[Reprinted from BULLETIN OF THE AMERICAN METEOROLOGICAL SOCIETY, Vol. 77, No. 6, June 1996]
Printed in U. S. A.

Description of a Complete (Interpolated) Outgoing Longwave Radiation Dataset

1. Introduction

Estimates of outgoing longwave radiation (OLR) from National Oceanic and Atmospheric Administration (NOAA) polar-orbiting satellites (Gruber and Winston 1978; Gruber and Krueger 1984) are often used to distinguish areas of deep tropical convection and to estimate the earth's radiation budget. The National Environmental Satellite Data and Information Service (NESDIS) of NOAA archives the data onto 2.5° lat × 2.5° long grids. The two grids per day correspond to the daytime and nighttime orbits. The local equatorial crossing times have changed over the years (Gruber and Krueger 1984).

One problem in using these data is that missing grids, and missing values within grids, are often present, presumably owing to satellite problems, archival problems, or incomplete global coverage. When a grid is incomplete, but not entirely missing, the missing values are completed by a nearest neighbor spatial interpolation. These interpolated values are flagged as negative. Missing values tend to occur in "swaths." Thus, when the number of missing values is large, the resultant errors are large as well, because of the large distance over which a value is interpolated. This problem is evident in Fig. 1a, which shows the NESDIS estimate of the OLR field for the daytime pass of 24 January 1979. On this date 6837 (out of 10 224) values were flagged as interpolated. Clearly, the field is too smooth east of 90°E. Figure 1b shows the field after missing values have been interpolated by the algorithm described below, in which missing data are primarily interpolated in time. This field is seen to appear more realistic and at least captures the low-frequency phenomena.

The point of this note is to describe a dataset that is available to the public. Missing values have been removed by temporal and spatial interpolation, except for the gap between 17 March and 31 December 1978.

2. Preliminary analysis

Before interpolation, we remove any data that are likely to be erroneous, subject to the following criteria:

1) less than 50 W m^{-2}
2) from 90°–60°N, and from 45°–90°S; daytime values greater than 325 W m^{-2}; nighttime values greater than 300 W m^{-2}
3) from 57.5°N–42.5°S; daytime values greater than 400 W m^{-2}; nighttime values greater than 300 W m^{-2}.

The data are next subjected to a spatial "buddy" check, such that any value cannot deviate too far from the surrounding grid points. The criterion differs depending on how many of the surrounding values are missing. If none of the eight values (including the grid points on the diagonals) surrounding a grid point are missing, then the center value is set to missing if it is more than 49 W m^{-2} above or below any of its neigh-

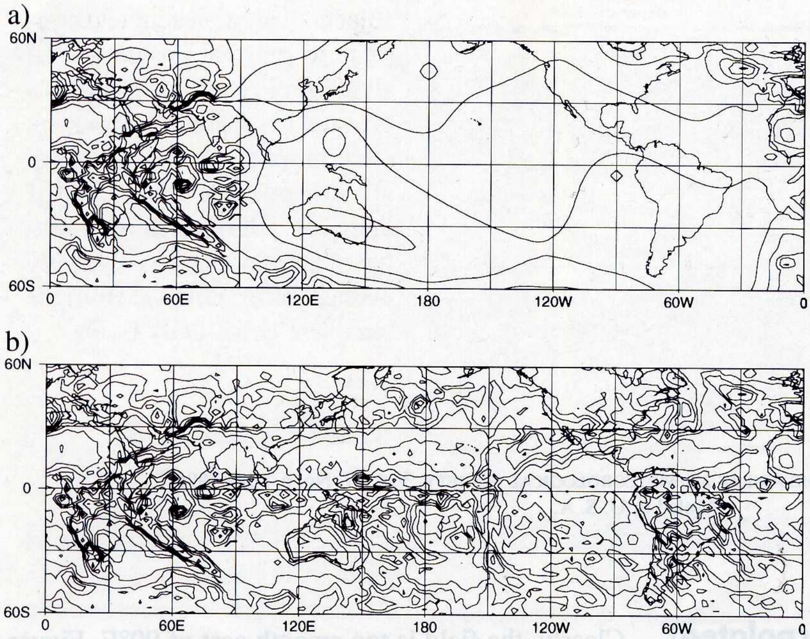
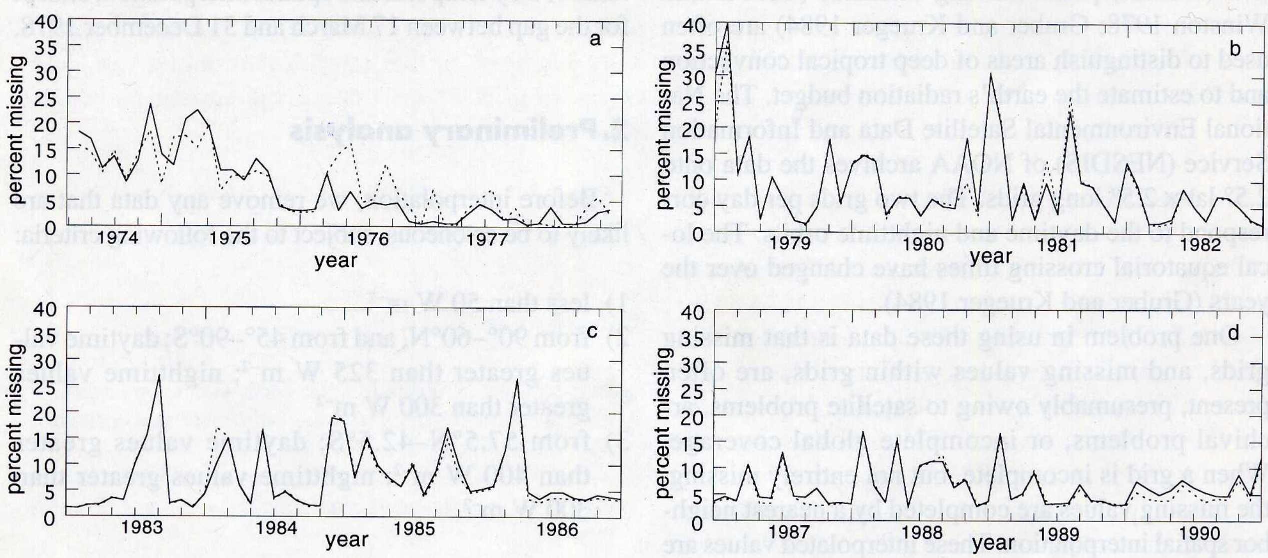


FIG. 1. OLR field for daytime crossing of 24 January 1979. Contour interval is 25 W m^{-2}. (a) Missing values interpolated spatially by NESDIS. (b) Missing values interpolated temporally and spatially by technique decribed in text.

bors. If one surrounding value is missing, the limit is ±52 W m^{-2}, two missing = ±55 W m^{-2}, etc. This check continues until a maximum of five surrounding values may be missing, in which case the limit is ±64 W m^{-2}. There is no check if six–eight points surrounding a point are missing. The rationale for increasing the difference limit as the number of surrounding values decreases is that we believe it is unlikely that if all surrounding values are close together, the value at the center point should deviate significantly from any of its neighbors. On the other hand, if a grid point is on the edge of a region of missing values, and it is also a region of large gradient (e.g., a land–ocean boundary), there is a distinct possibility that the value at a grid point should actually be quite different than its neighbors.

Figure 2 is a time series of the percent of missing values for each month from 87.5°N to 87.5°S for both daytime and nighttime crossings for 1974–94. The percent missing from 30°N to 30°S (generally considered to be the region where low OLR values are indicative of convection) is quite similar to the global value. Figure 3 shows the



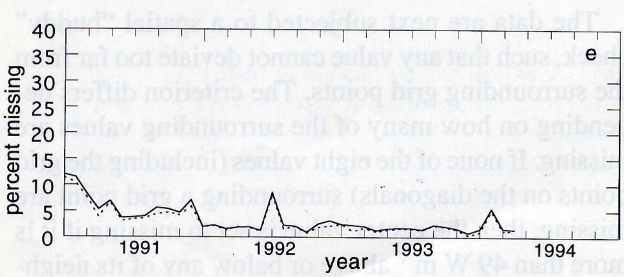


FIG. 2. Time series of percent of missing values for each month. Solid curve represents daytime crossing and dashed curve represents nighttime crossing: (a) 16 March 1978–1 June 1994, (b) 1979–82, (c) 1983–86, (d) 1987–90, (e) 1991–94. Count includes values at poles (put on grid) and values deemed unrealistic (see text).

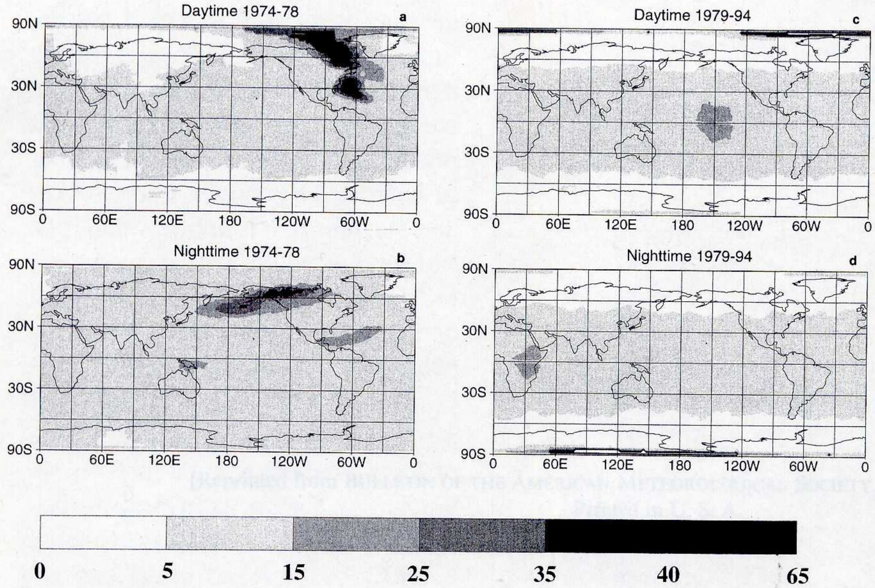


FIG. 3. Percent of missing values for (a) 1974–78 daytime, (b) 1974–78 nighttime, (c) 1979–94 daytime, and (d) 1979–94 nighttime.

percent missing over the record (divided into 1974–78 and 1979–94 data) for both daytime and nighttime crossings. About 50% and 30% of the missing values come from grids that are entirely missing for 1974–78 and 1979–94; the number is roughly equal for daytime and nighttime crossings. The remainder of the missing values were either flagged as spatially interpolated by NESDIS or have been screened by our algorithms.

3. Method of interpolation

The intent of interpolation technique is to minimize the distance in space or time over which a value is interpolated. All interpolations are done for daytime and nighttime crossings separately. We first linearly interpolate in time if there is a 1-day gap. This procedure fills in about 65% and 45% of the missing grid points for 1974–78 and 1979–94, respectively.

Next, we spatially interpolate by averaging the surrounding values (not including diagonal) if three of the four values are present. This is followed by another temporal 1-day interpolation. We then interpolate spatially if two or more of the closest grid points have values.

To fill in the few remaining missing values, we linearly interpolate temporally for 3 or fewer missing days. We then interpolate in space if there is but one adjacent value present and continue to run that program until all missing values are filled. We then run the spatial buddy check once more to make sure there are no spurious values and if there are, fill them in with spatial interpolations. Finally, daily averages are computed from the complete twice-daily fields.

4. Description and availability of data

The datasets are available in both SUN binary and netCDF formats from the anonymous FTP site ftp.cdc.noaa.gov under/Datasets/interp_OLR. See the README file in that directory for more information. Any changes made to the algorithm described above will be explained in the README file. Further information also is available at the URL http://www.cdc.noaa.gov/cdc/data.interp_OLR.html.

Other available datasets are the twice-daily interpolated and uninterpolated fields, and interpolated daily averages, further interpolated onto a 128 × 64 Gaussian grid. Please contact the authors to access these datasets.

Acknowledgments. We wish to thank Julia N. Paegle for suggesting this note.

References

Gruber, A., and J. S. Winston, 1978: Earth–atmosphere radiative heating based on NOAA scanning radiometer measurements. *Bull. Amer. Meteor. Soc.,* **59,** 1570–1573.

——, and A. F. Krueger, 1984: The status of the NOAA outgoing longwave radiation data set. *Bull. Amer. Meteor. Soc.,* **65,** 958–962.

BRANT LIEBMANN AND CATHERINE A. SMITH
COOPERATIVE INSTITUTE FOR RESEARCH IN ENVIRONMENTAL SCIENCES
NOAA/ERL/CDC
UNIVERSITY OF COLORADO
BOULDER, COLORADO