Ocean surface air temperature derived from multiple data sets
and artificial neural networks

Catherine Gautier 1,2, Pete Peterson 1 and Charles Jones 1

Abstract. This paper presents a new method to derive monthly averaged surface air temperature, $T_a$, from multiple data sets. Sea Surface Temperature (SST) from the National Centers for Environmental Prediction (NCEP) and total precipitable water (W) from the SSM/I sensor are used as inputs to Artificial Neural Networks (ANN). Surface air temperature ($T_a$) measurements from the Surface Marine Data (SMD) are used to develop and evaluate the methodology. When globally evaluated with SMD data, the bias of the new method is small ($0.030^\circ$ C $\pm 0.26^\circ$ C), and the accuracy expressed as root-mean square (rms) differences has a small global mean ($0.73^\circ$ C $\pm 0.37^\circ$ C). These biases and rms differences are smaller than those obtained using NCEP reanalyses and TIROS Operational Vertical Sounder (TOVS) data products. When evaluated with the TOGA-TAO array measurements over the tropical Pacific, the ANN mean bias and rms differences have similarly small values, $0.37^\circ$ C and $0.61^\circ$ C, respectively.

1. Introduction

Surface air temperature has long been the parameter of choice for characterizing and monitoring climate and is indeed often thought to be a fingerprint of global changes. For instance, the impact of enhanced greenhouse effect induced primarily by increased CO$_2$ concentration is mostly expressed as a change in surface air temperature (IPCC, 1993). Surprisingly, however, with the advent of space observations, the use of surface air temperature as a remote sensing parameter has somewhat retreated, mostly because it is rather difficult, if not impossible, to measure it from space to the few tenths of degree accuracy needed for climate change studies. Measurements from near surface layers to the radiation measured at the satellite level is difficult to deconvolve from the contributions of higher atmospheric layers.

Since air temperature ($T_a$) is key to climate studies, the search for improved approaches to determine it from space is an active research topic. A number of recent studies have suggested improved accuracy on surface air temperature retrievals from space observations (Konda et al., 1996; Prihodko and Goward, 1997, Anyamba et al., 1998). This paper contributes to this research effort by presenting a new method to compute monthly averaged $T_a$ over the ocean from multiple data sets on a global basis and using an artificial neural network (ANN) approach.

2. Data

Monthly averages of $T_a$ are obtained using two parameters best describing the lower atmosphere properties that are routinely available from satellite observations and numerical weather prediction analyses. Monthly averages of total precipitable water (W) are obtained from the Special Sensor Microwave Imager (SSMI) data (Wentz, 1992). The sea surface temperature (SST) field derives from the National Centers for Environmental Prediction (NCEP) operational analysis, which is based on the method of Reynolds and Smith (1994). The data record spans from January 1988 through August 1997 (1° latitude x 1° longitude resolution). The development and testing of the new method are accomplished by using observations from the Surface Marine Data (hereafter SMD, da Silva et al., 1994). The SMD is obtained from ship and buoys (some of the IAO, see below) reports and extends from January 1988 through November 1993. Additional comparisons are performed with $T_a$ fields obtained from NCEP reanalyses at 2.5° x 2.5°, and TOVS data. In addition to ship observations, this study uses $T_a$ observations from the moored buoys array from the Tropical Ocean Global Atmosphere Tropical Ocean Atmosphere (TOGA-TAO) program (McPhaden, 1995). Daily averages of $T_a$ measurements from the TAO array for the period January 1988-August 1997 are used to compute monthly averages. Since the number of available daily observations varies widely among the buoys during this period, we specified a cutoff criterion of at least 20 observations to compute monthly averages. Several studies have investigated the minimum number of observations necessary to obtain a reliable estimate of monthly averages. The results vary with parameters and there has not been any specific study for air temperature in the tropical Pacific, but 20 observations are consistent with the numbers chosen by other authors (e.g., da Silva, 1994, pp33).

3. Methodology

The methodology used in this paper to estimate $T_a$ is an extension of the study of Jones et al. (1998, hereafter JPD), which determines surface specific humidity and air temperature with artificial neural networks. The methodology involves two phases: data classification and artificial neural net (ANN) development.

a. Data Classification

The data classification requires some preliminary sorting and classification to ensure adequate observations for both development and testing analyses. First, the initial data sets of

1 Institute for Computational Earth System Science (ICESS)
2 Department of Geography
University of California, Santa Barbara, CA 93106

Copyright 1998 by the American Geophysical Union

0094-8276/98/1998GL900086S05.00
W (SSM/I), SST (NCEP), $T_a$ (SMD) were divided into two sub-samples. The 47-month period including January 1988-June 1988, July 1989-June 1992 and July 1993-November 1993 was used to train the neural networks (Sample-I). The 24 month period of July 1988-June 1989 and July 1992-June 1993 was reserved for evaluation of the methodology (Sample-II).

The main goal of this study is the derivation of empirical relationships relating W and SST to $T_a$. We developed an algorithm that takes triplets of observations [W (SSM/I), SST (NCEP), $T_a$ (SMD)] from Sample-I and from all grid cells in the analysis. The next step consists of computation of a sampling parameter, defined as $\alpha = \sigma / (N)^{1/2}$, based on surface marine observations of $T_a$. In this expression $\sigma$ is the standard deviation of the climatological monthly averages, and N is the effective number of ship observations used to compute the monthly average. This sampling parameter is defined to select triplets of (W, SST, and $T_a$) observations that will be used in the ANN training. Only triplets meeting a cutoff value of $\alpha = 0.6$ are used (280 x 10^3 triplets).

b. Artificial Neural Network (ANN) Approach

An artificial neural network (ANN) algorithm is applied to derive transfer functions relating the pairs of (W, SST) observations (inputs) to the surface marine observations of $T_a$ (outputs). The type of ANN used in this paper, feed-forward multilayer perceptron with backpropagation, is completely described by its architecture and the value of the weighted connections between nodes. By presenting the neural network with known input-output patterns, it "learns" a function which performs the mapping embedded in the presented patterns. Clearly, the quality of the input-output patterns used determine the performance of the final network. Ideally, this training set would be free of noise and span the expected range of both input and output parameters with a probability distribution similar to that found in nature. The derivation of the transfer functions is performed in two phases. In the first phase, the ANN used is a supervised fully connected feed-forward network employing a 2-6-2-1x architecture*. It consists of two inputs (W and SST), six nodes in the first hidden layer, two nodes in the second hidden layer and one output ($T_a$). The "x" indicates that there are additional connections between the input and output layers. In order to "train" the ANN, a sub-set of 70 x 10^3 triplets of (W, SST, and $T_a$) are randomly taken from the available data meeting the selection criteria described in the previous section. This sub-set is designated as "training set", whereas another randomly taken set of 10 x 10^3 triplets constitutes a "testing set". The training set is run through the network with the weights at each node of the ANN updated using a back-propagation algorithm (Riedmiller, 1994). Conversely, the testing set is used to independently monitor the mean error between the observed and estimated $T_a$ values. The first phase of the ANN training provides monthly mean maps of $T_a$ for the Sample-I data period. These maps are compared with observed $T_a$ from the surface marine data to produce a set of mean biases (SMD - ANN) for each calendar month. In the second phase, the mean biases are included as a third parameter to train a second ANN. The second ANN consists of a 3-6-2-1x architecture with inputs W (SSM/I), *SST (NCEP) and bias (ANN), and $T_a$ as output.

4. Results

a. Global Results

Taking SMD as our standard, we compared our ANN $T_a$ with other available $T_a$ products, NCEP and TOVS, over the ocean from 70°N to 70°S (even though some results are presented only for the 50°N to 50°S region). Since NCEP has 2.5° x 2.5° resolution, all four fields are averaged up to 5° x 5° resolution. Then using the 24 month Sample II period, rms and biases are calculated. These are then presented as zonal averages in Fig. 1a (bias) and Fig. 1b (rms) for ANN (solid line), NCEP (dotted line) and TOVS (dashed line). Figure 1c shows the spatial distribution of the rms difference between SMD and ANN. The rms differences have a global mean of 0.73°C and standard deviation of 0.37°C. The spatial distribution shows small magnitudes except over: 1) the Southern hemisphere, south of 30°S, where the differences reach large values, and 2) the equatorial Pacific and the Western boundary currents, where values of about 1°C are found. The zonal averages of both the bias and rms (Fig. 1a and b) show that the ANN method provides $T_a$ values closer to those measured by the ships than either the NCEP or TOVS products. The largest differences are in the high latitudes of both hemispheres. For the region within 50°N and 50°S the rms differences are less than 0.5°C. We contend that the degradation of our results in high latitudes is due to the lack of data coverage in these regions. This is supported by the distribution of ship observations over the 24-month period (not shown).

---

* This study uses computer software developed by 1veters (1996): Professional Basis of AI Back-propagation. Additional information can be found at the internet address: www.1veters.com

---

Figure 1. Evaluation of ANN methodology to derive surface air temperature. (a) mean bias for $T_a$ derived from ANN (solid), NCEP (dotted) and TOVS (dashed) by comparison with $T_a$ observations. (b) Root-mean-square differences (rms) between ANN, NCEP and TOVS, and SMD data. Statistics are computed for the Sample-II period (see text for explanation). (C) spatial distribution of rms difference between ANN and SMD.
5. Summary and Conclusions

We have presented a new method to derive monthly averaged surface air temperature. The innovative aspect of this study is the ANN approach, which is trained with satellite (precipitable water), numerical weather prediction analysis (SST) and ship observations (\(T_a\)). The accuracy of the present approach is as good as, if not better, than the best current methods. The next improvement in such methodology will come in the estimation of \(T_a\) with increased temporal resolution.

Figure 3. (a) Spatial distribution of mean bias between ANN methodology and buoys measurements (shading). (b) Spatial distribution of mean bias between NCEP products and buoys measurements (shading) hash marks denote negative values. c Spatial distribution of rms difference between ANN methodology and buoys measurements (shading). d Spatial distribution of rms difference between NCEP products and buoys measurements (shading). In all panels, contours indicate the mean sea surface temperature with 1°C interval. Triangles indicate the TOGA-TAO buoys position. Period: January 1988-August 1997.

Figure 2. (a) Scatter plot of monthly averaged surface air temperature from ANN methodology and buoys measurements (TOGA-TAO). Period: January 1988-August 1997 (n=4137 observations). (b) same as in (a), but for NCEP (January 1988 - December 1996). (c) Time evolution of monthly mean bias (squares) and rms (crosses) between ANN and buoys surface air temperature (averaged over all buoys). Solid lines denote 5-month running averages. Solid and dashed bars indicate the Sample-I and Sample-II periods, respectively.

b. Equatorial Pacific Results

The second evaluation and more stringent test of the new approach, has been performed for the tropical Pacific ocean by comparing \(T_a\) obtained from the ANN method and observations from the TOGA-TAO array. Figure 2a and b show scatter plots of monthly averages from TOGA-TAO versus the ANN retrievals and NCEP for the period January 1988 through August 1997. The comparison for the ANN data set is excellent and slightly better than that for the NCEP data, with a correlation squared of 0.93, and mean bias and rms differences of 0.37°C and 0.61°C, respectively. Figure 2c displays the time evolution of the bias and rms differences between the ANN and buoy data sets. The solid and dashed bars at the bottom indicate the Sample-I and Sample-II periods, respectively. The bottom figure clearly shows that, as time increases, the bias and rms decrease. These results indicate that, as both the time sampling and the quality of the buoy data set improve with time, the accuracy of the ANN method is among the best methods presently available to estimate \(T_a\) from remote sensing observations.

The underlying causes of the differences between the ANN, NCEP and the buoy data sets can be further understood by examining the spatial distribution of the bias and rms between the ANN and NCEP and the buoy measurements. Figure 3a, b, c and d present the spatial distribution of the respective biases and rms (shading) with the mean SST field (contours) to indicate the regions of strong horizontal temperature gradients. Both the ANN bias and rms have their largest values (1.0°C for mean bias; 1.2°C for rms) in the region of strongest mean horizontal temperature gradients (i.e. the northern part of the cold tongue in the eastern Pacific). Another region of large discrepancy for the ANN data, though much smaller, is the northern part of the western Pacific warm pool, where the SST does not change as much, \(T_a\) will most likely be affected by the prevailing deep convection, which has no specific characteristic feature on the northern side of the warm pool. These bias and rms of 0.6°C and 0.8°C, respectively, are rather difficult to explain. Both the NCEP bias and rms have larger values than the ANN in the central tropical Pacific.
Acknowledgments. We thank Dr. Tim Liu (NASA/JPL), Dr. J. Dodge (NASA), Dr. Arlindo da Silva (NASA/GSFC), Dr. Michael J. McPhaden (TOGA-DAO project) and Don Tveten for support in several parts of this work. This work has been supported by grants from NASA/JPL (959177) and NSF (ATM-9310483).

References


Corresponding author address: Dr. Catherine Gautier, Institute for Computational Earth System Science (ICCESS), University of California, Santa Barbara CA 93106-3060. Email: gautier@icess.ucsb.edu

(Received February 18, 1998; accepted September 21, 1998.)