P1.15 TOWARD DEVELOPING UNBIASED UPPER-AIR TEMPERATURE AND MOISTURE TRENDS FROM HISTORICAL RADIOSONDE DATA: VALIDATING AND COMPLETING RUSSIAN RADIOSONDE HISTORY

Steven R. Schroeder *

Texas A&M University, College Station, Texas

1. INTRODUCTION

Global variations and trends of temperature, moisture, and wind above the surface can potentially be determined from archived radiosonde data back to the late 1950s, and from global high-resolution satellite data since late 1978. However, all climate trends above the surface are suspect, whether computed from radiosonde data or derived from satellite radiances. Modern radiosondes are better protected from radiative effects, and react faster to temperature and humidity changes, than older types. So, climate time series computed from historical radiosonde data have artificial trends, expected to be mainly cooling and drying of unknown magnitude, superimposed on the actual climate trend.

Major problems in using historical radiosonde data are incomplete metadata describing station and instrument histories, and uncertain effects of instrument changes. Goals of this project (and resulting major steps) are as follows:

(1) Develop complete inferred historical station and instrument metadata, building on available metadata.

(2) Develop temperature and dew point adjustments to make each instrument type equivalent to a common "reference instrument" to compensate for biases.

(3) Apply the adjustments to each sounding and develop climatology and time series from the adjusted data.

The steps are basically sequential but are performed repeatedly to refine inferred instruments. An incorrect inferred instrument type at a station is likely to cause a larger discontinuity than in the uncorrected data.

Trends and other statistics computed from the adjusted radiosonde record should be more reliable. While this project focuses on determining global and regional trends and variations of total precipitable water back to 1973 and eventually back to the 1950s, the improved instrument adjustments should also help resolve controversies about atmospheric temperature trends over the last few decades.

This paper describes work to date and some very preliminary results. Section 2 describes the process of developing complete global station and instrument metadata. Section 3 describes the planned method of developing and applying temperature and dew point adjustments to remove instrument-related biases. Section 4 summarizes preliminary global trends of precipitable water since 1973. Section 5 describes the current status of the detailed Russian instrument history.

2. DEVELOPING COMPLETE STATION AND INSTRUMENT METADATA

While available radiosonde station and instrument metadata is extensive, it is quite incomplete and often conflicting. This project combines many metadata sources in one location, attempts to validate the metadata for accuracy, and infers missing or inaccurate metadata.

2.1. Data and metadata sources

Archived radiosonde observations have been obtained from 1973 through (currently) September 2004 from NCAR (National Center for Atmospheric Research) Data Set 353.4. Pre-1973 data and some additional data can be obtained from the Comprehensive Aerological Reference Data Set (CARDS) project at the National Climatic Data Center (NCDC)

While the goal of this effort is to infer complete historical metadata, it is best to start with all available metadata sources. Major global metadata compilations include Gaffen (1993, 1996), the *WMO Catalogue of Radiosondes and Upper-Air Wind Systems in Use by Members* (WMO, 2002 and earlier years), the CARDS online station history file (ftp://ftp.ncdc.noaa.gov/pub/data/cards/long_sonde.lst), and 31313 (instrument code) entries in the soundings. All sources are admittedly incomplete, with some errors and inconsistencies.

A systematic literature search has uncovered many additional sources of information about instrument types, procedures, and experiments. Some sources are *Monthly Weather Review*, *Bulletin of the American Meteorological Society*, other American Meteorological Society journals, conference preprints, *Journal of Geophysical Research*, *Geophysical Research Letters*, *EOS*, *IEEE Transactions on Military Electronics*, *Journal of Scientific Instruments*, *Electronics*, *Journal of Research of the National Bureau of Standards*, WMO publications (especially Instruments and Observing Methods reports) and station and observing ship catalogs, reports of field experiments and intercomparisons, advertisements in journals, manufacturer brochures and web sites, several dozen foreign journals, civilian and military manuals, radiosonde collections (at Texas A&M University, the Smithsonian, and the National Climatic Data Center), *Meteorological and Geoastrophysical Abstracts*, atlases (for station names and locations), and some books.

2.2. Developing a comprehensive list of radiosonde instrument types

To determine what instrument types were or are used at each location, a comprehensive list of radiosonde types is needed. This effort starts with published lists such as

---

* *Corresponding author address:* Steven R. Schroeder, Department of Atmospheric Sciences, Texas A&M University, 3150 TAMU, College Station, TX 77843-3150; e-mail: steves@ariel.met.tamu.edu.

Gaffen (1993), Smith (2002), and the WMO list of instrument codes. The WMO and Smith (2002) lists are undocumented, but this effort has uncovered references to most instruments in those lists.

The scope of the instrument catalog includes all atmospheric in-situ profiling instruments for which references were found, because CARDS should be comprehensive. Besides radiosondes, other categories are nonbroadcasting instruments, dropsondes, rocketsondes, ozonesondes, tethersondes, other specialized radiosondes, and wind-only instruments. So far, 1108 instrument type codes are assigned. Because the literature search is not complete, several hundred more instrument types will probably be added. Some entries are erroneous or are not actually radiosondes, some are doubtful, some are duplicate names for the same instrument (but entries are combined if possible), and some entries are codes to identify a family of instruments where the specific model is unspecified.

Some ambiguities in instrument codes result from the fact that a 2-digit WMO code includes some information about the ground processing unit. However, some metadata relevant for instruments is not in that code, such as whether Vaisala RS80 uses an A-Humicap or H-Humicap. Descriptions are very sketchy with no documentation, codes are not assigned in a systematic order, and few codes remain for future assignment as long as the instrument code must be 2 digits.

Some proliferation of instrument types occurs since important changes are listed separately even if they are not counted as different models. Such changes include the new VIZ carbon hygristor in June 1980, changes in ground calibration or data processing, changes in solar or radiation corrections, and changes in formulas which relate ordinates or other broadcast signals to values of meteorological variables.

2.3. Validating station elevations

Each observation contains a location and elevation from an operational catalog. Before February 1995, the catalog was infrequently updated, and even now, updates occur only after the actual station change. By computing the surface elevation from the first above-surface height, elevation errors of 300 meters or more have been found, with many erroneous elevations between 20 January 1976 and 4 February 1980, and between 4 June 1986 and 30 June 1989. Elevation changes of 5 meters can often be detected to the exact observation, and changes of 1 meter can usually be detected to within a month. An inconsistent error usually indicates that the surface observation is missing. Alduchov and Eskridge (2002) performed a similar analysis, but based on CARDS metadata files, it appears they only checked a small subset of all stations. Operational agencies compute persistent height errors relative to model analyses, and it is possible that some of these are actually surface elevation errors.

A complete elevation history has been prepared for each ship in this data base. While the operational archive shows a station elevation of 0 meters for ships, the launch elevation computed hydrostatically is typically from 5 to nearly 30 meters.

It is not possible to validate horizontal moves directly, or elevations at wind-only stations. However, many erroneous locations can be corrected, such as a wrong latitude or longitude sign, a reported location which differs from an atlas location, or a transient discontinuity in the path of a ship or Arctic ice island.

As an example of a detectable error, station 10384 (Berlin Templehof Airport) was reported to be in the Southern Hemisphere from September 1975 until January 1976. A probable error is the location of station 99877 in a very remote part of Afghanistan (which started reporting 30 June 2004, and is not in a WMO catalog). The claimed elevation is 673 meters, but that region has an elevation of 2500 to 3000 meters, and the surface elevation computes to an average of 303 meters. The correct station probably is Termiz, Uzbekistan, which is almost exactly 3° north of the stated location, with an elevation close to 300 meters (Termiz, or Termez, is also station 38927).

2.4. Inferring instrument types and transitions

This project validates reported instrument types and infers unreported instruments and changes. The consistency of hypothesized instrument signals can be best verified by cross-checking all stations. The same procedures can be applied to both land and ship stations. Most ships travel into different climate regions, but this has not made it difficult to identify instrument transitions.

The procedures used here to determine instrument transitions are most effective when some metadata is available for at least some stations in a country. Even if the metadata is incomplete and of suspected quality or timeliness, it provides useful clues that can help narrow down the possible radiosondes.

The basic method to develop a complete metadata history involves two processes, performed repeatedly. First, the observations and derived statistics at stations which appear well-documented are examined for common characteristics of instrument types. Second, similar instrument signatures are sought at stations or in time periods which are less well documented.

When seeking consistent signals of each instrument type, an instrument model should have smoothly-varying differences between stations, as expected by climatology. Three types of variables should be examined:

(1) Temperature-related variables, especially at high altitudes and differences between day and night. If solar and radiation corrections are effective, these variables may not distinguish instrument types because of large real stratospheric variations such as transient volcanic warming and the gradual cooling from ozone depletion.

(2) Moisture-related variables at all levels, such as the average dew point depression, largest dew point depression, and lowest and highest relative humidity. Moisture variables are often very effective in distinguishing radiosonde types. "Dew point censoring" as practiced in the United States from 1 April to 1973 to 30 September 1993 (continuing at a few military stations) can be detected by at least one dew point depression of 30° C, no other relative humidities under 20 percent, and no dew point depressions above 30° C.

(3) Statistics of sounding "quality" such as average

number of temperature and dew point levels, lowest pressure of the sounding, lowest pressure with wind reported, number of wind levels, and lowest pressure and temperature with dew point reported. These may indicate an instrument change, or a change in radar, ground processing equipment, software, or operating procedures. A discontinuity in temperature or moisture may occur even if the radiosonde does not change. For example, starting to report dew point regardless of temperature instead of only to a temperature around -40° may cause apparent upper-tropospheric moistening because the colder cases are no longer excluded.

Caution must be used in inferring that a change in a statistical indicator is a change in an instrument. If such a change is an administrative policy, it is likely to be adopted the same day or in a short time by all (or at least many) stations in a country, and should affect all observations after the change is adopted. If a station gradually applies a practice to more and more observations (such as reporting dew points only to a temperature around -40°), it is likely to be introducing a new instrument gradually. If the new practice applies to all soundings starting a certain date, but this date varies considerably from one station to another, that may indicate that each station introduces a new radiosonde after using up all of the old ones. If the dew point ends at random points in different soundings (not consistently to a temperature just above -40° or some other threshold), the station may simply have a defective batch of humidity sensors. Even for the same instrument, the reporting practice may differ between countries, or between agencies such as civilian versus military.

2.5. Developing a consolidated metadata file

The main output of this portion of the project is a single text file called “station.master” which is designed to be both human-readable and suitable for computer processing. This is accomplished by placing formatted data at the beginning of lines. Comments either follow formatted data or are on a separate line. A line can be a comment line simply by violating the format (any nonnumeric character in a numeric field), or by starting with a semicolon, or by being blank. Comments can include the station name, operating agency, a reference or metadata source, data characteristics, reasons for certainty or uncertainty of the metadata, or anything else that helps explain what is observed in the data.

Having formatted metadata, comments, and references in the same file reduces duplication in maintaining the metadata. Even for people who do not plan to use radiosonde data directly in their research, this file should still be useful as documentation of station history and instrument or location changes.

The major portion of the metadata file lists all stations with their location and instrument type, with the date and time when that combination of parameters begins. A new data line is used for any change. Station metadata includes 3 lists, with the following formatted data:

(1) Land stations in numerical order, listing each station with its 5-digit ID, latitude, longitude, surface elevation, instrument type, and starting date and time of this combination of parameters. Temporary field experiment stations are assigned 5-digit IDs, and one “land” station is actually the Ekofisk oil platform (station 01400).

(2) Fixed ship stations in alphabetical order by ship station (for example, the “Ship M” location reported using station IDs 4YM, then C7M, and now LDWR), listing each station with its ID, nominal latitude and longitude, elevation from which the radiosonde is launched, instrument type, and starting date and time of this combination of parameters. In the archived data, a ship station reports an elevation of 0 meters, but here the computed launch elevation is listed. Because a fixed ship can drift slightly, individual observations specify the actual latitude and longitude, not the nominal location which is in this metadata file.

(3) Moving ship (or ice island) stations in alphabetical order by reported station ID, listing each station with its ID (4 to 6 characters), computed launch elevation, instrument type, and starting date and time of this combination of parameters. Moving ships report their location with each observation, so this part of the file does not specify latitudes and longitudes.

The metadata file contains some optional sections which are used to correct errors that are specific to this data set. Operational processing accepts all station IDs that are not a 5-digit number because there is no consolidated list of alphanumeric identifiers, and accepts all 5-digit station IDs which are in surface or upper air station catalogs (For example, there are over 100 observations for station 28552 which are mostly fragments of reports from station 28952). Over 3000 station IDs are simply typing or communication errors. Optional metadata sections rename or reject certain observations or station IDs. One section lists Arctic ice islands and their approximate locations for each month. Observations given other names (often SHIP) near these locations are actually Arctic ice island data. A final metadata section lists land stations reported by alphanumeric IDs which are assigned 5-digit station IDs here. Most of these are temporary field sites, so reporting the station as a ship allows reporting the latitude, longitude, and elevation

Many WMO station names are out of date. For example, many name changes of India cities since the 1970s (Madras is now Chennai, Bombay is now Mumbai, Poona is now Pune, and Calcutta is now Kolkata) are not in the catalogs yet.

3. DEVELOPING ADJUSTMENTS FOR INSTRUMENT TYPES

After substantially complete metadata is developed for all stations, it is possible to systematically search for consistent differences between instrument types and apply adjustments to statistically correct for instrument biases. Steps below are not applied strictly in sequence. For example, preprocessing of archived data was done before preparing the metadata file above, and is repeated and refined as the metadata file is developed.

3.1. Processing of archived observations

All data since 1973 is reprocessed almost every month to improve the evaluation of observations and collection of statistics. The program writes files of processed observations in a readable (and also computer-readable) format. This file has no comments, but has indicators of the quality of the sounding and computed statistics such as precipitable water (total and in layers), sea level pressure, free-air lapse rate, and hydrostatically-computed surface elevation and heights.

A separate statistical file lists each station ID, the starting date and time of each new location or elevation, and the number of observations and other statistics for each year. This file reveals many problems with the archived data, especially erroneous station IDs.

Each sounding is evaluated using over 30 tests. Many errors reject a sounding, and others only cause rejection of that data value. Errors which might be corrected include temperature sign errors (if reversing a sign makes the temperature realistic for the level and does not create an unrealistic inversion or superadiabatic cooling) and height errors of 500 or 1000 meters (quite often, a 1000-mb height of 0 meters is reported as 500 meters).

Reported 1000-mb heights at high altitudes in the Andes, Antarctic, and Himalayas are often unrealistically high, indicating that the assumed subsurface air column is too cold. Here, subsurface heights and the sea level pressure are computed by projecting the free-air temperature (above the surface inversion, if any) downward using the free-air lapse rate. This gives much more realistic sea level pressures, especially in the Andes where reported heights often imply a sea level pressure above 1040 mb.

Validation assigns a "quality indicator" to each sounding, which, if a letter, indicates a reason (actually, the first reason found) for rejecting a sounding. For example, "Z" indicates a sounding with no heights reported, "U" means there is a substantial superadiabatic layer, "c" indicates an 850-mb height that is too low or high, or "P" indicates that the reported surface elevation is not hydrostatically realistic. Rejected soundings are still processed as much as possible to see if the information can be rescued by, for example, correcting the surface elevation. For soundings which are not rejected, the quality indicator is a punctuation symbol, and indicates a minor defect of the sounding ("_" indicates that no defects are found). For example, ";" indicates that dew points are not reported if the temperature is below about -40°.

3.2. Inferring instrument types and identifying transitions

The most labor-intensive phase of this project is to examine the data record at each station to validate any existing metadata, to determine consistent characteristics of each instrument type, and to infer a complete instrument history where documentation is missing or inaccurate. Other researchers generally do not try to generate a complete history of specific instrument types, but simply determine that a discontinuity is presumably instrument-caused (Lanzante et al. 2003)

The *first* step is to prepare time series and monthly values (counts, averages, and some extremes) of data variables for each station from preprocessed observations produced in the preceding step. A "station" can be a sequence of station IDs where one station replaces another (This is subjectively determined. Some station replacements, such as in the United States in the late 1990s, are too far apart to be considered homogeneous). Including stations with sparse data, ships, wind-only stations, and stations which are probably erroneous, there are 2378 stations or sequences from 1973 to August 2004. There are 3 files produced for each station or sequence:

(1) A file with one line of statistics (data elements listed below) for each observation.

(2) A file containing monthly averages or totals based on these statistics.

(3) A similar monthly file that uses observations within 3 hours of 0000 and 1200Z to compute statistics. The reason for this file is that soundings around 0600 or 1800Z are often less detailed (for example, terminating at 100 mb or containing only mandatory levels) than those at 0000 and 1200Z, so starting or stopping 0600 or 1800Z observations may produce spurious changes in statistics such as the average number of levels per sounding.

Data elements listed for each observation are station ID, reported latitude, reported longitude, reported elevation, archived and reported instrument type, quality indicator, computed elevation, computed sea level pressure, surface pressure, pressure at top of sounding, lowest pressure with dew point reported, number of temperature levels (*), number of significant temperature levels (*), number of dew point levels (*), coldest temperature in the sounding, coldest temperature with dew point reported, coldest dew point, largest dew point depression (*), lowest relative humidity (*), number of levels with dew point and average dew point depression in 3 layers (800 to 600, 600 to 400, and 400 to 200 mb), number of wind levels by pressure (*), number of significant wind levels (*), lowest pressure with wind reported, and highest height with wind reported. Data elements with (*) are computed only from the surface to 100 mb, and some minor indicators not listed here are also included.

The *second* step is to examine available metadata, these files, and the processed observations, to infer instrument types and transitions. Automated methods to identify discontinuities are unsuccessful. Stations are examined in groups according to likely instrument histories. Histories for Russian and Indian instruments are nearly completed. First, the time series at stations which appear to be well-documented are examined for consistent signals of reported instruments and discontinuities coinciding with reported transitions. Then, time series at stations with little or no documentation are examined for similar signatures of instrument types and discontinuities. These steps are repeated many times, especially as variables are identified that seem to be consistent indicators of instrument types, to refine inferences of instrument types and transition dates. Differences between nearby stations may also help identify instrument transitions.

The inferred instruments and any other station changes are entered in the "station.master" file as the history is developed, along with any comments such as the apparent consistency of signals for the stated instrument type.

3.3. Designating a "reference" instrument

The basic adjustment philosophy here is to make "absolute" adjustments to a chosen "reference instrument," rather than "relative" adjustments from an earlier period to the latest period. This means that stations which have had no instrument changes are still adjusted to be statistically equivalent to the reference instrument. For example, most Russian stations still use older types of instruments, with goldbeater's skin hygrometers, and adjustments make dew points drier even at stations which have not yet started using newer instrument types.

An ideal reference instrument is correct and unbiased, but no such operational instrument exists. Here, the reference instrument is defined as the average of certain VIZ and Vaisala models, specifically VIZ models 1190 and higher through VIZ B (with a carbon hygristor, but excluding models where the hygristor was excessively heated in sunlight) and Vaisala RS21 and RS80 (except for a variety of RS21 used from the late 1970s to mid-1980s which was excessively dry). These models were widely used from the 1970s to about 2000, so most other instruments can be directly compared with one or more of the reference models.

VIZ tends to be wetter than Vaisala. Recent papers indicate a growing consensus that VIZ had a moist bias and Vaisala had (and may still have) a dry bias. So, the average of the two instruments may be nearly correct.

It would appear to be desirable to average the latest models (such as VIZ/Sippican B2 and Microsonde II, and Vaisala RS90 and RS92) as the reference, but those models are not yet as widely used as the models mentioned above, so fewer instruments are able to be directly compared with such a reference model. However, those models may be used as the reference eventually. In any case, while current models can provide much more detailed atmospheric profiles than earlier instruments, the lost information from smoothed profiles cannot be restored regardless of the chosen reference instruments.

Even the reference models are inhomogeneous because of changes in sensors or their exposure. VIZ used a new carbon hygristor starting June 1980 and the case enclosing the hygristor changed several times. For Vaisala, RS80 is much smaller than RS21, the radiation corrections changed several times, and some stations used the H-Humicap humidity sensor, which initially produced artificially dry readings due to contamination in storage. These differences will need to be investigated further and initial corrections to make the reference instruments homogeneous may be needed.

3.4. Developing temperature adjustments

A goal of this project is to adjust the temperatures and dew points in all archived observations to be statistically unbiased with respect to the reference standard. Other researchers use station-specific corrections to develop homogeneous radiosonde data. That is a legitimate approach because the climate environment is different at each station. Here, the basic principle is to develop adjustments for each instrument type, with a scheme that accounts for differences in environment. Both approaches use the data itself, before and after a transition or otherwise comparing periods using different instruments, to determine adjustments.

While several hundred of the more than 1000 instrument models are expected to be identified or inferred to be used, probably there will be about 50 to 60 distinct types with different temperature or humidity characteristics since 1973, because many of the different models do not have changes to the sensors.

To compare any pair of instrument types, 3 types of comparisons can be made using archived data:

(1) A transition from one type to another at a station, in either order, possibly with a brief gap (or temporary use of another type) between these instruments. The time period of comparison for a station should include an integer number of years (up to 3 years) with each instrument type.

(2) Simultaneous use of the instrument types at nearby stations. The stations should be in very similar environments, such as Berlin and East Germany. The time period of comparison should be as long as possible, but preferably an integer number of years, even if a real climate change is suspected in that period.

(3) Frequent alternations of the instrument types at the same station. Again, the comparison should be as long as possible. Caution must be used if one type is used in daytime and the other is used at night.

A fourth type of comparison, a formal intercomparison with different instruments on the same balloon, is not considered because such comparisons involve a very small number of radiosonde launches.

For each instrument type, it is ideal if stations are found with transitions or other comparisons directly to or from a reference instrument. Some instruments may not have a direct transition to a reference instrument (for example, "Type A" to "Type B" to a reference model). Before 1973, some instruments may have a longer "chain" of transitions or comparisons to a reference. Some types, such as (possibly) Indian or Japanese instruments, may have no direct "chain" of transitions to or from a reference. Poor-quality (long-distance) comparisons may be required in such cases. The chain of transitions should be as short as possible because the statistical uncertainty of the adjustment rises as the number of adjustments applied increases. (Adjustments applied by other researchers such as in Lanzante et al. (2003) have a similar problem, because each preceding time segment is adjusted to be statistically equivalent to the latest time segment.)

Here, the first step is to develop differences between reference instrument models, and the second step is to develop differences between each instrument type and a reference instrument. The inferred metadata must be substantially complete at this point so all stations with appropriate instrument comparisons can be included. If only stations with the most obvious instrument discontinuities are included, the adjustments will be too large.

Because an instrument type is assigned to each sounding, a single list of stations and time periods can be used to generate comparisons for each instrument pair. Instrument pairs should include VIZ models in the reference, Vaisala models in the reference, comparisons of VIZ and Vaisala, and each instrument pair in each chain of comparisons to a reference model. If it turns out that the instruments in any pair are homogeneous, those instruments can be consolidated.

For each instrument pair, a computer program would use this list and, for each station and time period (and for all stations), would accumulate statistics considering only observations with each specified instrument type, and would ignore soundings not assigned the specified type (or observations where the instrument type is uncertain).

Temperature adjustments are prepared first. The specific statistics to accumulate for each instrument type in a pair are the cumulative distributions of temperatures in specified pressure layers (usually, the surface is not adjusted because surface data should be obtained from permanently-installed instruments) and intervals of sun angle above the horizon. Such statistics should be accumulated for each station and for all stations as a group. Because each station (or adjacent station pair) should have a similar climate environment, if the cumulative temperature distributions are quite similar for both instrument types (with small, unsystematic differences and little difference in the means), then the instrument types should be considered homogeneous, with no adjustments made.

A temperature adjustment (if needed) is the amount to add to the archived temperature, to make the probability distribution the same as for the instrument type adjusted to, within the same interval of pressure and sun angle. Because the reference instruments are probably well-protected from radiative errors, most adjustments will probably be negative (causing the readings to become cooler). Adjustments are expected to be larger in the stratosphere than nearer the surface, and probably will be small at night.

Even VIZ and Vaisala instrument models in the reference are adjusted. A VIZ observation is adjusted by adding half of the difference from VIZ to Vaisala, and a Vaisala observation is adjusted by adding half of the difference from Vaisala to VIZ. It is possible that both the VIZ and Vaisala series are inhomogeneous enough that small adjustments may be needed to correct some VIZ models (possibly before and after the hygristor change in 1980) to a "VIZ average" and to correct some Vaisala models to a "Vaisala average" before defining the characteristics of the reference instrument. If any VIZ or Vaisala model requires a large adjustment, then that model should be excluded from the reference.

3.5. Developing dew point adjustments

When adjusting temperatures in the step above, dew point depressions are not changed. This changes the relative humidity slightly. For example, with a negative temperature correction (cooling), the relative humidity with a constant dew point depression decreases. However, a major difference between inhomogeneous instrument types is in the response of different humidity sensors. So, a comparison of dew points between instrument types after applying temperature adjustments includes all factors which cause the distribution (including bias) of moisture data to differ.

For each instrument type in a pair, cumulative probability distributions of dew point depressions are obtained in intervals of pressure (with the surface excluded), temperature, and possibly sun angle, using the same lists of stations and time periods which were used to develop temperature adjustments. As with temperature distributions, if the two instrument types are unbiased, or equally biased, they should show similar probability distributions in each interval.

Even though many instrument types are not expected to need temperature adjustments, most or all instrument pairs are expected to show systematic dew point differences. Dew point depression adjustments are more complex than temperature adjustments because for most older and less sensitive instruments, the difference from the reference instrument is usually a narrow probability distribution (infrequent reporting of dry or nearly saturated conditions) as well as a bias. So, the adjustment for an instrument type is stored as a 3-dimensional array. Specifically, within a pressure layer and temperature interval, an adjustment amount (to be added to the reported dew point depression) is specified for different intervals of reported dew point depression. If intervals of sun angle are included because the humidity sensor is affected by solar radiation, then the adjustment array has 4 dimensions. To widen a narrow probability distribution, the dew point depression adjustment is negative in moist cases (a low dew point depression is decreased further, closer to saturation) and positive in dry cases (a high dew point depression is increased, which lowers the reported relative humidity).

As with temperature adjustments, the adjustment from VIZ to the reference is half the difference between VIZ and Vaisala, and the adjustment from Vaisala to the reference is half the difference between Vaisala and VIZ. However, the adjustments are not necessarily the same magnitude with the opposite sign because the shape of the probability distribution is changed by the dew point depression adjustment.

At the end of this step, the temperatures and dew points at each level of each observation are adjusted to be statistically unbiased with respect to the hypothetical reference instrument.

4. PRELIMINARY GLOBAL PRECIPITABLE WATER VARIATIONS SINCE 1973

Because unadjusted and adjusted soundings have the same format (except the adjusted data will have both original and corrected metadata, so the original sounding can be traced), climatology and statistics can be readily developed from either unadjusted or adjusted data.

Time series of the same variables used to infer instrument types are first produced to compare to unadjusted time series. The time series can be used to evaluate the inferred instruments, because a discontinuity is often made worse if an inferred instrument type is incorrect. Detailed data examination may also show that some instrument types need to be split into two or more types. After repeating preceding steps until the inferred instruments appear satisfactory, climatology and other statistics can be prepared.

4.1. Grids and climatology of precipitable water

While the approach to develop grids and statistics is the same for all variables, this research is focused on atmospheric moisture trends, so grids and climatology have been prepared only for total precipitable water.

Daily 2.5° grids of the desired variable are produced first. Spacing of 2.5° is appropriate because few areas have more closely-spaced stations. Procedures are not complex and are not described in detail here. Some special considerations are as follows:

(1) Observations around 0000Z (2100 to 0300Z) are weighted half to the day before 0000Z and half to the day starting 0000Z. Daily grids include the weighted number of observations as well as the value of the variable in each grid box.

(2) For variables such as total precipitable water, where the surface elevation affects the column amount, the quantity at a station is adjusted to the average elevation in the grid box.

(3) For variables with large diurnal variations such as near-surface temperatures, grids of daily average values may be inappropriate.

(4) In daily grids, empty grid boxes are not filled in.

A monthly average grid is simply prepared by summing the weighted values, divided by the sum of the weights, from the daily grids. Such a grid is still sparse and empty grid boxes need to be filled in to produce climatological averages. The grid filling process is summarized as follows, with underlying assumptions stated:

(1) If a grid box has at least as many observations as some defined threshold, the grid box value is accepted.

(2) At grid boxes with few or no observations, surrounding boxes are searched and their values are accumulated with weights declining with distance. When the sum of the weights reaches a threshold, the grid box value is the weighted sum divided by the sum of the weights. The empirical part of this process is that a roughly diamond-shaped area is scanned (farther east and west than north and south from the grid box, except near the poles) because the climate varies less in the zonal than meridional direction. Also, for variables depending on elevation, other grid box values are adjusted to the elevation of this grid box before weighting, and when filling in a low-elevation grid box, the scan in a direction stops if a grid box with an elevation over 750 meters is encountered.

(3) Each annual average is simply the average of the 12 filled-in monthly grids. A global or regional average of a variable weights the grid boxes by area.

This study uses a 30-year climatology from 1973 to 2002, although any sufficiently long period could be used. A climatology is built in the same way as the monthly grids are built, using a grid for each month of the year, but observations in that month for all years are included before filling in the grid. A grid box is accepted without weighting from surrounding points if observations are available on at least 5 percent of all days. The annual climatology is the average of monthly filled climatology grids.

With the empirical scheme of filling in empty grid boxes, the quality of gridding is good even with large data-sparse areas, and can be evaluated by comparing patterns with known climatological processes. With total precipitable water, the main potential problem area is the eastern Pacific. The ITCZ and SPCZ are moderately well reproduced in the western and central Pacific, but the eastern Pacific ITCZ is broader in meridional width and has a lower peak value of total precipitable water than in satellite climatologies, because of the lack of suitable stations in that area.

Climatological averages are most likely to be correct when the spatial patterns are reasonable (within the limitations of sparse station coverage) and the values are as accurate as possible at individual locations. With this analysis, the global annual average is 2.518 cm of precipitable water from 1973 to 2002, close to averages obtained by others.

To develop time series of global or regional averages of a variable, with sparse data it is usually best to construct grids of anomalies and then fill in the anomaly grids, from which the spatial averages are computed. This is because a filled-in average of surrounding anomalies is a conservative estimate for an empty box, but a filled-in average of surrounding absolute values can be very extreme at that location. So, climatological grids are actual variable values, but monthly and annual grids are expressed as anomalies. For some variables, monthly grids of percentages of mean values should be constructed and then filled in.

4.2. Observed global precipitable water variations since 1973

Monthly and annual grids and time series are produced from the unadjusted radiosonde data fairly frequently to look for unexpected trends and variations, and ensure that there are no data problems. Figure 1 shows the latest time series of monthly global average precipitable water anomalies, ending September 2004.

A previous project (Schroeder 2003) developed very preliminary assessments of instrument types and the resulting adjustments, covering 1973 to July 1996. The adjustments in that effort are much less detailed than in the current project, and no temperature adjustments were made. The time series of monthly anomalies of global precipitable water from that project is superimposed on the unadjusted time series, and shows the approximate effect of instrument adjustments on the global trend in the last few decades.

In Figure 1, the black lines are not adjusted for instrument differences, and the blue lines are computed from preliminary adjustments. Note that adjustments

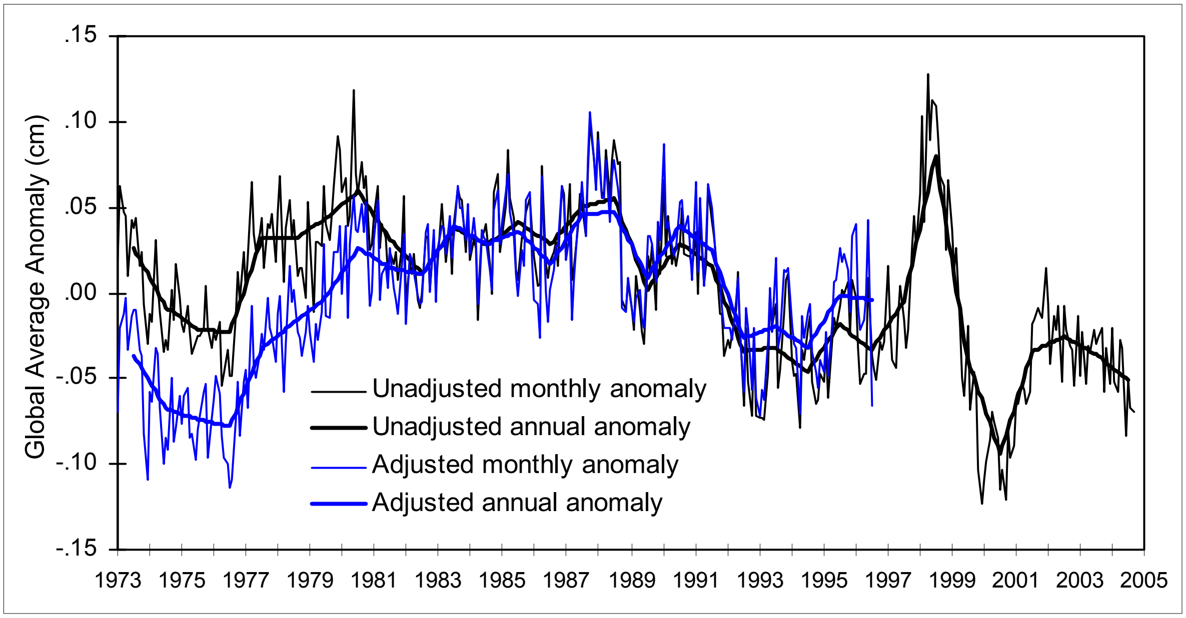

FIGURE 1. Monthly (thin lines) and annual (thick lines) global area-averaged anomalies of precipitable water. The unadjusted values for January 1973 to September 2004 are relative to the 30-year climatology for 1973 through 2002. The adjusted values for January 1973 to July 1996 are based on a previous project and are relative to the average for January 1973 to July 1996.

have a gradual effect on the time series because the transition to drier instrument types has not been over a short period in any substantial part of the world. The adjustments have little effect on the size of short-term or interannual variations, such as the effects of El Niño.

With either unadjusted or adjusted data, the basic trend of global precipitable water shows nearly-steplike changes coinciding with documented climate shifts. Starting from 1973, the initial period was dry, the period from the late 1970s to about 1990 was moist, and the period since then has been generally dry, except for a very large moistening and drying from 1997 to 2001. The 3 climate regimes are more distinct in tropical averages (30° N to 30° S, not shown), with the latest dry regime starting in the tropics from 1988 to 1989.

The moistening in the late 1970s is intensified by the instrument adjustments, since a moistening trend was occurring at the same time that much of the world was transitioning to drier instrument types. The rate of change to drier instrument types slowed since the late 1980s, so it is unlikely that the adjusted data will completely eliminate the dryness of the last 15 years relative to the 1980s. The final transition to drier instrument types in the Russian Federation, India, and China (which has recently begun in all of these countries) will cause some additional drying in the global averages. After those transitions are completed, there will still be fluctuations in future adjustments, but they should be of smaller magnitude and they should not have a "one-way" (exclusively drying) effect on the global averages.

During the transition from the 1997-98 El Niño to the following La Niña, 1998 was exceptionally moist from the eastern Pacific into the Caribbean and in much of the Indian Ocean, and very dry in much of the western Pacific. The basic pattern in the Pacific and Indian Oceans was similar in 1983 as the 1982-83 El Niño decayed. During the persistent La Niña of 1998-2002, 2000 was drier than usual almost globally except from Australia to east of Japan, with a 7 percent decline in annual average precipitable water from 1998 to 2000. After adjusted data is prepared, the pattern of 2000 should be compared with previous dry periods such as 1974 to 1976, and possibly before 1973.

Even with the incomplete state of instrument metadata and adjustments, Figure 1 shows that it is unlikely that there has been a consistent global moistening trend during the persistent global warming since the early 1980s. However, moistening in the late 1970s was large enough that the dry period since the early 1990s is more moist than the 1970s dry period. The main goal of this research is to quantify the moisture trend more accurately. It should then be more feasible to investigate physical mechanisms and feedbacks involved in both interannual and decadal moisture changes, and to relate these shifts to the ongoing global warming trend.

5. DETAILED RUSSIAN RADIOSONDE HISTORY

As an illustration of the considerations and problems in developing complete global historical radiosonde metadata, the process of developing a detailed history for the stations operated by the Russian Federation is described here. The Russian radiosondes illustrate instruments primarily distinguished by statistical characteristics as well as by data characteristics, well-documented and sparsely-documented instruments (the existence of one variety was inferred before documentation was found), accurate documentation at some stations, no documentation at other stations, and an ongoing transition to newer instruments. Lanzante et al. (2003) find serious systematic problems at stations using Russian instruments, although many of the problems occur before 1973.

From 1973 to 2004, stations operated by the Russian Federation include 203 land stations (with 13 sequences where one station replaced another), 33 ships, 3 ice islands, and 6 Antarctic stations. Other stations using Russian instruments, such as in eastern Europe and Vietnam, will be considered in later research.

5.1. Validating existing metadata and inferring complete metadata through early 1990s

Gaffen (1996) metadata gives a detailed history up to about 1992 for almost all of the land stations (a few such as Moscow are omitted), provided by personal communication. The poor metadata quality suspected by Lanzante et al. (2003) may come from disagreement with other sources within Gaffen (1996), which say that every station used A-22 radiosondes in 1976 and MARS in 1986 and 1989, in contrast to detailed station histories which show a variety of instruments in those years.

Russian instruments used up to the early 1990s are well-documented in Zaitseva (1993), and most are on display at the NCDC Weather Museum. Starting 1973, four major instrument series were used:

(1) A-22, with a bimetal thermometer and goldbeater's skin hygrometer.

(2) RKZ, with a rod thermistor and goldbeater's skin hygrometer.

(3) MARS, with a rod thermistor and goldbeater's skin hygrometer.

(4) MRZ, with a rod thermistor and goldbeater's skin hygrometer.

Even though Lanzante et al. (2003) have the same metadata, they report considerable problems with attempting to determine when instrument changes caused data discontinuities, by looking at monthly average temperature anomalies at different levels. They looked at 11 of the Russian Federation stations, plus 2 Russian Antarctic stations, although they do not mention if the Antarctic stations support or do not support their instrument-related breakpoints.

Based on physical examination, the hygrometer appears identical on all of these models, so it is not expected to show much difference in data characteristics. The rod thermistor is the same on all models starting with RKZ-2, so no systematic difference is expected between models, except for a possible difference from A-22. The lack of differences in sensors between most models may help explain why Lanzante et al. (2003) had trouble relating temperature discontinuities to instrument changes. Systematic differences are best supported by examining a large number of stations.

The first main task is to determine if there are any differences in either data characteristics or data reporting, and if these differences correspond to reported dates of instrument transitions, at stations with apparent detailed metadata. The most obvious discontinuity is that almost all stations show a period in the 1980s when the dew point is reported only down to a temperature around -40°, with dew points reported to or near the top of each sounding before and after this period. However, some stations still report dew points to a temperature around -40°. Often, the period with dew points reported to temperatures around -40° matches the claimed period of use of MARS radiosondes. Varying patterns are seen at different stations, but the most common is a few soundings with dew points to temperatures around -40°, followed by a period of all soundings reporting dew points to temperatures around -40°, and a similar later gradual transition to reporting dew points to near the top of the sounding. Other stations have periods of dew points reported to the top of the sounding interspersed with other periods of dew points reported only to temperatures around -40°, and a few stations show only a few soundings with dew points reported to temperatures around -40°. Such behavior would be extremely unlikely if an administrative policy accounted for reporting or not reporting dew points when the temperature is colder than about -40°. The hypothesis that reporting dew points to temperatures around -40° corresponds to Mars soundings is supported at stations which report that instrument type in the 31313 group since 1996. There are few exceptions, which may occur from misreading codes when typing the report manually.

While some stations appear to have slight gradual drying, few steplike discontinuities are found. Therefore, these instrument families are best distinguished by data reporting policies. After examining and cross-checking many data variables, the following are the most robust distinguishing characteristics between instrument families. It is probably not feasible to distinguish between models in a family, such as distinguishing RKZ-2 from RKZ-5, and few stations report such transitions:

(1) A-22 family:

No significant wind levels reported, except possibly the tropopause, highest wind, and top of the sounding.

Relatively few temperature levels reported, but more levels when automation is introduced around 1976.

Dew point almost always reported to the top of the sounding.

Dew point depressions quite small at each level (instrument reads quite moist), and are smaller at 300 than at 500 mb (relative humidity does not drop rapidly with altitude).

Minimum relative humidity in an average sounding quite high, from about 50% in the Arctic to around 30% in southern areas.

Transition from A-22 often coincides with a change in the computed surface elevation.

(2) RKZ family (RKZ-1A, RKZ-2, RKZ-5):

Most reliable indicator: Most soundings have significant wind levels (except in the periods into 1975 and from 1989 to 1991 where no stations show significant wind levels in the archives).

Usually more temperature levels than with A-22.

Dew point usually reported to the top of the sounding, but slightly less often at the top level than with A-22.

Dew point depressions usually slightly larger than for A-22 (slight drying), but not at all stations.

Minimum relative humidity in an average sounding about 4 to 8% lower than with A-22.

RKZ-2 precedes Mars-2-1, and RKZ-5 precedes Mars-2-2, because they use the same radar.

(3) Mars family (Mars-2-1, Mars-2-2):

Most reliable indicator: Dew point reported to a temperature around -40°, with the level at which the dew point terminates often being quite consistent. At very high altitudes, if the temperature warms to above -40°, the dew point is usually reported.

Usually about the same number of temperature and wind levels as with RKZ, often with a slow increase in the number of levels per sounding over several years.

Dew point depressions usually slightly larger than earlier instruments at levels for which dew points are omitted from many observations, but this is simply a consequence of omitting colder cases.

Minimum relative humidity in an average sounding considerably wetter than with RKZ and usually slightly wetter than A-22 due to omission of colder cases.

Mars-2-1 follows RKZ-2 and Mars-2-2 follows RKZ-5 because they use the same radar.

(4) MRZ-3A:

Most reliable indicator: Dew point reported to or near the top of the sounding. This is not a useful indicator when the transition is to MRZ directly from RKZ or A-22 instruments.

Usually about the same number of temperature levels as with MARS, and slightly more wind levels than earlier models (but this indicator cannot be used when no significant wind levels are archived).

Dew point depressions almost always slightly drier than with A-22, often slightly more moist than with RKZ, and usually more moist than with Mars (because Mars omits dew points in colder cases).

Dew point depression at 300 mb usually about the same as at 700 mb.

Minimum relative humidity in an average sounding considerably drier than with Mars, but usually about the same as or only 1 to 2% drier than with RKZ.

If a station used Mars radiosondes, it is often possible to construct an exact history of the use of RKZ, Mars, and MRZ instruments. If a station did not use Mars, or did not make a complete transition to Mars, the only ambiguity is determining when MRZ replaced RKZ. It appears that RKZ was not used after 1989. If a transition from A-22 to RKZ or MRZ occurs in a period when the archive includes significant wind levels, the beginning of RKZ or MRZ is assumed to coincide with the first observation with significant wind levels, and that A-22 is discontinued when most observations have significant wind levels (which is often a year or more after the first significant wind levels).

5.2. Inferring current transition to updated instrument types

Starting in the late 1990s, some Russian stations show steplike drying, but two different amounts of drying are seen. So, it is hypothesized that at least two new radiosonde models are being introduced. Currently there are 6 WMO instrument codes in the 31313 group for the Russian Federation:

27 = AVK-MRZ
28 = Meteorit Mars-2-1
29 = Meteorit Mars-2-2
53 = AVK-RF95
75 = AVK-MRZ-ARMA
76 = AVK-RF95-ARMA

Eventually, documentation of new Russian radiosondes was found in Balagurov et al. (1998, 2002), with some different names than in the WMO codes. The 1998 paper mentions MRZ-3A (goldbeater's skin hygrometer), MRZ-3AM, MRZ-6, and METEOR-1 (the last 3 with capacitive humidity sensors, and the METEOR-1 from Ukraine). Based on the 2002 paper, MRZ-3A is the same as MRZ, and MRZ-6 and METEOR-1 are not operational by 2004. MRZ-3AM uses a DVR capacitive humidity sensor, of "medium" dryness, with some relative humidity reports around 10 to 20%. RF95 uses a Vaisala capacitive humidity sensor (probably the A-Humicap, but this is unstated), and data is even drier with many relative humidity values as low as 1%. Since ARMA is a new ground processing system, the following 7 combinations produce the most consistent metadata:

Level of dryness	WMO code	Instrument type	Ground unit
Moist	28	Mars-2-1	Meteorit-1
Moist	29	Mars-2-2	Meteorit-2
Moist	27	MRZ-3A	AVK
Medium	27	MRZ-3AM	AVK
Medium	75	MRZ-3AM	ARMA
Dry	53	RF95	AVK
Dry	76	RF95	ARMA

Some stations for a while did not report a WMO instrument code that appeared to be appropriate for the level of dryness, but the new WMO codes (53, 75, and 76) may not have been assigned until 1999 or 2000. While Balagurov (2002) states that 10 stations use RF95, it does not list the stations, but using the 31313 reports and checking the level of dryness, exactly 10 stations were using the RF95 by early 2002, plus 3 more stations starting in 2003. In 2004, none of these stations use RF95 exclusively. That reference states that MRZ-3AM is still in testing. Based on detailed data examination, it appears that 15 stations have tested MRZ-3AM (including 10 stations using RF95), with some usage confirmed by WMO code 75 and some inferred by sudden drying while WMO code 27 is reported. Some stations report codes 27, 53, 75, and 76 in the same period, and the behavior of individual observations is consistent with the data characteristics inferred above. So, it should be possible to continue to develop station metadata confidently as the transition continues.

5.3. Summarized findings of examination of Russian-operated stations

The main findings from this detailed examination of the metadata and data for stations of the Russian Federation are as follows:

(1) Detailed station histories for Russian stations from Gaffen (1996) are of fairly good quality, but many transitions occur about 1 or 2 years earlier or later than the stated time.

(2) Many transitions are gradual, some transitions are not complete (the earlier model continues to be used for several years), and in some cases a station goes back to an earlier model for a while.

(3) By checking all stations, a complete instrument history can be derived with very high consistency of signals of different instrument types, even if physical differences in sensors are minor.

(4) The 31313 group is usually but not always correctly used to report instrument changes, and it will become less useful in identifying future models unless new codes can be assigned.

(5) Instrument signals and transitions appear just as distinct for stations without documentation (a few land stations, stations in Antarctica, ships, and ice islands) as for documented stations.

(6) From 1973 to 2004, most stations show 20 to 100 instrument transitions because of gradual transitions to new instrument models.

(7) Only 39 land stations, including station 89592 in Antarctica, have nearly continuous records from 1973 to 2004, with no gaps over a few months long.

(8) Since the early 1990s, 4 stations in the Russian Federation (26038, 26422, 26629, 37789) have switched to Vaisala radiosondes. The transitions are easily seen as sudden drying (with the dew point depression larger at 300 mb than at lower levels), and are more recently confirmed in the 31313 reports. Similarly, Zaitseva (1993) says that 6 Russian ships used Vaisala RS80 by 1991. Based on dryness, these ships are UBNZ, UJFO, UQYC, UUPB, UUQR, and UWEC. Also, URWW was inferred to use Vaisala RS21 in 1981, and ENQT, EOGW, UBNR, and UCKZ used Vaisala RS80 (confirmed by 31313) in the late 1990s. No Russian ships have reported radiosonde observations since 2000.

5.4. Examination of stations in LKS network

The 13 Russian Federation stations (including 2 in Antarctica) in the Lanzante, Klein, Seidel (LKS) network, listed in Lanzante et al. (2003), were looked at in detail because of the problems they report. While the stations illustrate most problems of Russian instruments, it is possible to develop detailed histories for all stations.

For example, at Pechora (station 23418) Lanzante et al. (2003) could not confidently relate temperature discontinuities to documented instrument changes. They chose breakpoints in 1979 and 1987, which are years when no instrument changes occurred. For the period starting 1973, the reported history and a condensed inferred history (the full history shows 55 entries with transitions back and forth between models) are as follows. For the reported instruments, the last 2 entries are from the 31313 reports in the soundings, and all other entries are from Gaffen (1996). Date formats below are YYYY, YYYYMM, or YYYYMMDD/HH (for example, 20040522/00 is 0000Z on 22 May 2004). Reasons for the inferred instrument type, and other notes, are given on the line below each inferred instrument type.

Reported instrument	Inferred instrument
195901 - start A-22	1973 - using A-22
(No sig wind levels, few temperature levels)	
1976 - use A-22	
197611 - start RKZ-5	19780918/12 - start RKZ-5
(Sig wind levels start for almost all observations)	
	19820721/06 ~80% A-22, 20% RKZ-5
(Few obs have significant wind levels)	
	19821209/00 - all RKZ-5
(Sig wind levels resume in most observations)	
198411 - start Mars	19840704/12 - start many obs Mars-2-2
(Dew point reported to temperature around -40°, indiv obs distinguishable: RKZ with dew point reported to or near top, or otherwise Mars)	
	19841120/00 - almost all obs Mars-2-2
(indiv obs distinguishable: Mars or RKZ)	
1986 - use Mars	19860725/12 - all Mars-2-2
(Dew point reported to temperature around -40°)	
1989 - use Mars	
199212 - use Mars or MRZ	
19960507/12 - use Mars-2-2	
20040522/00 - start MRZ	20040522/00 - MRZ-3AM
(31313 says MRZ, dew point reported to or near top of sounding, drier lowest rel hum indicates new sensor)	

For some other Russian stations in the Lanzante (2003) network, findings are as follows:

(1) Station 21504 has an unusual pattern of computed elevation. The elevation averages 35 meters, but starting 19901201/00 the elevation averages 60 m at 0000Z and 35 m at 1200Z, and starting 19910401/00 the elevation averages 60 m. It is inferred that soundings at 35 m used A-22 radiosondes, and soundings at 60 m used MRZ, because significant wind levels are reported for the first time at this station in late 1991 when the archives resume including significant wind levels.

(2) At station 23472, the starting month of MRZ appears correct except that very few MRZ are used (other observations are Mars-2-2) until 13 months later.

(3) Station 24266 transitions directly from A-22 to MRZ (like station 21504) during the 1989 to 1991 period while the archives contain no significant wind levels. The transition is uncertain to less than a year even in this case, but is inferred to be 19901031/12 based on an elevation increase, decreased minimum relative humidity, and beginning of the top wind level often not being a mandatory level, which almost exactly matches the reported Gaffen (1996) transition of 199011.

(4) At station 34731, Gaffen (1996) omits the date of the transition to MRZ, and Lanzante et al. (2003) assign no breakpoints after the 197003 transition from RKZ-1 to RKZ-2. Mars-2-1 (dew points reported to temperature around -40°) is inferred to begin 19860218/12 (the same month as in Gaffen 1996), but only about a quarter of the observations are Mars, with the others being RKZ-2. Since RKZ and Mars usually report complete dew points, the main question is when MRZ-3A replaces RKZ-2. The

most likely date appears to be 19881216/12, after about 4 weeks using only Mars, because the average number of temperature levels per observation increases. Only Mars is used starting 19930506/00 but mixed MRZ and Mars resumes 19940701/00. The 31313 code, starting 19960401/00, reports Mars-2-1 whether the dew point is complete or terminates at a temperature around -40°, but starting 20010710/12, the 31313 code has almost always been accurate. The fact that RKZ, Mars, and MRZ all use the same sensors may help explain why Lanzante et al. (2003) assign no breakpoints since 1970.

(5) Station 38880 reports an unusual transition from MRZ to Mars in Gaffen (1996), with MRZ starting 198505, over a year earlier than at any other station. Also it reports a transition from RKZ-1A to RKZ-2 in 1977, which is much later than most other stations transitioning from RKZ-1A. The transition to RKZ-2 is inferred to be 19770104/12, when more temperature and wind levels per sounding are reported. Based on dew points reported to a temperature around -40°, some Mars-2-1 soundings start 19870826/12, but most soundings continue to be RKZ-2. In May and June 1989, many observations have no dew point data above the surface, and starting 19890630/18, dew points resume to the top of the sounding with a slight increase in the number of temperature levels. This is interpreted as the beginning of MRZ-3A, with defective RKZ-2 units used in May and June 1989.

(6) Station 89542 has no specific history in Gaffen (1996), so as with ships the station history must be entirely inferred. From 1973, RKZ-2 is inferred because a typical observation has a moderately large number of temperature levels, and some significant wind levels are reported starting in 1975 (if no significant wind levels were reported starting in 1975, A-22 would have been inferred). Almost all observations starting 19870620/00 report dew points to a temperature around -40°, and this is inferred to be the beginning of Mars-2-1. (Mars-2-1 rather than Mars-2-2 is inferred, because the archived instrument type starting in 1992, when the codes become specific, is Mars-2-1). When dew points are again reported to or near the top of the sounding starting 19871216/00, this is inferred to be the beginning of MRZ-3A. Based on dew point reporting, the station reverts to Mars-2-1 starting 19950503/00, and then resumes using MRZ-3A (generally complete dew points) on 19960720/00. The 31313 group starting 19970802/00 (shortly before station closure at the end of January 1998) says that Mars-2-1 is used. However, because few stations have reported that they are using Mars while the observations have complete dew points, it is more likely that the 31313 group is incorrect than that a few stations decide to modify their observations to report complete dew points.

Acknowledgements

This project is supported by NOAA Office of Global Programs grants NA16GP2034 and NA04OAR310125.

References

Alduchov, O. A., and R. E. Eskridge, 2002: Determination of radiosonde station elevation from observational data. *J. Applied Meteor.*, **41**, 357-362.

Balagurov, A., A. Kats, and N. Krestyannikova, 1998: Implementation and results of the WMO radiosonde humidity sensors intercomparison (Phase I laboratory test). A. Van Gysegem, ed., *Papers Presented at the WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation (TECO-98)* (Instruments and Observing Methods Report No. 70, WMO-TD-No. 877). WMO, Casablanca, Morocco, 181-184.

Balagurov, A., N. Krestyannikova, and V. Bogov, 2002: Results of humidity sensor intercomparisons of new Russian radiosondes. *Papers Presented at the WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation (TECO2002)* (Instruments and Observing Methods Report No. 75). WMO, Bratislava, Slovakia, online at http://www.wmo.ch/web/www/IMOP/publications/IOM-75-TECO2002/Posters/P1.2(01)Balagurov.doc.

Gaffen, D., 1993: *Historical Changes in Radiosonde Instruments and Practices* (Instruments and Observing Methods Report No. 50, WMO/TD-No. 541). WMO, Geneva, 128 pp.

Gaffen, D., 1996: *A Digitized Metadata Set of Global Upper-Air Station Histories* (NOAA Technical Memorandum ERL ARL-211). NOAA Air Resources Laboratory, Silver Spring, MD, 38 pp.

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

Schroeder, S. R., 2003: Completing instrument metadata and adjusting biases in the radiosonde record to allow determination of global precipitable water trends. *Preprints, 12th Symp. On Meteorological Observations and Instrumentation*, Long Beach, CA, Poster P1.6.

Smith, D. P., 2002: *Data Documentation for Data Sets SSI 6301 - DSI 6379, NCDC Upper Air Digital Files*. National Climatic Data Center, Asheville, NC, http://www1.ncdc.noaa.gov/pub/data/documentlibrary/tddoc/td6300.txt.

WMO, 2002 and earlier years: *Catalogue of Radiosondes and Upper-Air Wind-Finding Systems*. Geneva, WMO, online through http://www.wmo.ch/web/www/ois/ois-home.htm.

Zaitseva, N. A., 1993: Historical developments in radiosonde systems in the former Soviet Union. *Bull. Amer. Meteor. Soc.*, **74**, 1893-1900.