data (ERA-20C and 20CR).
reconstructed heat flux (both latent and sensible) is highly correlated with the ERA-20C. Correlations between the reconstructed and 20CR latent heat flux show a pattern of positive values in the west and east of the North Pacific, and negative values in the central of the North Pacific. That may the result of inconsistency of the reconstructed and 20CR relative humidity. The reconstructed turbulent heat flux can reproduce the air-sea interaction in the North Pacific—the strengthening of the surface westerlies enhancing the underlying SST in most areas, and the ocean driving the atmosphere in the SAFZ. These examinations indicate the rationality of the reconstruction.

Long-term datasets of the reconstructed turbulent heat flux can be used to study the physical processes of the air–sea interaction in the North Pacific on multidecadal timescales. Future work will focus on the related physical processes and we will try to identify which predominates in the air–sea interaction in the North Pacific, whether the atmosphere influences the ocean (Hasselmann 1976; Frankignoul and Reynolds 1983) or the ocean forces the atmosphere (Deser et al. 1999; Qiu et al 2010). This will improve our understanding of multidecadal variability in the North Pacific.
Acknowledgments. This work was supported by the National (Key) Basic Research and Development (973) Program of China (2012CB955204) and National Natural Science Foundation of China (41575104). The authors express their gratitude to the Woods Hole Oceanographic Institution (WHOI), NOAA, The National Oceanography Centre Southampton (NOCS), Japan Meteorological Agency (JMA), ECMWF and Hadley Centre for providing the datasets. We also appreciate NOAA for providing the code of COARE3.5. Special thanks are given to Elizabeth Kent and Chris Fairall for their comments and discussions on this paper. Finally, we acknowledge the anonymous reviewers for their constructive criticism and helpful comments during the review process.
References


Cayan D. R., 1992: Latent and sensible heat flux anomalies over the northern oceans:


Huang J., X. Zhang, Q. Zhang, Y. Lin, 2017: Recently amplified arctic warming has
contributed to a continual global warming trend. Nature Climate Change, https://doi.org/10.1038/s41558-017-0009-5


Kent, E. C., N. A. Rayner, D. I. Berry, M. Saunby, et al., 2013: Global analysis of night


Qiu B., S. Chen, 2010: Eddy-mean flow interaction in the decadally modulating Kuroshio


Zhang Y., J. M. Wallace, and D. S. Battisti, 1997: ENSO-like interdecadal variability:
Fig. 1. (a) Mean number of observations for each 5°×5° grid square in every season for different time periods. The left-hand column shows the original numbers, whereas the right-hand column shows the numbers after reconstruction of the relative humidity. (b) Time series of the number of observations for each 5°×5° grid square in different years. The blue line shows the original number and the red line showing the number after reconstruction of the relative humidity. It should be noted that the number in (b) is in logarithmic scale.
Fig. 2. (a) Annual mean climatology of the reconstructed relative humidity for the period 1961–2014 (units: %). (b) Difference in climatology between the reconstructed relative humidity and the original data (the reconstructed data minus the original data) (units: %). (c) Scatter plot of the latent heat flux based on the original relative humidity and the reconstructed relative humidity with a slope of 0.94 and an RMSE of 8.76 (units: W m$^{-2}$). (d) same as (c), but for sensible heat flux, with a slope of 0.99 and an RMSE of 0.94 (units: W m$^{-2}$). (e) Scatter plot of the latent heat flux based on the original relative
humidity and the relative humidity using a climatological value, with a slope of 0.91 and an RMSE of 11.98 (units: W m$^{-2}$). (f) same as (e), but for sensible heat flux, with a slope of 0.98 and an RMSE of 1.22 (units: W m$^{-2}$).
Fig. 3. Annual mean climatology of the parameters (α and β) in the modified Fisher-Tippett distribution: (a) α of latent heat flux; (b) α of sensible heat flux; (c) β of latent heat flux; and (d) β of sensible heat flux.
Fig. 4. Comparison of different sampling methods (20 and 45 for each 5°×5° grid square).

(a) Difference in climatology between the 20 and 45 sampling method for the latent heat flux from 1950 to 2010 (20 sampling method minus 45 sampling method) (units: W m\(^{-2}\)).

(b) same as (a), but for sensible heat flux. (c) Scatter plot of the latent heat flux for the 20 and 45 sampling methods from 1950 to 2010 with a slope of 0.99, an intercept of 0.2 and an RMSE of 5.5 (units: W m\(^{-2}\)); (d) same as (c), but for sensible heat flux, with a slope of 1.01, an intercept of -1.2 and an RMSE of 2.8. (units: W m\(^{-2}\)).
Fig. 5. (a) Time series of the missing number of total grid squares for the whole domain (the total number of grids without missing is 146). (b) Distribution of the missing number during the period 1921-2014.
Fig. 6. Root mean square errors for different numbers of EOF modes while DINEOF process. (a) Latent heat flux, with minimum RMSE of 15.51 in the 13th EOF mode. (b) same as (a), but for sensible heat flux, with minimum RMSE of 7.48 in the 12th EOF mode. (units: W m$^{-2}$).
Fig. 7. Spatial distribution of the latent and sensible heat fluxes averaged over the 33-year period (1974–2006) of the reconstructed dataset and other five available turbulent heat flux datasets (OAFlux, NCEP, NOCS2.0, JRA-55 and CORE2) (positive values indicating the direction from ocean to atmosphere) (units: W m$^{-2}$). (a) Reconstructed latent heat flux; (b) OAFlux latent heat flux; (c) NCEP latent heat flux; (d) NOCS2.0 latent heat flux; (e) JRA-55 latent heat flux; (f) CORE2 latent heat flux; (g-l) same as (a-f), but for sensible heat flux (units: W m$^{-2}$).
Fig. 8. Correlation coefficients between the reconstruction and other datasets for both latent and sensible heat fluxes. (a) Correlation between the reconstructed and the OAFlux latent heat flux; (b) same as (a), but for the reconstructed and NCEP latent heat flux; (c) same as (a), but for the reconstructed and NOCS2.0 latent heat flux; (d) same as (a), but for the reconstructed and JRA-55 latent heat flux; (e) same as (a), but for the reconstructed and CORE2 latent heat flux; (f-j) same as (a-e), but for sensible heat flux. Dotted are the areas passing 95% significant level test.
Fig. 9. Year to year variability of six latent and sensible heat flux datasets averaged in the four different domains marked in Fig. 8e. (a) Latent heat flux averaged in domain A; (b) latent heat flux averaged in domain B; (c) latent heat flux averaged in domain C; (d) latent heat flux averaged in domain D; (e-h) same as (a-d), but for sensible heat flux.
Fig. 10. Trends in the reconstruction, OAFlux, NCEP, NOCS2.0, JRA-55 and CORE2 heat flux data during the period 1974-2006 (data are anomalies from seasonal mean). (a) Trends in the reconstructed latent heat flux; (b) same as (a), but for OAFlux latent heat flux; (c) same as (a), but for NCEP latent heat flux; (d) same as (a), but for NOCS2.0 latent heat flux; (e) same as (a), but for JRA-55 latent heat flux; (f) same as (a), but for CORE2 latent heat flux; (g-l) same as (a-f), but for sensible heat flux (units: W m$^{-2}$ yr$^{-1}$). Dotted are the areas passing 95% significant level test.
Fig. 11. Correlation coefficients between the reconstructed dataset and long-time series dataset (ERA-20C and 20CR) for both latent and sensible heat fluxes. (a) Correlation between the reconstructed and the ERA-20C latent heat flux; (b) same as (a), but for the reconstructed and 20CR latent heat flux; (c-d) same as (a-b), but for sensible heat flux. Dotted are the areas passing 95% significant level test.
Fig. 12. Correlation coefficients between the turbulent heat flux (latent plus sensible) and SST in boreal winter. (a) Reconstructed turbulent heat flux leading SST for one month (DJF average for turbulent (latent plus sensible) heat flux and JFM average for SST). (b) same as (a), but SST leading turbulent heat flux for one month (NDJ average for SST and DJF average for turbulent heat flux). (c-d) same as (a-b), but for ERA-20 turbulent heat flux. Dotted are the areas passing 95% significant level test.
Table 1. Correlation coefficients between adjusted and other source data for the flux-related variables (all passing 95% significant test)

<table>
<thead>
<tr>
<th></th>
<th>Air temperature</th>
<th>SST</th>
<th>Wind speed</th>
<th>Humidity</th>
</tr>
</thead>
<tbody>
<tr>
<td>corr(Adjusted, OAF)</td>
<td>0.94</td>
<td>0.86</td>
<td>0.74</td>
<td>0.80</td>
</tr>
<tr>
<td>corr(Adjusted, NOCS)</td>
<td>0.95</td>
<td>0.87</td>
<td>0.70</td>
<td>0.87</td>
</tr>
<tr>
<td>corr(Adjusted, NCEP)</td>
<td>0.91</td>
<td>0.78</td>
<td>0.58</td>
<td>0.37</td>
</tr>
<tr>
<td>corr(Adjusted, JRA)</td>
<td>0.94</td>
<td>0.83</td>
<td>0.45</td>
<td>0.66</td>
</tr>
<tr>
<td>corr(Adjusted, CORE)</td>
<td>0.90</td>
<td>0.82</td>
<td>0.72</td>
<td>0.48</td>
</tr>
</tbody>
</table>