Overview of the Integrated Global Radiosonde Archive

By

Imke Durre, Russell S. Vose, and David B. Wuertz

National Climatic Data Center, Asheville, North Carolina

Submitted to the

Journal of Climate

October 2004

Corresponding author's address:

Imke Durre
National Climatic Data Center
151 Patton Avenue
Asheville, NC 28801
Phone: 828.271.4870
Fax: 828.271.4328
E-mail: imke.durre@noaa.gov

Abstract

This paper provides a general description of the Integrated Global Radiosonde Archive (IGRA), a new radiosonde dataset from the National Climatic Data Center (NCDC). IGRA consists of radiosonde and pilot balloon observations at 1536 globally distributed stations with varying periods of record during 1938-present. Observations include pressure, temperature, geopotential height, dewpoint depression, wind direction, and wind speed at standard, surface, tropopause, and significant levels.

IGRA was created by merging data from 11 different sources and applying a suite of quality assurance procedures to the resulting dataset. During the merging process, duplicate levels within soundings were eliminated, and one sounding was selected for every station, date, and time. The quality assurance algorithms checked for format problems, physically implausible values, internal inconsistencies, climatological outliers, and runs of values across soundings and levels. The performance of the various checks was evaluated by careful inspection of selected soundings and time series.

In its final form, IGRA is the largest and most comprehensive dataset of quality-assured radiosonde observations freely available. Its temporal and spatial coverage is most complete over the United States, western Europe, Russia, and Australia. The vertical resolution and extent of soundings improve significantly over time, with approximately two thirds of all soundings reaching up to at least 50 mb by 2003. IGRA data are updated on a daily basis and are available online from NCDC as both individual soundings and monthly means.

1. Introduction

Upper-air observations of temperature, humidity, and wind are critical to many applications in the atmospheric sciences, including numerical weather prediction, operational weather forecasting, model verification, climate monitoring, and research. Although satellite measurements have become a major source of information on the atmosphere, radiosonde observations continue to play an important role in all of these applications and are critical to the verification of satellite measurements and satellite-derived products (Finger and Schmidlin 1991; NRC 2000; Free et al. 2002; Durre et al., submitted to Bull. Amer. Meteor. Soc.). Furthermore, they constitute the only source of upper-air observations prior to the 1970s and frequently provide a higher vertical resolution in critical layers of the atmosphere such as the planetary boundary layer.

Radiosondes have been launched on a daily or twice-daily basis at points around the globe since the 1940s. During its one- or two-hour ascent from the surface into the stratosphere, a radiosonde transmits its measurements to ground receiving stations where they are processed into pressure, temperature, dewpoint depression, and geopotential height. Wind direction and speed are obtained by tracking the position of the balloon during its ascent. Thermodynamic and wind observations may be provided at mandatory pressure levels, additional required levels, significant levels, and certain fixed height increments. Mandatory pressure levels include those specified by the World Meteorological Organization (WMO) (1000, 925, 850, 700, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30, 20, and 10 mb) as well as those additional levels suggested by the US National Weather Service (7, 5, 3, 2, and 1 mb) (FCM-H3 2004; WMO 2004). Surface observations taken at or near the launch site are included in the sounding as a "surface level". Conforming to standards set forth by the WMO, the radiosonde, wind, and surface measurements are compiled into a report that is transmitted as a binary-coded message over the Global Telecommunications System (GTS) to various regional and national meteorological centers around the world where they are processed, archived, and redistributed to other locations (WMO 1996).

The goals of the Integrated Global Radiosonde Archive (IGRA) project are (1) to combine as many sources as possible into one radiosonde data archive, (2) to develop and apply quality-assurance algorithms that remove any gross errors in the data, (3) to put into place an automatic system for updating the resulting archive on a daily basis, and (4) to provide unrestricted online access to the data. For Version 1 of the dataset, we have focused our efforts on creating robust procedures for merging data from various sources and on developing algorithms for quality assuring temperature. The emphasis on temperature was motivated by the fact that this element is both central to climate monitoring activities as well as less variable, and therefore more easily quality assured, than humidity and wind measurements. This paper provides an overview of the methods used to merge the various sources and quality-assure the data values. A more detailed description of the quality-assurance procedures applied to temperature will be provided in a separate paper.

The remainder of the article is organized as follows. Section 2 describes the selection of data sources and stations for IGRA as well as initial processing steps applied to each source dataset. The strategies employed in eliminating duplicate levels and soundings and merging the various sources are presented in Section 3. An overview of the quality assurance procedures is given in Section 4. Section 5 contains a description of the resulting dataset. A summary of the work is provided in Section 6.

2. Selection of Data Sources and Stations

IGRA constitutes a compilation of eleven source datasets (Table 1). The core of IGRA is formed by four GTS-based datasets that were preprocessed at one of three locations in the United States: National Climatic Data Center (NCDC, 1963-1970 and 2000-present), the National Center for Atmospheric Research (NCAR, December 1970-1972), and the National Centers for Environmental Prediction (NCEP, 1973-October 1999). These datasets have nearly consecutive periods of record and will be collectively referred to as the core data sources. Two additional sources are collections of GTS reports that were pre-processed at the Australian Bureau of Meteorology between 1990 and 1993 (Australian GTS) and at the All-Russian Institute for Hydrometeorological Information between 1998 and 2001 (Russian GTS). For a variety of reasons, including differences in decoding practices, some messages transmitted over the GTS are decoded only at certain receiving centers and not at others. Thus, even though extensive duplication generally exists among the core, Australian, and Russian GTS data, the latter two sources occasionally supply soundings that are either not present or incomplete in the core data. Another data source with extensive spatial coverage is a dataset compiled by the United States Air Force that consists of both GTS reports and non-GTS data. With a period of record of 1946-1973 for stations included in IGRA, this dataset extends the records of many stations back in time from the 1960s to the 1950s or 1940s. The temporal completeness and vertical resolution of data at stations in the United States, Australia, Argentina, and South Korea is further enhanced by four country-specific sets of data that were archived before their transmission over the GTS and thus may contain levels and variables not found in the GTS reports.

The above-mentioned sources were selected for inclusion in IGRA based on timely availability of the data, availability of documentation for codes and conventions used, and the level of agreement with data from other sources during periods of overlap. Updates to three of the sources (NCDC real-time GTS, country-specific U.S., and Russian GTS datasets) continue to be received at NCDC on a regular basis. The remaining source datasets were obtained through contacts with the relevant institutions and archived at NCDC at various times during the 1990s, and so their records end during or prior to that decade. The types of variables, number and types of data levels, and precision of data values vary with data source, station, and time. Furthermore, the source datasets were subjected to different sets of quality checks prior to or during their ingest at NCDC.

Each source dataset contains up to three types of stations: (1) land-based stations with WMO numbers, (2) other (non-WMO) land-based stations, and (3) fixed and mobile ships. In most data sources, stations are identified only by their station number and location, and no station inventory is available. The process of identifying stations can be complicated by errors in station location and differences in station numbers assigned to the same station across different data sources. Since the most reliable station information is available for active WMO stations with fixed coordinates, we chose to focus on land-based stations with data in the NCDC real-time GTS and to supplement these with other identifiable stations that significantly enhance the temporal and spatial coverage of the dataset. Three primary sources of station information were used to determine the name, country, latitude, and longitude of each station: the list of stations used during the processing of GTS messages at NCEP and NCDC; the station inventory of the Global Historical Climatology Network (GHCN, Peterson and Vose 1997); WMO Publication 9 Volume A (WMO 2004). When entries were found in more than one list, the various sets of coordinates were intercompared. In the rare case in which a station number could not be located in the three primary lists, latitudes or longitudes from the various lists differed by more than half a degree, or the coordinates placed the station in an incorrect country, online searches were used to determine any necessary corrections.

For each station to be included in IGRA, any data from the four consecutive GTS datasets (Table 1) were concatenated into one core time series maximally spanning the period September 1963 to present. In those cases in which data for December 1970 are available from both the NCDC historical GTS and the NCEP/NCAR GTS, only the data from the NCDC historical GTS were used for that month. December 1970 data from the NCEP/NCAR GTS were utilized only when no data were available from the NCDC historical GTS. Many of the concatenated core records contain a 2.5 month break between the end of the NCEP/ NCAR GTS in October 1999 and the beginning of the NCDC real-time GTS in January 2000. This gap was, in many cases, filled in with data from other sources.

In order to check that station numbers are assigned consistently across data sources, e.g., that station number 01001 refers to the same station in the core and Russian GTS datasets, the data for each station in the core GTS data were compared to the data for each station number in every other data source. For this comparison, the two highest-precision variables, temperature and wind speed, at the five most common mandatory levels, 850, 700, 500, 400, and 300 mb were used and the percentage of identical values was computed for each station pair with a common period of record. Since differences in processing procedures among the various data sources result in minor differences among their data values, temperature and wind speed values were considered to be identical if their differences fell within the similarity thresholds listed in Table 2. A mismatch was identified either when fewer than 90% of the values in two records for the same station number were found to be identical or when a significant percentage of compared values from two records of different station numbers were identical to each other. An investigation of such mismatches revealed that they generally occur for one of the following reasons: (1) a site changed station number at some time during its operation, and the timing of the change differed in the source datasets; (2) a site was assigned to a WMO station number in one data source and to a non-WMO number in another; (3) extensive duplication of data exists between two stations within the same data source. Depending on the nature of the mismatch, the availability of station history information from reliable sources (e.g., Elliott et al. 2002), and the relative importance of the stations involved, data from all or part of one station were reassigned to the other station number or one or both stations were excluded from the dataset. To facilitate data comparisons and subsequent quality assurance, a station was further required to have at least 100 soundings and a valid station elevation. The spatial and temporal coverage of the resulting IGRA station network will be discussed in Section 5.

3. Duplicate Elimination

The goals of duplicate elimination are to generate a dataset in which there are no occurrences of (1) multiple sets of values for the same station, date, time, and level, (2) duplication of data in consecutive soundings, and (3) duplication of data in concurrent soundings from different stations. Cases of level duplication within one sounding are addressed by removing any data values that differ among the duplicate levels and combining the remaining data into one level. Duplication of soundings may stem from transmission and processing errors or from the recurrence of the same sounding in multiple data sources. When data for a particular station are available from two or more sources, they are combined into one record. Where multiple soundings with the same time stamp occur, the sounding with the largest number of values is chosen. Of the more than 30 million soundings processed, approximately one quarter contained duplicate levels, with an average of three such levels per sounding. Discrepancies in data values, however, are found only at a few percent of these duplicate levels.

Some data sources report the nominal observation time (e.g., 00 UTC) as the observation hour, while others report the hour closest to the launch time (e.g., 23 UTC), resulting in the existence of residual duplicate soundings. This fact necessitates the retention of the sounding with the largest number of values when identical soundings occur consecutively within two hours of each other. Consecutive identical soundings whose time stamps are more than 2 hours apart are discarded as the duplication of their data is considered erroneous. In order to allow for differences in data processing and precision, two soundings from different sources are considered identical if at least 90% of the absolute differences between values at levels common to both soundings fall within specified similarity thresholds (Table 2).

With the purpose of identifying cases in which identical soundings are reported simultaneously at more than one station, the mandatory-level 850-to-300 mb data of concurrent soundings from all stations are compared. Approximately 60 000 soundings (0.2%) were identified as interstation duplicates and removed from the dataset.

4. Quality Assurance

The quality of radiosonde data is compromised by a variety of observation, transmission, and processing problems that manifest themselves in: the occurrence of physically implausible or inconsistent values; inconsistencies between the reported surface pressure and the archived station elevation; the presence of various types of inconsistencies among data amounts or levels;

incomplete soundings; and artificial breaks or drifts in time series (Schwartz and Doswell 1991; Gandin et al. 1993; Gaffen 1994). Quality assurance procedures for sounding data generally rely on principles of internal consistency, basic physical relationships such as the hydrostatic balance, or statistical methods (Kahl et al. 1992; Loehrer et al. 1996; Collins 2001a,b). Some approaches employ a decision-making algorithm that takes into account the results of multiple tests, while others apply a sequence of independent checks. For its greater transparency, we chose the latter approach.

The IGRA quality assurance procedures can be grouped into six general categories: basic plausibility checks, internal consistency checks, checks for the repetition of values, climatologically-based checks, checks on the vertical and temporal consistency of temperature, and data completeness checks. Various format checks are also performed to ensure that the structure of the data files as well as any codes employed are consistent across all files and adhere to the standards set forth in the documentation. The first three categories of checks eliminate any values that might compromise the performance of subsequent algorithms. Most of the procedures in these categories have the added advantage of using universal parameters and, therefore, can be applied regardless of the temporal resolution or completeness of records. The remaining procedures are based on station-specific climatological means and standard deviations (STDs) and are applicable only when sufficient data are available for computing the required statistics. Climatologies are computed only for geopotential height, temperature, and surface pressure since their frequency distributions tend to follow the normal distribution more closely than do the distributions of dewpoint depression and wind measurements. Some additional checks on vertical and temporal consistency that also require climatological statistics were developed specifically to further improve the quality of temperature data. The data completeness checks ensure that only those levels, soundings, and station records with a sufficient amount of data are included in IGRA.

The following subsections describe the IGRA quality assurance procedures. Table 3 lists the variables involved in each check and the types of items removed. In constructing IGRA, we have taken the approach of moving a value, level, or sounding that fails any of the quality assurance procedures from the data file to a log file. Except in the case of elevation, no replacement value is provided. The IGRA log files are available upon request from the author.

a. Basic plausibility checks

The basic plausibility checks determine whether the date, observation hour, and actual launch time of a sounding as well as the data values at each level fall within certain gross plausibility limits (Table 4). The date check, which is applied during the process of reformatting each source dataset, identified occasional instances of invalid days of the month (e.g., April 31) and soundings with missing observation hour. Soundings with such invalid dates or times were excluded from further processing. In addition, 0.25% of all release times and 0.025% of all data values were found to be implausible.

b. Internal consistency checks

Internal consistency checks compare the release time and observation hour, geopotential heights at different levels within a sounding, pressure and geopotential height at any one level, as well as temporal variations in station elevation and surface pressure. An additional check removes wind values when the wind speed is equal to 0 and the wind direction is neither 0 nor 360 degrees. In the release time consistency check, soundings are deleted if the actual launch time deviates by more than three hours from the observation hour. Differences of such magnitudes are found in approximately 0.25% of all soundings. The South Pole station is excluded from this check because, in recent years, launch times consistently have differed from the observation hour by three to four hours.

Two algorithms were developed to evaluate the physical consistency of the pressure and geopotential height at any level and eliminate problems such as those found in the sounding shown in Figure 1. The "hypsometric check" is similar to a hydrostatic check (Gandin et al. 1988), but is independent of the temperature profile within the sounding examined. The range of plausible pressure values for any given height is determined from the hypsometric equation using the extreme values of the average temperature of the atmospheric layer between the surface and the level in question. The extremes of the layer-average temperature are computed using the lapse rates from the 1976 U.S. standard atmosphere and assuming surface temperatures of -60°C for the cold extreme and 60°C for the warm extreme. Given these parameters, the hypsometric check removes levels whose pressure and geopotential height are grossly inconsistent with each other, such as 30-mb levels with geopotential heights of 0 and surface levels with geopotential heights of 3000 m (Fig. 1). This check, however, does not guarantee the monotonic increase of geopotential height with decreasing pressure. To ensure that this basic relationship holds true in all soundings, an iterative multi-step procedure was designed in which the height of each pressure level k is compared to the height of every level j with a higher pressure. If the geopotential height of level k is found to be less than or equal to the geopotential height of level j, the numbers of violations for levels j and k are each incremented by 1. Once all possible pairs of levels within the sounding have been compared, levels with the largest number of violations are removed from the sounding. The comparison and level deletion process is repeated until no more violations are found. Based on both the hypsometric and height sequence checks, approximately 0.1% of the 800 million levels in the dataset were removed.

Another series of internal consistency checks insures that a sounding contains at most one valid surface level. In soundings in which more than one surface level remains after the hypsometric check, all such levels are deleted. In the rare case in which the only level marked as a surface is a level with only height and wind values, the surface level indicator of that level is changed to a generic wind level indicator, so that in the final dataset, a sounding contains either one surface pressure level or no surface at all. Monthly median elevations are generated for each station, year, and month from the elevations provided in the soundings. Depending on the source of the sounding, the elevation and surface height information was either provided by the source or taken from various station lists during initial processing at NCDC. In IGRA, the surface elevation in each sounding is replaced by the monthly median elevation, or set to missing if no adequate monthly median elevation can be determined. A height-only wind level having a height equal to either the original elevation or the monthly median elevation is merged with a surface pressure level.

The use of the monthly median eliminates isolated errors in individual soundings and reduces sporadic variations caused by merging data sources reporting different elevations. However, as illustrated by the time series of monthly median elevations at Atyran, Kazakhstan (Fig. 2), spurious variations may remain. To minimize the number of such spurious elevations, we identified grossly inaccurate elevations by comparing each station's mean elevation to the elevation expected from the station's mean surface pressure, the elevation reported in WMO Publication 9 Volume A, and the elevation of the nearest grid point in the Global Land One-kilometer Base Elevation (GLOBE) dataset (NGDC 2004). Atyran, for example, has a WMO elevation of -28 m, a GLOBE elevation of -37 m, and a mean surface pressure of 1018 mb. Thus, the elevations around 3000 m up to the early 1960s, around 500 m in the mid-1970s, and around 10 000 m in 1982 (Fig. 2) are incorrect for the station. In addition, time series of monthly median elevations for each station were examined for significant (>~50 m) discontinuities, for cases in which a station seemed to remain at the same elevation for no more than a few months, and for inconsistencies with corresponding timeseries of surface pressure. Using the various pieces of information, we identified specific periods at certain stations with implausible elevations, setting the respective surface level heights to missing. In the process of inspecting surface time series for all IGRA stations, we further noted gross temporal inconsistencies in the variations of surface pressure and temperature during 1968-70 at stations in the former Soviet Union and during 1991-97 at Chinese stations and remove all affected surface levels from IGRA.

Several of the data sources contain levels whose pressure or geopotential height is clearly below the surface pressure or elevation of the station. In most cases, these levels consist of data that have been extrapolated from the surface down to any mandatory pressure that happens to fall below the surface. When such extrapolated levels are identified as below-surface levels by the level type indicator in the original sounding, they are automatically deleted by the IGRA QC process. Since not all extrapolated levels are properly identified in the source datasets, and transmission errors can also produce below-surface levels, an additional check was designed to identify and remove all types of below-surface levels. In a sounding with a surface level that passes the above mentioned consistency checks, a pressure level is considered to fall below the surface if its pressure is higher than the surface pressure or its geopotential height is less than the surface height. In a sounding without a valid surface level, any pressure level whose geopotential height is at least 10 m below the median elevation of the current month is removed.

c. Checks for the repetition of values

The next set of checks looks for runs of values in time and in the vertical. A run is defined as the repetition of a value over a certain number of consecutive soundings or levels, ending with a change to another non-missing data value; the absence of a value in a sounding or level does not interrupt a run. Before choosing thresholds for the identification of runs of unreasonable length, we generated summary statistics on the frequencies of runs extending over four or more values, manually inspected selected runs, and examined the types of values being repeated in the context of the surrounding data.

The generic check for runs across soundings eliminates runs in surface pressure, surface- and mandatory-level temperature, and mandatory-level geopotential height that are more than 15 values in length. The same temporal runs check was also performed separately for each hour of the day, e.g., 0 UTC only. Dewpoint depression, wind speed, and wind direction are excluded from the check for runs in time as the lower precision of their valuees results in the more common occurrence of runs in these variables.

Two procedures identify runs of values across levels in a sounding, namely, a single-variable check and a pairwise check. The single-variable check analyzes surface/mandatory and significant levels separately and looks for temperatures of the same value across at least five consecutive levels in either group. This procedure is applied only to temperature since the duplicate removal and height sequence checks have already eliminated any vertical runs in pressure or height, and runs in the other variables are considerably more plausible than equivalent runs in temperature. The pairwise vertical run check identifies the repetition of the same value in either temperature and dewpoint depression or wind direction and speed over at least five consecutive levels. This check is performed separately on pressure levels and height only levels.

The inspection of individual temporal and vertical runs revealed the existence of five peculiar data problems. These problems consisted of excessively frequent occurrences of certain temperature or geopotential height values within specific geographical regions, periods, data sources, and atmospheric levels. Respective values were eliminated by specifically designed checks as they might otherwise seriously impact the quality of IGRA data. All in all, the various procedures for identifying excessive repetition of values removed approximately 0.02% of all data values. Among the more interesting runs identified are cases of 40 consecutive 1000-mb surface levels, -7.5°C temperatures at nine consecutive mandatory levels between 850 and 30 mb in a sounding, 10 24.4°C temperatures at significant levels between 937 and 429 mb, and 0 wind speed and direction throughout an entire sounding.

d. Climatological checks

A two-tiered set of climatological checks remove geopotential height, temperature, and pressure values that deviate by more than a certain number of STDs from their respective long-term means. In the first phase, the climatological means and STDs are calculated for station and pressure-level, whereas in the second phase, the climatological statistics are further categorized by time of year and time of day. Due to their less stringent data requirement, the tier-1 checks can be applied to a larger number of data values than the tier-2 checks. On the other hand, the tier-2 checks allow for the use of tighter thresholds in the identification of outliers because their STDs do not reflect the seasonal and diurnal variations included in the tier-1 statistics.

Furthermore, the tier-2 statistics are not computed until after the tier-1 checks have been applied and thus are based on a cleaner set of data.

The means and STDs are calculated for the surface and mandatory levels using biweight statistics as described by Lanzante (1996). The biweight statistics tend to be more resistant to outliers which are likely to be present in data that have not undergone advanced quality assurance. For the tier-1 checks, a mean and STD are produced as long as at least 120 values are available for a given station, level, and variable during the station's period of record. For the tier-2 checks, statistics are calculated for 45-day windows centered on each day of the year and in 3-hour windows centered on 0, 3, 6, 9, 12, 15, 18, and 21 UTC, provided that at least 150 values are available for any station, level, and variable in a given time interval. The means and STDs at other pressures were derived as needed by interpolating linearly with respect to the logarithm of pressure between the nearest adjacent mandatory or surface levels. Recognizing that actual changes in temperature with height are not always linear, we compared the statistics derived by linear interpolation to those computed using all available data in 1-mb slabs throughout the troposphere and stratosphere. Inspection of the two types of climatological profiles at a set of 87 globally distributed stations (Lanzante et al. 2003a) revealed that the vertical structure is rather insensitive to the method of computation.

To choose thresholds for labeling values as outliers, we compared soundings and time series prior to the climatological checks to those following the application of the tier-1 and tier-2 checks, using various thresholds between 3 and 7 STDs. We chose thresholds such that the algorithms neither remove a disproportionate number of values within the normal range of variability nor fail to remove a significant number of points that are clear outliers. In the tier-1 check, a threshold of 6 STDs was chosen for all three variables. For the tier-2 check, the threshold is set to 5 STDs for geopotential height, temperature, and below-normal surface pressure and to 4 STDs for an above-normal pressure. The asymmetric thresholds for above- and below-normal surface pressure were set in recognition of the fact that high-pressure anomalies tend to be smaller in magnitude than low-pressure anomalies. These thresholds resulted in removal of approximately 0.1% of all pressure, temperature, and geopotential height values by the tier-1 and tier-2 checks.

e. Additional checks on temperature

The inspection of various time series and soundings revealed that, particularly for temperature, the climatological check alone is incapable of satisfactorily removing all outliers without also removing realistic extreme temperatures. Figures 3 and 4 show examples of a time series and a sounding with outliers that are clearly erroneous when viewed in context with other temperatures within their temporal and vertical vicinity. However, to address outliers that pass the climatological checks, but are vertically or temporally inconsistent, a number of additional vertical and temporal consistency checks were developed specifically for temperature. These checks will be described briefly here and in more detail in a subsequent paper.

Using vertical profiles of temperature z scorers derived from the tier-2 climatological means and STDs, a set of vertical consistency checks eliminate the most common problematic characteristics of temperature profiles, including soundings whose entire temperature profile deviates significantly from normal, soundings whose temperatures fluctuate wildly from level to level, and temperatures that are clearly inconsistent with either the entire profile or the portion of the profile in their immediate vicinity. These procedures augment each other since each has unique strengths as well as specific data requirements.

Two variants of a temporal consistency check are applied to surface and mandatory level temperatures and are based on z scores derived using the overall mean and STD for any station and level, provided that at least 120 such values remain following the climatological and vertical consistency checks. The algorithm looks for temperatures whose z score both exceeds a specified absolute value and differs by a specified number of STDs from all other temperature z scores within a specified time window. Both variants examine only those temperatures whose absolute z score is greater than 2.5 and require that temperatures be available on at least half of the days in the time window. The first identifies outliers that differ by more than two STDs from temperatures within 22.5 days before and after, while the second variant uses a difference threshold of one STD and time window of 2.5 years on either side of the potential outlier. The vertical and temporal consistency checks together remove approximately 0.08% of the temperatures.

f. Checks for data completeness

The IGRA quality assurance process also ensures that the dataset adheres to certain minimum requirements for completeness. Wind speed and direction must always appear together, and a dewpoint depression may exist only if it is accompanied by a temperature at the same level. A pressure level is retained if it contains valid thermodynamic data, valid wind data, or both. Levels with a height, but no pressure, are permitted to exist if they contain valid wind data. A sounding may consist of any combination of pressure levels and height-wind levels as long as there is at least one non-surface level. The "isolated sounding check" eliminates groups of fewer than three soundings surrounded by at least 31 days without data, groups of fewer than

15 soundings surrounded by gaps of three months (92 days), groups of fewer than 28 soundings flanked by gaps of half a year (182.5 Days), and records consisting of fewer than 100 soundings.

IGRA contains a number of stations whose radiosonde observations were reported under two or more station numbers over time. Many such changes in station number occurred without discernible change in station location and were the result of changes in the numbering system used by the WMO (e.g., at Canadian stations in 1977). The compositing procedure merges the records of such stations into one record which is assigned the station number of the most recent station. At stations in the contiguous United States during the 1990s, radiosonde observations were moved from one site to another site close enough to reflect the same regional atmospheric conditions. The records of such stations are also combined, as long as they are located within 150 km of each other, and their periods of record do not overlap. Composite stations are identified in the IGRA station list. For each of the 151 composite stations, the dates and times of the first and last soundings of each original station record and the composite record are listed in an auxiliary documentation file. Users engaged in climate change studies are advised to consider the potential impact of the compositing on their specific analysis, particularly when the emphasis is on the planetary boundary layer.

5. Description of the Dataset

The final quality-assured IGRA dataset contains slightly more than 28 million soundings at 1536 stations during the period 1938 to present. As shown in Table 1, most of the soundings originate from GTS messages. The core data sources contribute approximately 82% of the soundings, while the other large-scale sources contribute 6%, and country-specific datasets contribute 12%. Before the beginning of the first core data source in September 1963, the U.S.

Air Force and country-specific U.S. sources each supply nearly half of the soundings (approximately 48% and 44%, respectively), with the remainder provided by the country-specific sources for Australia (nearly 8%) and Argentina (<1%). From September 1963 through December 2001, between 70 and 95% of the soundings come from the core data sources in any one year, depending on the availability and relative completeness of the other sources. Due to ease of processing, all soundings for the years after 2001 have been taken from the NCDC real-time GTS. Data from the Russian GTS and country-specific U.S. datasets for those years will be added in future annual updates.

As indicated by the map of all IGRA stations (Fig. 5a), the stations are distributed across most areas of the globe. The spatial coverage is most complete in Europe and sparsest in northern Canada, interior Antarctica, and equatorial Africa. However, both the total number and spatial distribution of stations vary considerably over time (Fig. 6). The top line in Fig. 6 tracks, for each year, the number of stations where at least one sounding is reported. The dataset coverage begins with one station in Tasmania in 1938, reaches a peak of 1180 stations in 1991, and then declines to 937 stations active in 2003. A comparison of the map of all stations (Fig. 5a) to the map of stations active in 2003 (Fig. 5b) shows that the distribution of active stations is generally less dense. The most significant difference between the two maps is in Western equatorial Africa, where many stations have closed.

The jumps in the number of stations in 1946, 1963, and 1973 (Fig. 6) are related to changes in the number or type of data sources contributing to IGRA (Table 1). Before the beginning of the core data sources in 1963, IGRA stations are concentrated in the contiguous United States, Alaska, and the former Soviet Union, with additional stations in parts of the North Atlantic, southeast Asia, Argentina, and coastal Australia. With the jump in 1963, coverage of Western

Europe, China, and Japan begins, while many stations in Africa, Brazil, Central Asia, and India do not become available until the late 1960s or early 1970s. No data were available for many Chinese stations between 1973 and 1990. In this period, the core data source, i.e., the NCEP GTS, does not contain any data for most Chinese stations. The sudden increase in the number of stations from 1990 to 1991 in Figure 6 is the result of the resumption of data for these stations.

Figure 6 also shows, for each year, the number of stations where data are available on at least 10%, 50%, and 80% of the days in the year. During the first 20 or 30 years of the dataset, most of the available stations report data on at least 80% of the days. The subsequent decrease in the proportion of stations with rather complete records is a reflection of the addition of stations in parts of the world where observations tend to be more sporadic. At the peak of the time series in the early 1990s, approximately 840, or more than two-thirds, of the available stations report at least one sounding on at least 80% of the days.

Before 1958, the most frequent hours for launching radiosondes were 3 and 15 UTC. Since that year, the WMO has stipulated that observations be made near the times of 0 and 12 UTC. However, a lack of equipment or observers restricts a considerable number of stations to one observation per day, while others are able to take observations more than twice a day. The geographical variation in temporal resolution is illustrated by a comparison of the map of all stations active in 2003 (Fig. 5b) to the map of active stations with a median time difference between consecutive soundings of half a day or less in 2003 (Fig. 5c). While about two thirds of all active stations release radiosondes at least twice daily, most stations in Russia, South America, and parts of Africa, central Asia, and eastern Europe report fewer observations. Some 267 stations in central Europe, parts of Asia, Australia, northwest Africa, and a few other locations launched at least half of their radiosondes in 2003 every six hours. The nominal frequency of observations can also vary over time at any one station. For example, a station in the United Kingdom and another in the western tropical Pacific reported observations every hour for a portion of their record in recent years, but took observations much less frequently in other years.

The variables available at mandatory levels in IGRA generally include geopotential height, temperature, wind direction, and wind speed. Dewpoint depression is usually also available in the lower and middle troposphere, but becomes scarcer in the upper troposphere because of the general practice to discontinue humidity measurements at temperatures less than -40°C (Elliott and Gaffen 1991; Garand et al. 1992). Temperature and dewpoint depression are also available at significant thermodynamic levels. These levels may also contain wind observations, but except in certain country-specific sources, geopotential height is generally absent. The location of wind-only levels in the atmosphere is defined by pressure, by pressure and height, or only by height.

Figures 7-9 display, for each year, the average number of mandatory and total levels per sounding (Fig. 7), the percentage of soundings reaching up to at least 100, 50, or 10 mb (Fig. 8), and the percentage of soundings containing each of the different types of variables (Fig. 9). In the period from 1938 through 1945 during which data are available for only one station, soundings consists of only wind measurements at 1 to 2 mandatory levels (Fig. 7 and 9). As the number of the data sources and stations increases during the next two to three years, the number of mandatory levels begins to increase, soundings occasionally reach as high as 100 mb, temperature becomes a dominant variable, and wind measurements are available in 60 to 70% of the soundings. The average extent of soundings continues to increase from the late 1940s into the 1960s, as some observations at levels other than mandatory pressure levels become available.

During the same time, wind measurements become more and more widespread. By the early 1960s, most of the soundings consist of temperature and wind observations up to the 100-mb level. The addition of large numbers of stations with varying degrees of data completeness accounts for the overall drop in the number of soundings reaching into the stratosphere during the late 1960s and 1970s, even though the vertical resolution of soundings continues to improve, as indicated by the rather monotonic rise in the total number of levels per sounding. This period also features the onset of dewpoint depression observations, first only at a few stations in the early 1960s, then in the vast majority of soundings in 1969. The abrupt decrease in the percentage of soundings containing temperature and dewpoint depression measurements from more than 90% to less than 70% in 1973 coincides with the onset of wind only soundings around the hours of 6 and 18 UTC. From the late 1970s up to present, the availability of thermodynamic observations as well as the vertical resolution and extent of soundings improve significantly. By 2003, the average sounding consists of 11 mandatory and 35 additional levels, more than 80% (35%) of all soundings reach at least a 100-mb (10-mb) level, and 74% of soundings contain temperature observations.

6. Summary

IGRA consists of quality-assured soundings at over 1500 globally distributed stations. The archive was created by merging data from different sources with varying periods of record and areas of coverage. Although the overall period of record is 1938 to present, the length and completeness of record vary widely among stations, and the vertical resolution, vertical extent, and completeness of soundings generally improve considerably over time. An extensive set of algorithms was developed and carefully evaluated to eliminate duplicates and gross errors. The highest-quality data in IGRA are temperature, geopotential height, and surface pressure at stations with relatively complete records in these variables.

IGRA is available freely to all users in various forms via the Internet. In addition, monthly means of geopotential height, temperature, as well as zonal and meridional wind components at the surface, tropopause, and mandatory levels are provided for the nominal times of 00 and 12 UTC. The data are updated on a daily basis using GTS messages received at NCDC on the previous day. Using the same procedures that were applied to the historical data, the update process restructures the original files into the IGRA format, ensures that soundings and levels are properly sorted, removes duplicate levels and soundings, and employs all applicable quality assurance procedures. Checks that require data for periods of time longer than a few days, such as the runs-in-time check and the check for temporal consistency in temperature, are not applied as part of the daily update process. These algorithms will instead be employed at the beginning of each calendar year.

At the time of the annual updates, we will consider revisions to the quality assurance procedures as well as the addition of historical data from previously unused stations and sources. When adding data, efforts will concentrate on the enhancement of spatial coverage in regions outside the contiguous United States and the former Soviet Union prior to 1963, improvements in the temporal completeness and geographical distribution of stations in Africa, and the completion of time series at Chinese stations. Potential enhancements to the quality assurance process include climatologically-based temporal and vertical consistency checks on geopotential height, an algorithm for the identification of invalid tropopause levels, procedures for detecting unrealistically large wind speeds, and additional checks on dewpoint depression.

Currently, IGRA data have not been adjusted for inhomogeneities resulting from changes in instrumentation or observing practices. However, as part of the Radiosonde Atmospheric Temperature Products for Assessing Climate (RATPAC) project, a collaborative effort between scientists at NCDC, NOAA's Air Resources Laboratory, and Geophysical Fluid Dynamics Laboratory, work is under way to construct homogeneous global and hemispheric temperature time series using a combination of the adjustment methodology developed by Lanzante et al. (2003a,b) and the first difference method (Peterson et al. 1998). The resulting time series will be updated on a monthly basis and used for climate monitoring activities at NCDC. Recognizing the limitations of the first difference method which are particularly noticeable in data-sparse regions (Free et al. 2004; Menne and Williams, submitted to J. Climate), other future work will explore the identification of stations and time periods whose temperature data are sufficiently homogeneous for climate studies, using a combination of station history information and multiple change-point tests.

Acknowledgments

We thank Byron Gleason and Claude Williams for providing helpful and constructive comments on an initial version of the manuscript.

References

Collins, W.G., 2001a: The Operational Complex Quality Control of Radiosonde Heights and Temperatures at the National Centers for Environmental Prediction. Part I: Description of the Method. *J. Appl. Meteor.*, **40**, 137–151.

_______, 2001b: The Operational Complex Quality Control of Radiosonde Heights and Temperatures at the National Centers for Environmental Prediction. Part II: Examples of Error Diagnosis and Correction from Operational Use. *J. Appl. Meteor.*, **40**, 152–168.

Durre, I., and co-authors, 2004: Report on the Workshop to Improve the Usefulness of Operational Radiosonde Data. *Bull. Amer. Meteor. Soc.*, submitted.

Elliott, W.P., and D.J. Gaffen, 1991: On the utility of radiosonde humidity archives for climate studies. *Bull. Amer. Meteor. Soc.*, **72**, 1507-1520.

_______, R.J. Ross, and W. Blackmore, 2002: Recent changes in NWS upper-air observations with emphasis on changes from VIZ to Vaisala radiosondes. *Bull. Amer. Meteor. Soc.*, **83**, 1003-1017.

FCM-H3, cited 2004: Federal Meteorological Handbook No. 3: Rawinsonde and Pibal Observations. [http://www.ofcm.gov/fmh3/text/default.htm .]

Finger, F.G., and F.J. Schmidlin, 1991: Meeting review: Upper-Air Measurements and Instrumentation Workshop. *Bull. Amer. Meteor. Soc.*, **72**, 51-56.

Free, M., I. Durre, E. Aguilar, D. Seidel, T.C. Peterson, R.E. Eskridge, J.K. Luers, D. Parker, M. Gordon, J. Lanzante, S. Klein, J. Christy, S. Schroeder, B. Soden, and L.M. McMillin, 2002: CARDS Workshop on Adjusting Radiosonde Temperature Data for Climate Monitoring: Meeting summary. *Bull. Amer. Meteor. Soc.*, **83**, 891-899.

_______, J. K. Angell, I. Durre, J. R. Lanzante, T. C. Peterson, and D. J. Seidel, 2004: Using first differences to reduce inhomogeneity in radiosonde temperature datasets. *J. Climate*, in press.

Garand, L., C. Grassotti, J. Hall, and G.L. Klein, 1992: On differences in radiosonde humidity-reporting practices and their implications for numerical weather prediction and remote sensing. *Bull. Amer. Meteor. Soc.*, **73**, 1417-1423.

Gaffen, D.J., 1994: Temporal inhomogeneities in radiosonde temperature records*. J. Geophys. Res.*, **99**, 3667-3676.

Gandin, L.S., 1988: Complex quality control of meteorological observations. *Mon. Wea. Rev.*, **116,** 1137–1156.

_______, L.L. Morone, and W.G. Collins, 1993: Two years of operational comprehensive hydrostatic quality control at the National Meteorological Center. *Wea. Forecasting,* **8,** 57-72.

Kahl, J.D., M.C. Serreze, S. Shiotani, S.M. Skony, and R.C. Schnell, 1992: In-situ meteorological sounding archives for Arctic studies. *Bull. Amer. Meteor. Soc.*, **73**, 1824-1830.

Lanzante, J.R., 1996: Resistant, robust and nonparametric techniques for analysis of climate data: Theory and examples, including applications to historical radiosonde station data. *Int. J. Climatol.*, **16**, 1197-1226.

_______, S.A. Klein, and D.J. Seidel, 2003a: Temporal homogenization of monthly radiosonde temperature data. Part I: Methodology. *J. Climate*, **16**, 224-240.

_______, _______, and _______, 2003b: Temporal homogenization of monthly radiosonde temperature data. Part II: Trends, sensitivities, and MSU comparison. *J. Climate*, **16**, 241-262.

Loehrer, S.M., T.A. Edmands, and J.A. Moore, 1996: TOGA COARE Upper-Air Sounding Data Archive: Development and quality control procedures. *Bull. Amer. Meteor. Soc.*, **77**, 2651–2672.

Menne, M.J., and C.N. Williams, Jr., 2004: Detection of undocumented change points: On the use of multiple test statistics and composite reference series. *J. Climate*, submitted.

National Geophysical Data Center (NGDC), Cited 2004: Global Land One-kilometer Base Elevation (GLOBE). [Available online from http://www.ngdc.noaa.gov/seg/topo/globe.shtml .]

National Research Council (NRC) Panel on Reconciling Temperature Observations, 2000: *Reconciling Observations of Global Temperature Change*. National Academy Press, Washington, DC, 85 pp.

Peterson, T.C., and R.S. Vose, 1997: An overview of the Global Historical Climatology Network Temperature Database. *Bull. Amer. Meteor. Soc.*, **78**, 2837-2849.

_______, T. Karl, P. Jamason, R. Knight, and D. Easterling, 1998: First difference method: Maximizing station density for the calculation of long-term global temperature change. *J. Geophys. Res.*, **103**, 25,967-25,974.

Schwartz, B.E., and C.A. Doswell III, 1991: North American rawinsonde observations: Problems, concerns, and a call to action. *Bull. Amer. Meteor. Soc.*, **72**, 1885-1896.

WMO, 1996: *Guide to Meteorological Instruments and Methods of Observation, 6th ed.* WMO Rep. 8, World Meteorological Organization, Geneva, Swit zerland.

WMO, cited 2004: Weather Reporting, Observing Stations. Publ. 9. [Available online from http://www.wmo.ch/ .]

List of Tables

List of Figures

Table 1. Data sources for IGRA. Source names are listed by categories referenced in the text. For each source, IGRA contains the periods of record, number of stations, and percentage of soundings listed.

	Data Source	Period of Record	Area of Coverage	Number of Stations	Percent of IGRA Soundings
Core	NCDC Historical GTS	1963-1970	Global	820	7.94
	NCAR/NCEP GTS	1970-1972	Global	848	3.01
	NCEP GTS	1973-1999	Global	1517	64.06
	NCDC Real-Time GTS	2000-present	Global	1093	7.13
Other Large Scale	Russian GTS	1998-2001	Global	923	1.59
	U.S. Air Force	1946-1973	Global	292	4.49
	Australian GTS	1990-1993	Southern Hemisphere	170	0.15
Country-specific	U.S.	1946-2001	U.S. + U.S. military	150	9.81
	Australian	1938-1989	Australia and its territories	17	1.63
	Argentine	1958-1991	Argentina	8	0.18
	South Korean	1984-1992	South Korea	4	0.01

Table 2. Similarity thresholds used when comparing data from different sources.

Variable	**Threshold**
Geopotential Height	10 m
Temperature	0.2 °C
Dewpoint depression	0.5 °C
Wind Speed	2 m/s
Wind Direction	10°

Table 3. IGRA quality assurance procedures and their impact.

Category	QA Procedure	Items Checked	Items Deleted
Basic Plausibility Checks	Date Check	Year, month, day, hour	Sounding
	Release Time Check	Release time	Release Time
	Observation Value Check	p, z, T, d, ws, wd	Individual values
Internal Consistency Checks	Hypsometric Check	p, z	Individual levels
	Height Sequence Check	z	Levels
	Obs Hour/Release Time Check	Obs Hour – Release Time	Sounding
	Elevation Inspection (manual)	Surface height	Surface height
	Multiple Surface Levels Check	Level type indicator	Levels
	Surface Inspection (manual)	p, T	Level
	Below-surface Level Check	p, z	Level
	Zero-speed Wind check	ws, wd	ws, wd
Checks for Repetition of Values	Temporal runs check (generic)	p, z, T	Levels or values
	Temporal runs check (by hour)	p, z, T	Levels or values
	Vertical runs check	T	values
	Joint vertical runs check	T, d, ws, wd	values
	Frequent erroneous values check	z, T	values
	Fixed Geopotential Height	z (Russian GTS only)	values
Climatological Checks	Tier 1	p, z, T	levels or values
	Tier 2	p, z, T	levels or values
Additional Checks on Temperature	Crazy Profile Check	T	T-soundings
	Generic vertical outlier check	T	values
	Vertical sore-thumb check	T	values
	Temporal sore-thumb check	T	values
Data Completeness Checks	Lone dewpoint depression check	d. T	values
	Lone wind value check	ws, wd	values
	Incomplete level check	p, z, d, T, ws, wd	levels
	Surface-only sounding check	level-type indicator	sounding
	Isolated sounding check	date and time	sounding

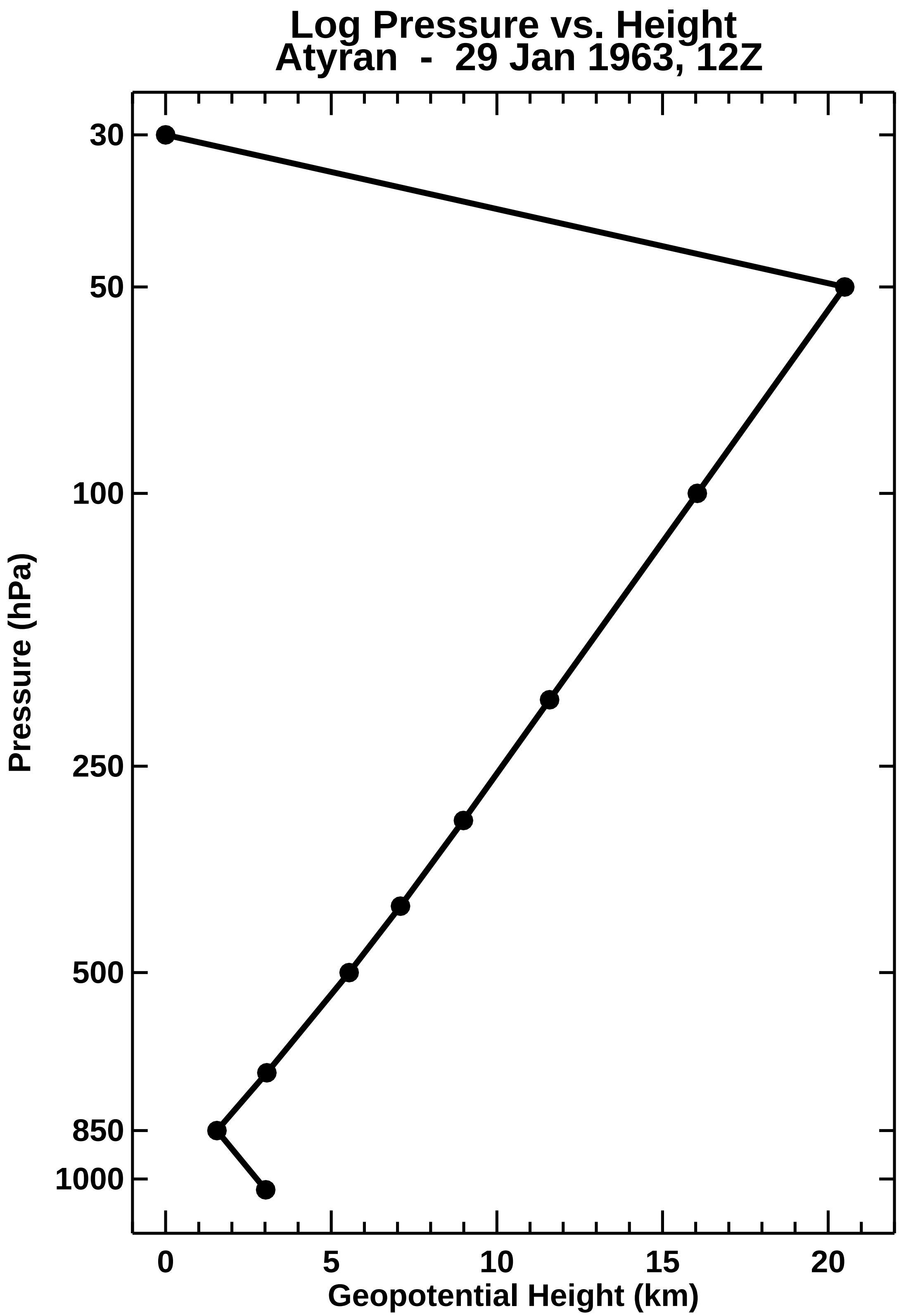

Log Pressure vs. Height
Atyran - 29 Jan 1963, 12Z
Pressure (hPa)
30
50
100
250
500
850
1000
0
5
10
15
20
Geopotential Height (km)

Figure 1

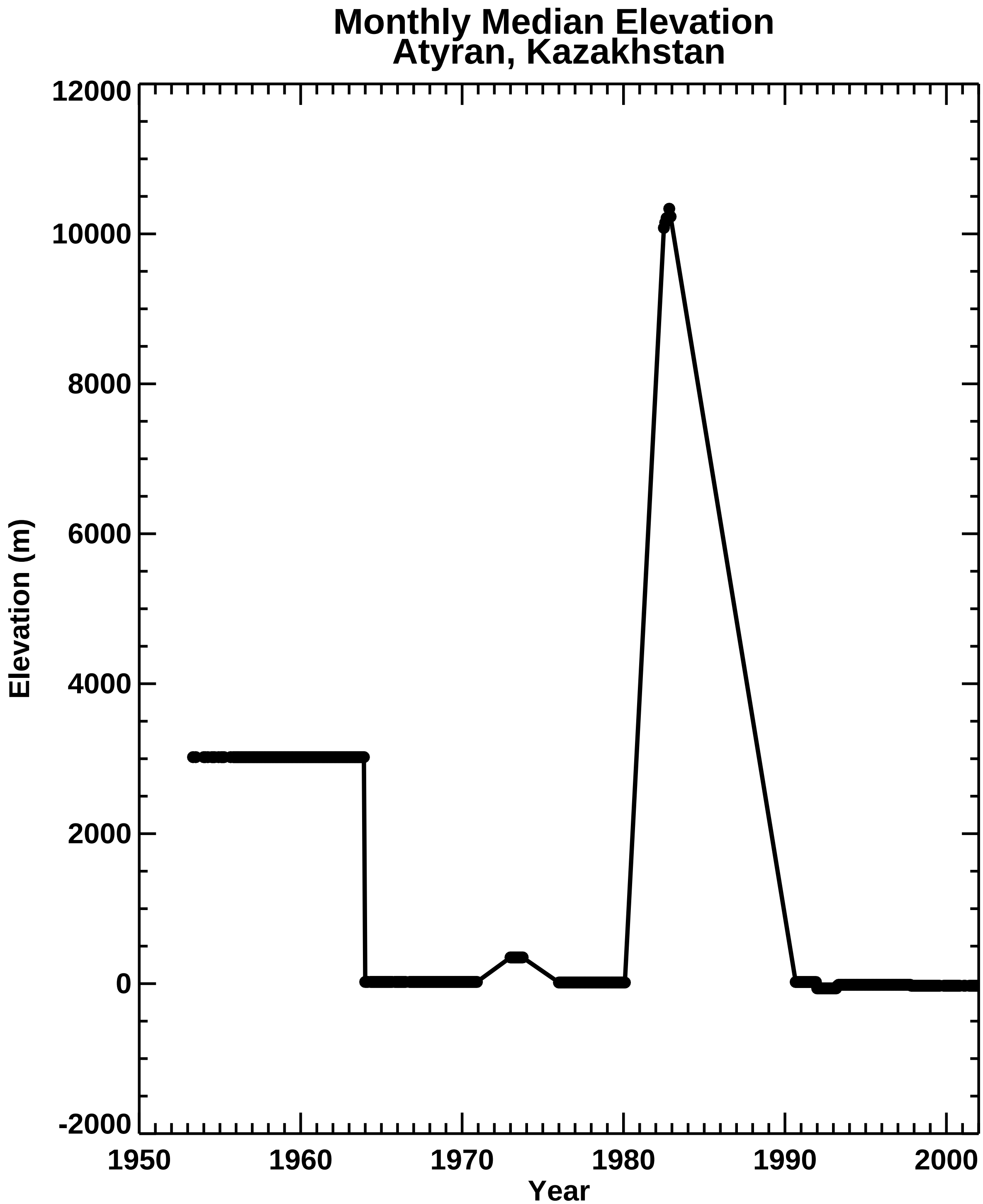

Monthly Median Elevation
Atyran, Kazakhstan
12000
10000
8000
6000
4000
2000
0
-2000
Elevation (m)
1950
1960
1970
1980
1990
2000
Year

Figure 2

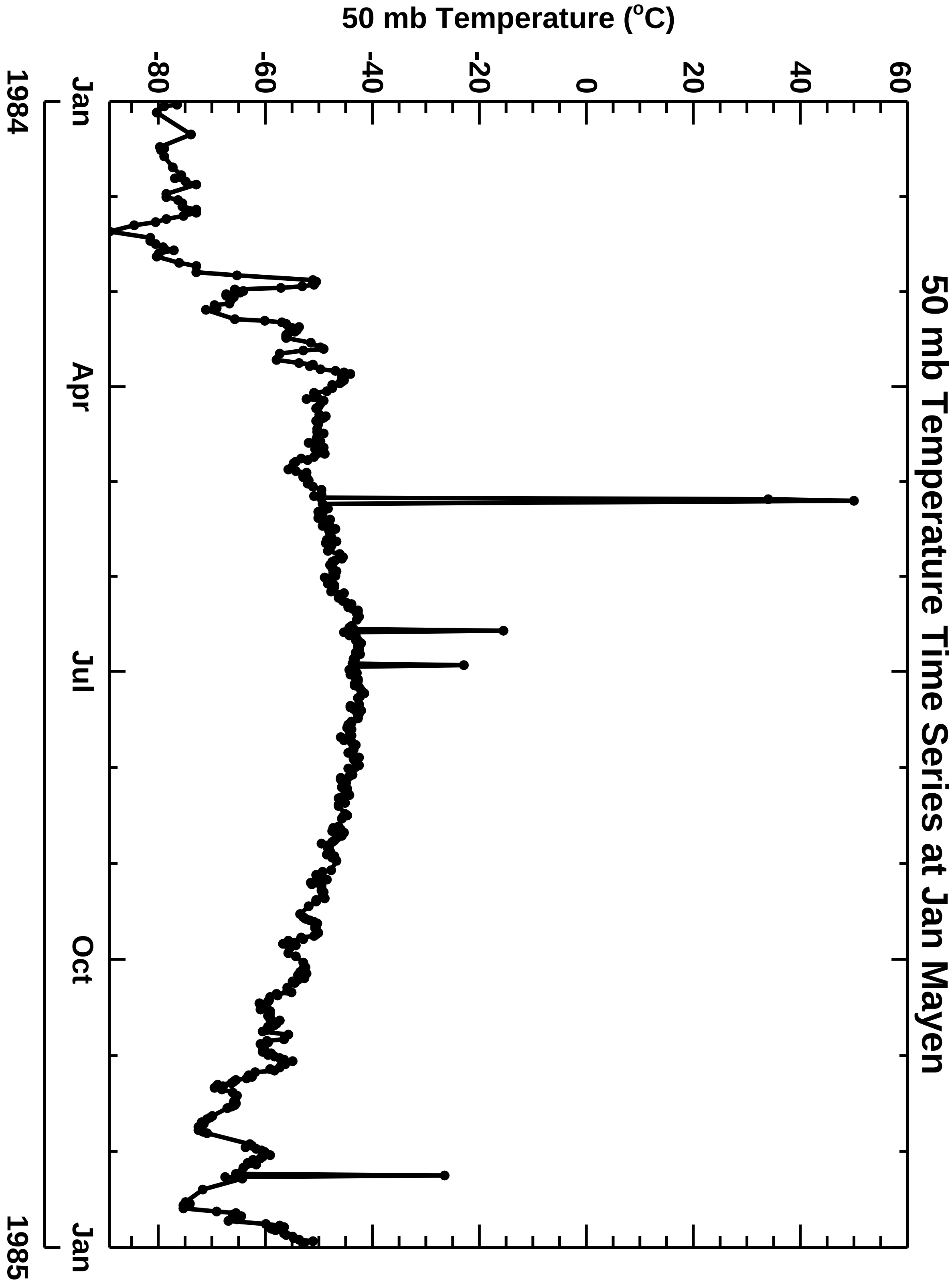

50 mb Temperature Time Series at Jan Mayen
50 mb Temperature (oC)
-80
-60
-40
-20
0
20
40
60
Jan
1984
Apr
Jul
Oct
Jan
1985

Figure 3

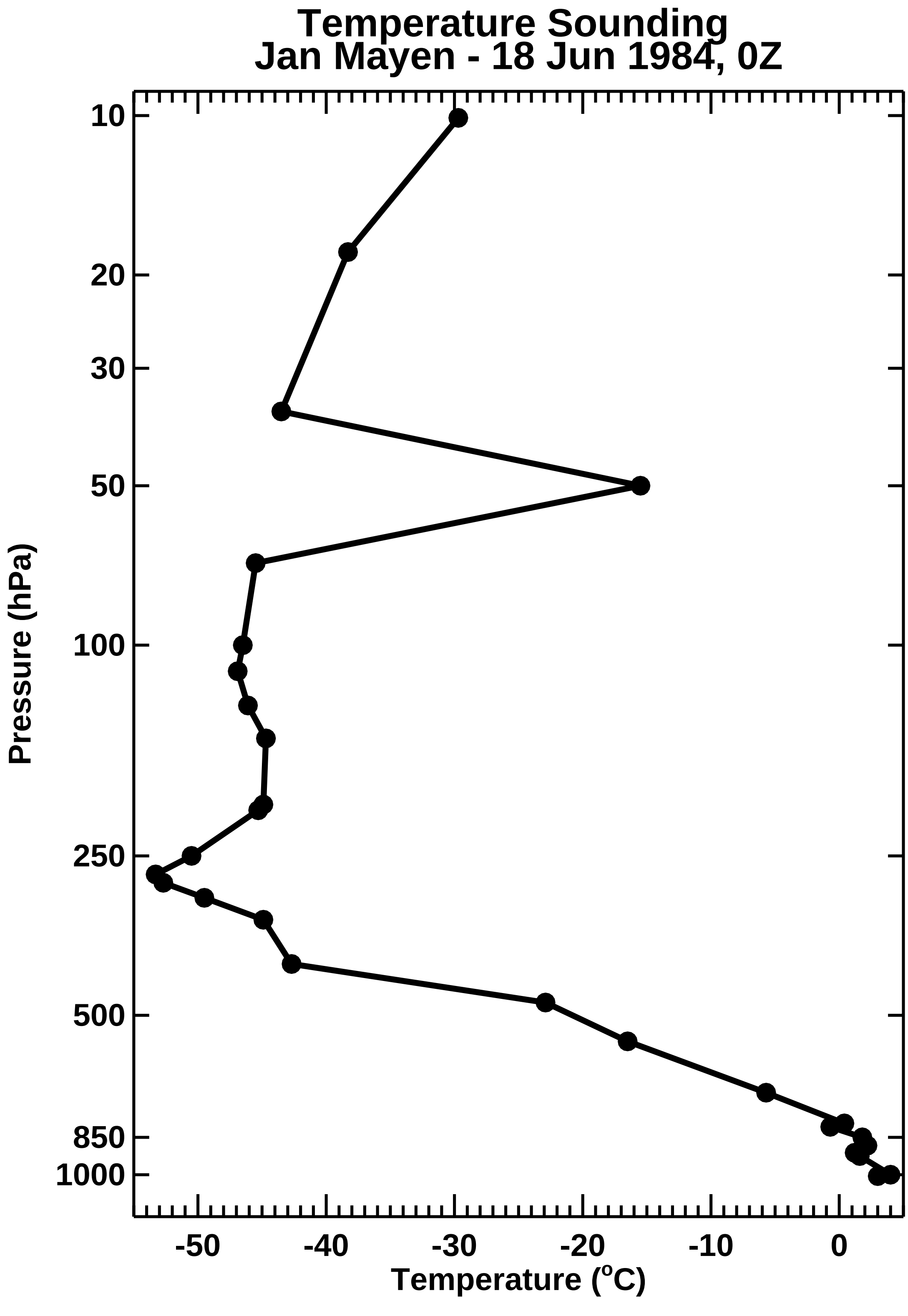

Temperature Sounding
Jan Mayen - 18 Jun 1984, 0Z
10
20
30
50
100
250
500
850
1000
Pressure (hPa)
-50
-40
-30
-20
-10
0
Temperature (°C)

Figure 4

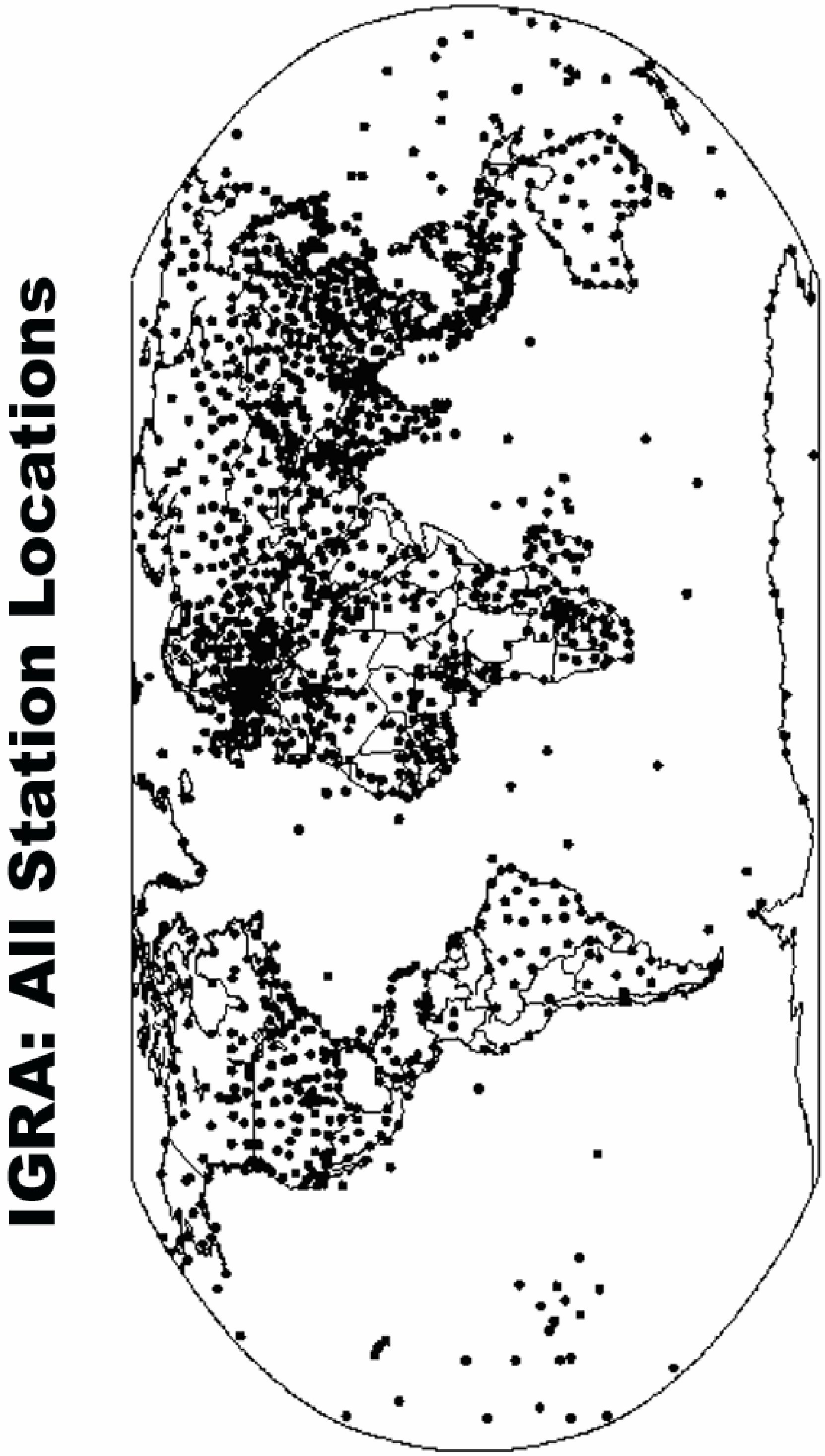


Figure 5a

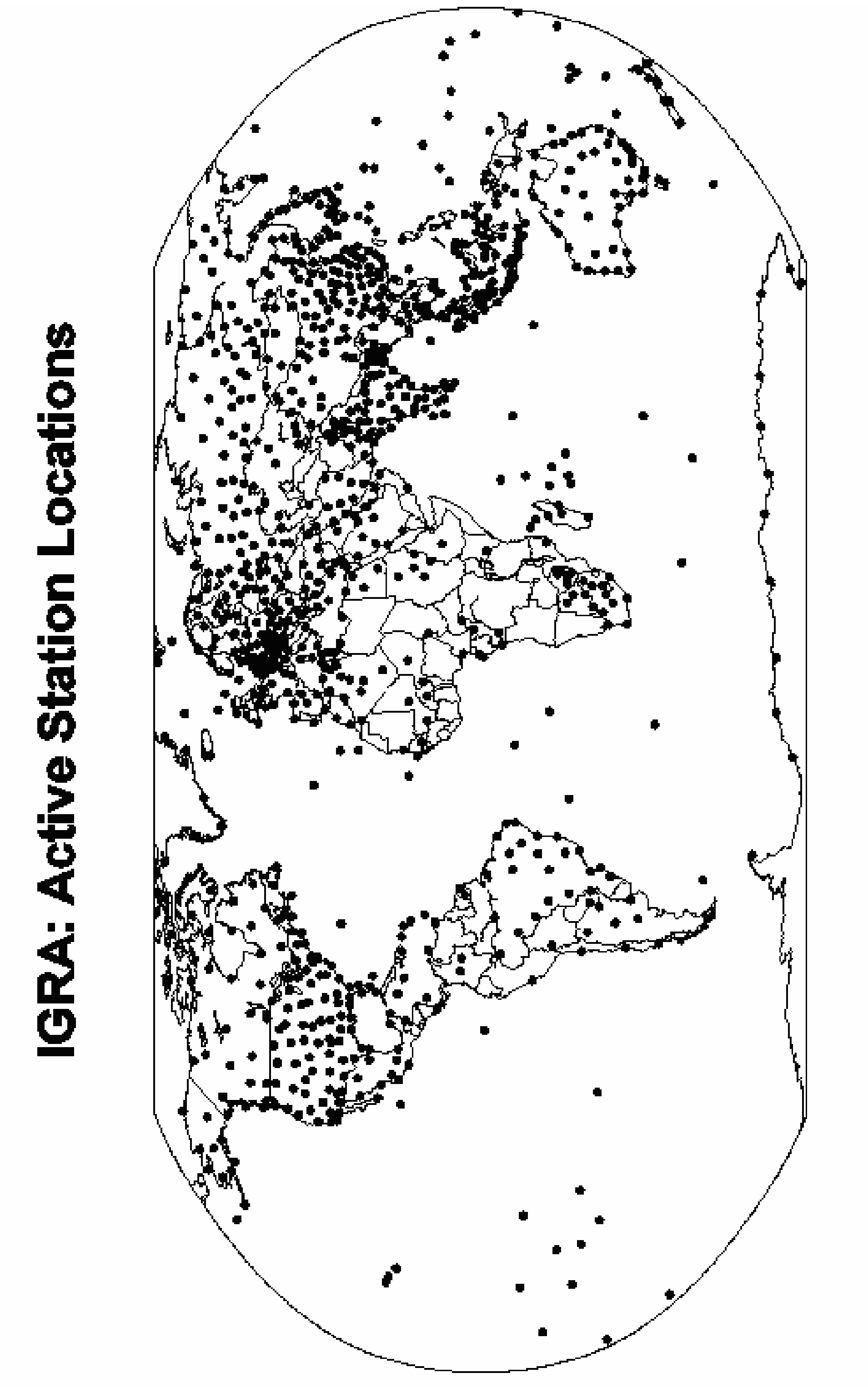


Figure 5b

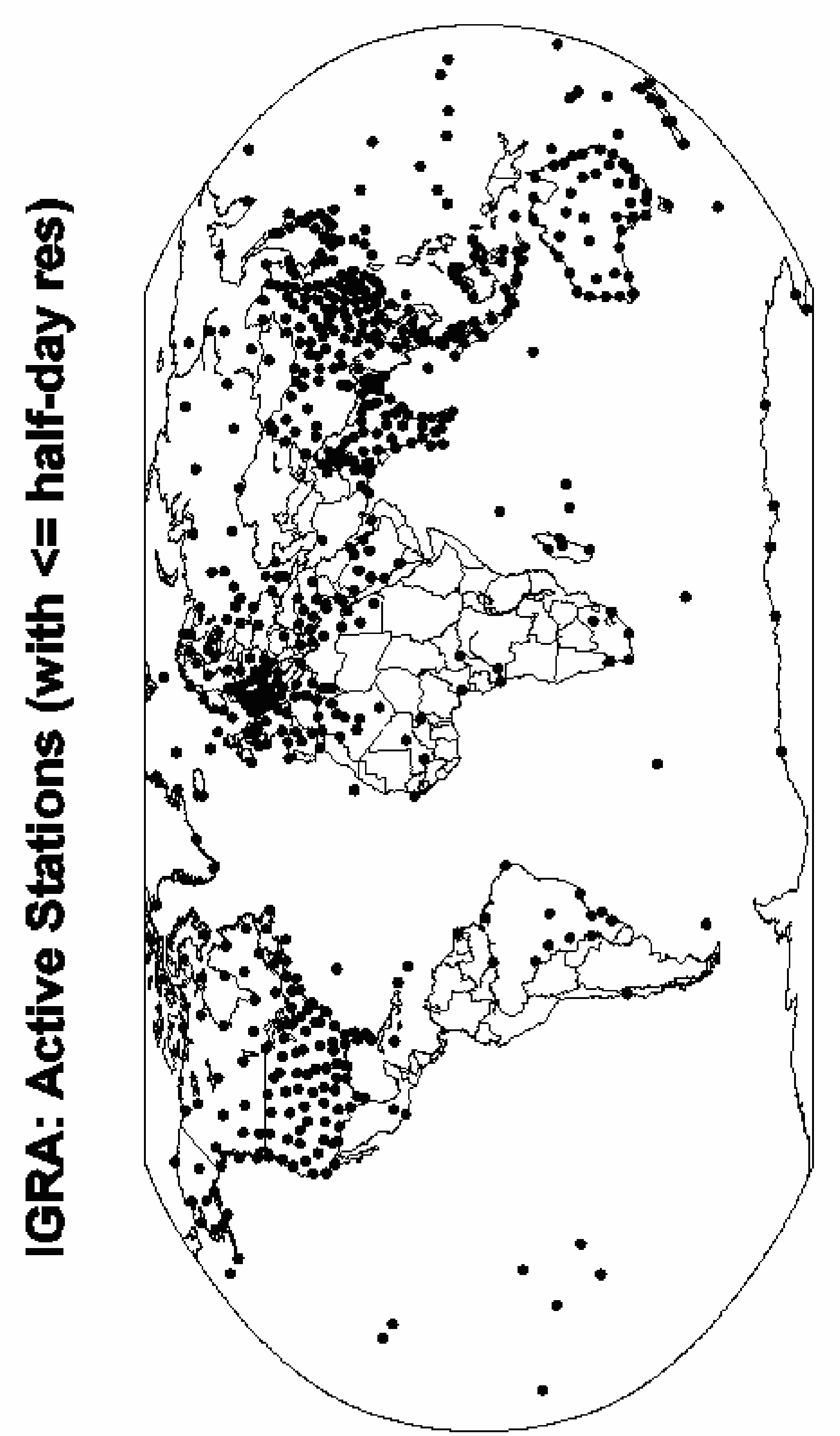
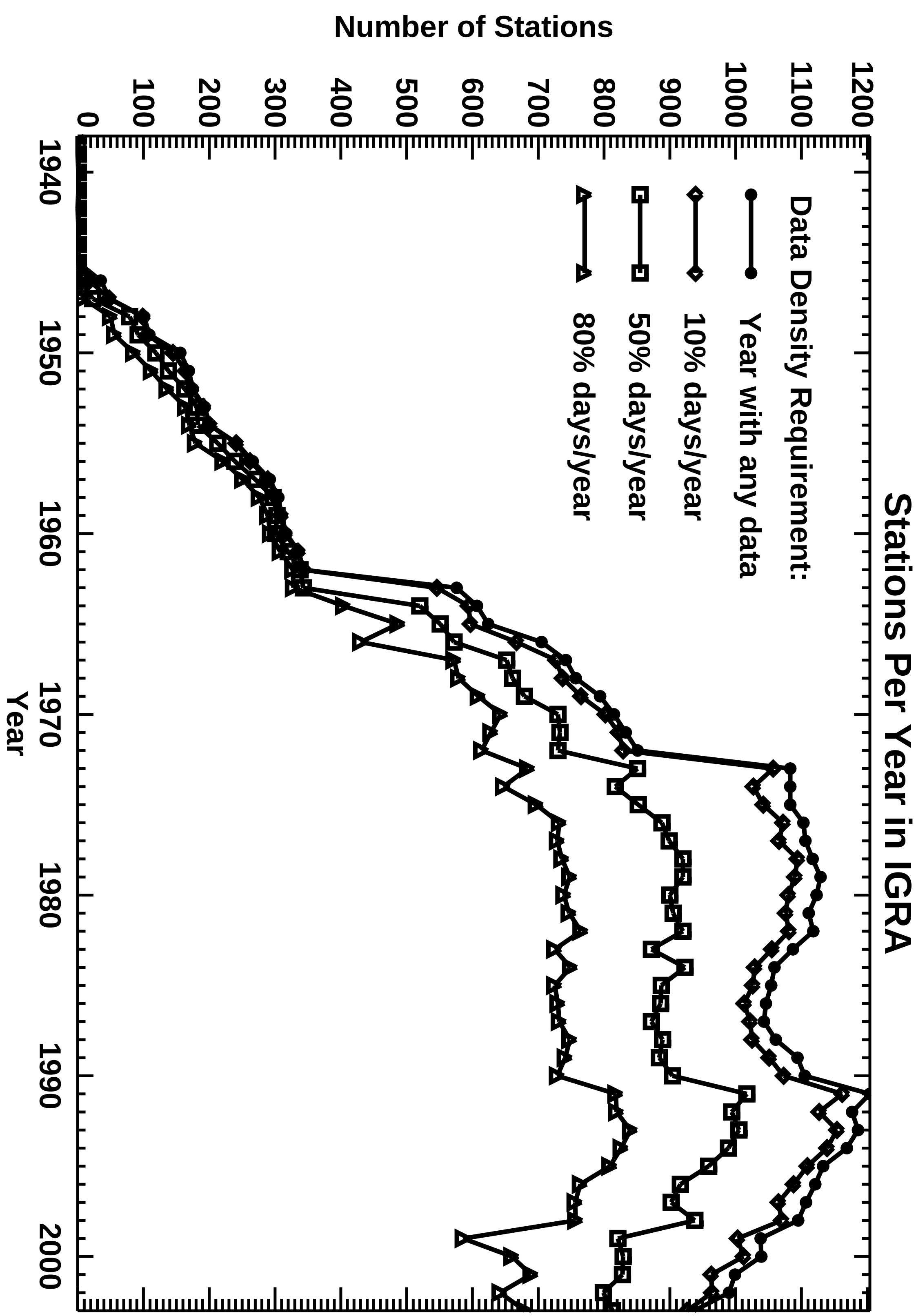
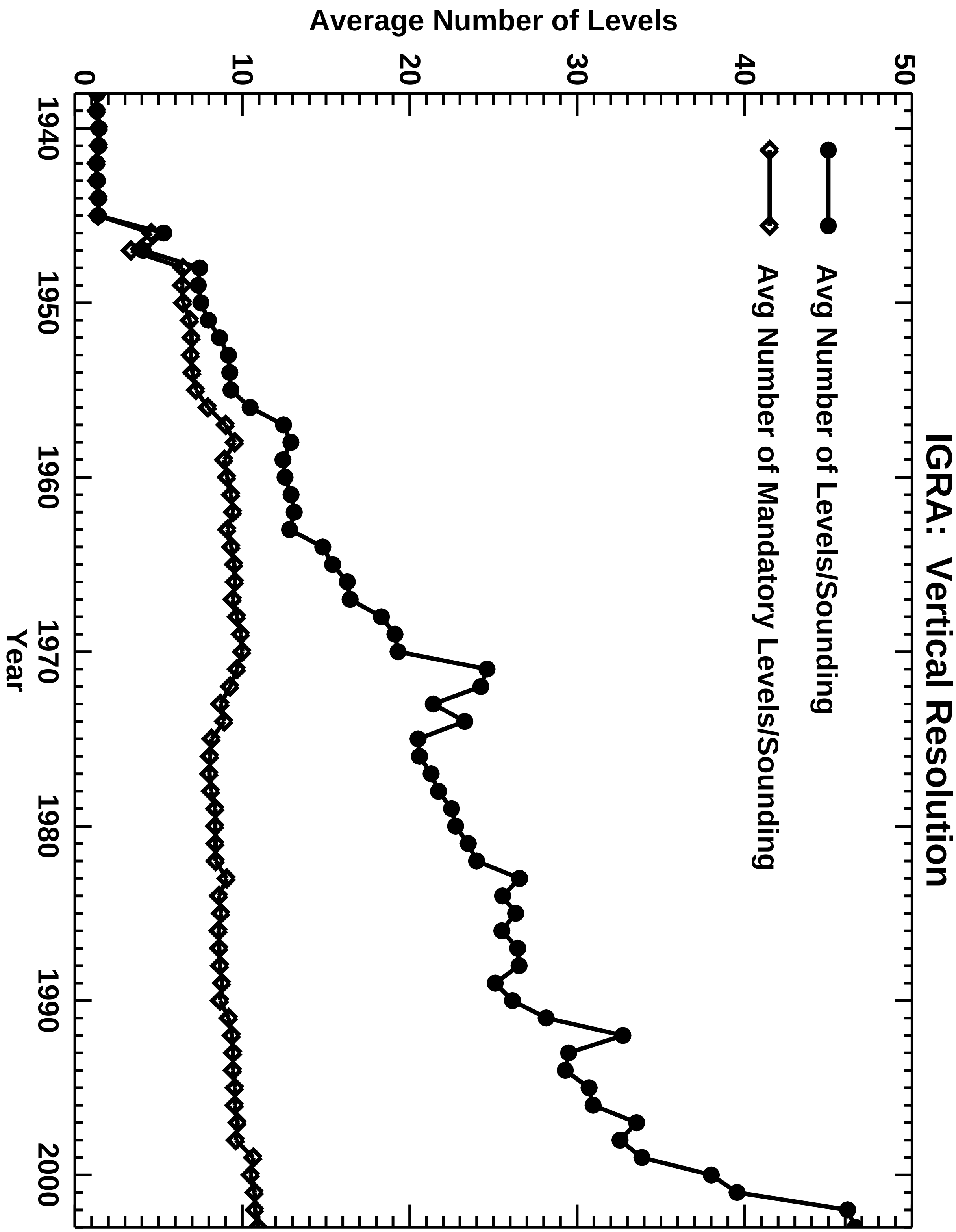


Figure 5c



Stations Per Year in IGRA
Data Density Requirement:
Year with any data
10% days/year
50% days/year
80% days/year
Number of Stations
0
100
200
300
400
500
600
700
800
900
1000
1100
1200
1940
1950
1960
1970
1980
1990
2000
Year

Figure 6



IGRA: Vertical Resolution
Avg Number of Levels/Sounding
Avg Number of Mandatory Levels/Sounding
Average Number of Levels
0
10
20
30
40
50
1940
1950
1960
1970
1980
1990
2000
Year

Figure 7

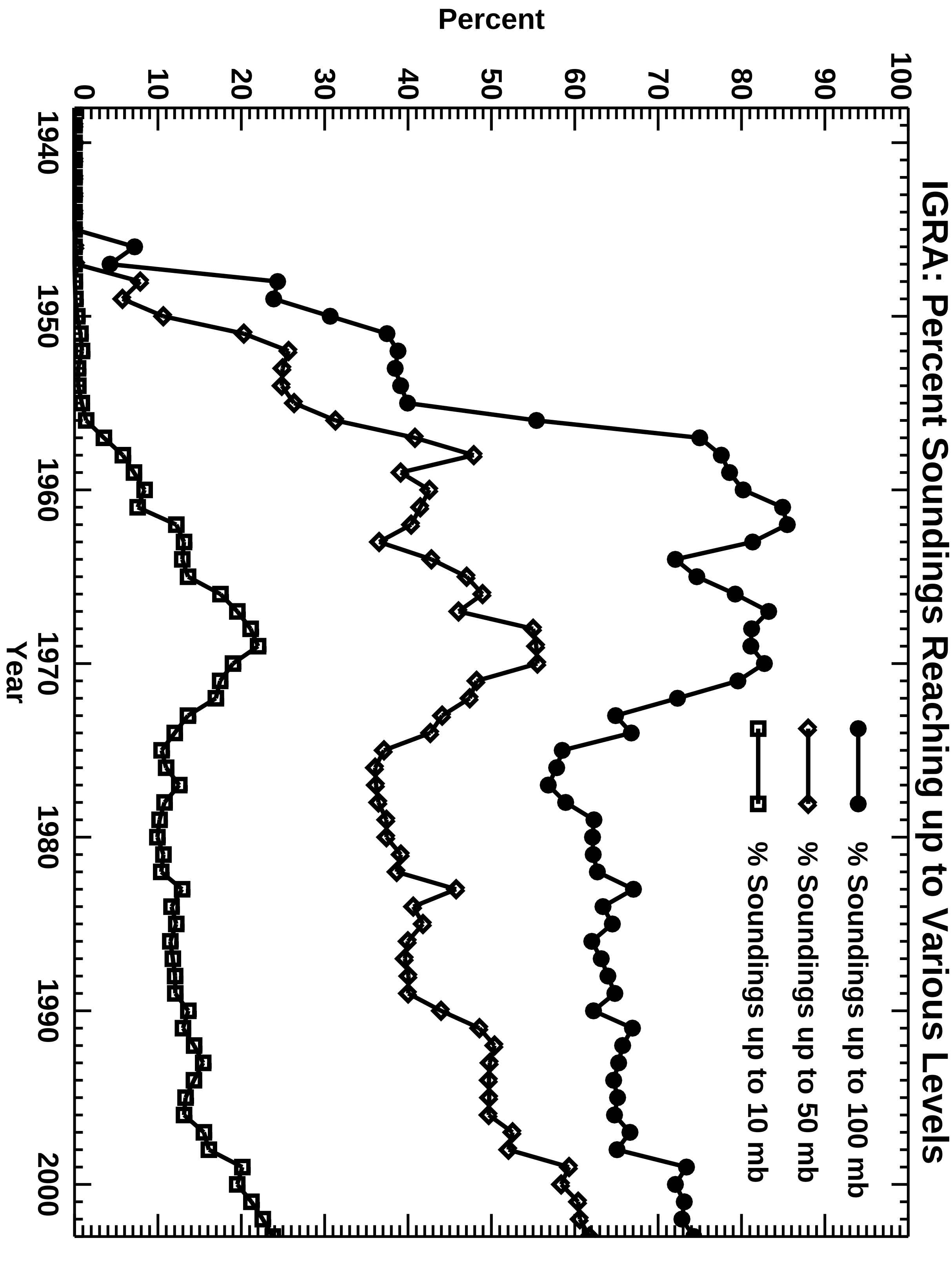
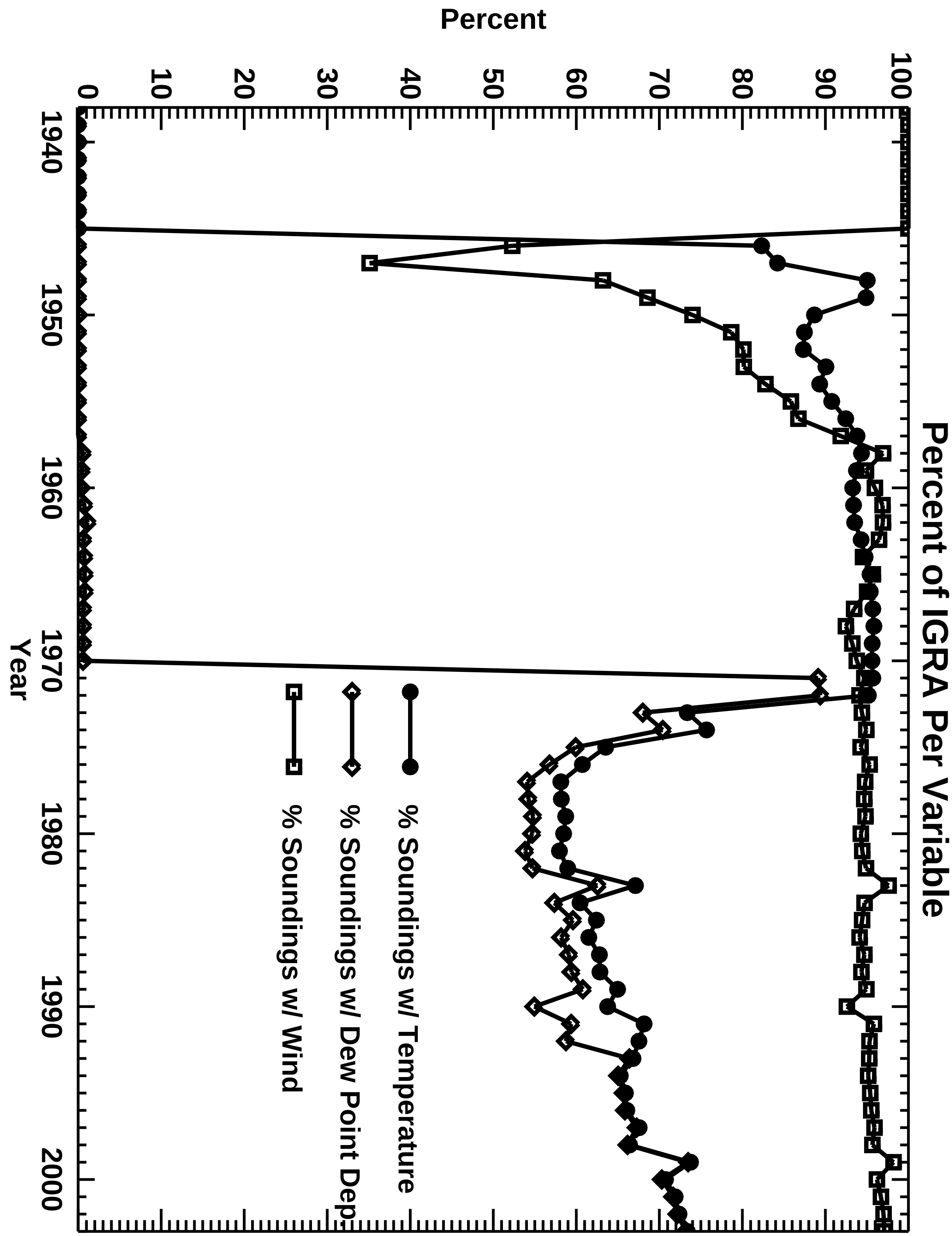

IGRA: Percent Soundings Reaching up to Various Levels
% Soundings up to 100 mb
% Soundings up to 50 mb
% Soundings up to 10 mb
Percent
0
10
20
30
40
50
60
70
80
90
100
1940
1950
1960
1970
1980
1990
2000
Year

Figure 8



Percent of IGRA Per Variable
Percent
Year
0
10
20
30
40
50
60
70
80
90
100
1940
1950
1960
1970
1980
1990
2000
% Soundings w/ Temperature
% Soundings w/ Dew Point Dep.
% Soundings w/ Wind

Figure 9