5

Weather Data: Cleaning and Enhancement

Auguste C. Boissonnade; Lawrence J. Heitkemper and David Whitehead

Risk Management Solutions; Earth Satellite Corporation

High quality weather data is used for pricing weather trades, for managing the risk associated to these trades, and for settlement. Forecast weather data is used for evaluating the weather risk up to the time where forecast data ceases to show significant skill. After this period, a combination of seasonal forecasts and historical data is used in pricing weather trades. In addition, near-real time weather data provides the market participants with discrete weather data values to mark the market model for daily risk management purposes and, ultimately, for settlement purposes at the expiration of the weather trades.

The weather market requires data that meets key criteria for use in valuation. This chapter identifies and discusses these criteria, major weather data issues and means for resolving them.

Introduction

The main source for weather data in each country is the national meteorological service (NMS). Each NMS operates independently in setting its standards for weather data collection, archival and distribution.

Weather data are generally classified as either synoptic data or climate data. Synoptic data is the real time data provided for use in aviation safety and forecast modelling. Climate data is the official data record, usually provided after some quality control is performed on the data. Special networks also exist in many countries that may be used in some cases to provide supplementary climate data. These topics and related issues will be further discussed in the next section.

Weather data released by the NMSs often must be "cleaned", ie, replacing missing and erroneous data before the weather market can use it. The types of errors to be expected and methodologies for treating them are discussed in the "Data Cleaning" section.

After weather data are cleaned, problems might still exist that need to be corrected before the data can be used for modelling weather risk. These problems are known as "inhomogeneities" or "discontinuities" caused by station relocations, changes in instrumentation or changes in the surrounding environment near the station site. Treatment of inhomogeneities in the weather data time-series is generally referred to as "homogenisation" or "homogeneity adjustment". This chapter provides a thorough discussion on data enhancement, which is data adjustment resulting from discrete changes in a station's weather data time-series caused by station changes. Non-discrete inhomogeneities such as urbanisation effects are treated as trends.

There are many applications for weather data in weather risk modelling, including the potential use of remotely-sensed data for settlement. Conclusions on weather data use in risk management are presented in the final section.

Weather data

Ideally, the market needs timely and accurate weather data. In order to achieve this, data should be continuously recorded from stations that are properly identified, manned by trained staff or automated with regular maintenance, in good working order and secure from tampering. The stations should also have a long history and not be prone to relocation.

Unfortunately, the main charter of the NMSs in most, if not all countries is "protection of life and property". The emphasis of these offices is focused on accurate diagnosis, tracking and forecasting of severe weather events. The collection and archiving of weather data is important because it provides an economic benefit but the local/national economic needs are not as dependent on high data quality as is the weather risk market. Thus, most government weather service offices do not emphasise data quality issues.

The World Meteorological Organization (WMO) is the international association of NMS members but has no enforcement authority.[1] While various organisations have standardised products, it is apparent that each organisation stands alone in its practices in setting its standard for data collection, archival and distribution. For example, although the WMO has established a standard schedule for the reporting of data at specific times of each day, it has no ability to acquire what data is reported at these times. Daily minimum and maximum temperatures are reported from midnight to midnight of the same calendar day in the US and Germany. However, other European countries report these values in different ways (for example, the United Kingdom Meteorological Office (UKMO) reports the daily minimum temperature from 09.00 Coordinated Universal Time (UTC) of the previous day to 09.00 UTC of the current day, and the current day's maximum temperature from 09.00 UTC of that day to 09.00 UTC of the following day).[2] These individual standards and reporting practices create difficulties in dealing with weather data on a global scale and necessitate "country-dependent" quality control procedures. Knowledge of these standardisation issues and other issues, including the maintenance of historical station information (metadata), maintenance of observation networks, data availability, data usage rights and data pricing, are extremely important to the weather market; without understanding how weather data is collected and processed and without knowledge of the station history, it is very difficult to assess the reliability of these data and to use them. As the weather derivative industry continues to expand globally, these issues are being resolved by the private weather industry. Agreements are made with NMSs for redistribution of raw data and for accessing information on local weather data collection practices and station history.

METEOROLOGICAL NETWORKS

Weather data are recorded from several meteorological networks. The primary purpose of weather networks is to provide synoptic data and climate data. Generally, synoptic data are data obtained simultaneously over a large area of the atmosphere. SYNOP data (data found at stations reporting in accordance with the WMO synoptic reporting standards) undergoes very little quality control and is exchanged among NMSs in order to produce forecasts. Climate data, which are sometimes derived from synoptic data, are data generally used for climatological studies. This data is subjected to quality control procedures and most often forms the official record of historical weather and climate. Climate and synoptic weather data sets may or may not be the same. The same stations may be used for collecting both data sets. However, more quality control checks will be implemented with climate data. In other instances, physically separate observation instruments are used for recording synoptic and climate data.

There are various conventions for identifying stations established by various organisations. The WMO has established a five-digit numbering system for identifying stations for international reporting of synoptic data from these stations. There is also a naming convention for hourly METAR (the acronym roughly translates from the French as Aviation Routine Weather Report) reports consisting of four characters. In the US, the WBAN (Weather Bureau Army Navy) five-digit convention is often used in writing contracts to identify the station. The individual country meteorological networks are identified using the convention selected by the individual country. The lack of a comprehensive international identification standard sometimes causes confusion for the market.

United States

The official source of weather data in the United States is the National Climatic Data Center (NCDC[3]), an agency of the National Oceanic and Atmospheric Administration (NOAA[4]). The US does not significantly differentiate between synoptic and climate networks. US network classification is based upon a tier system with the highest quality stations called "first-order" stations (primary stations). The differences in the class of observing stations are in the variables measured, the frequency of observations being reported, generally the length of data history, and the type of operators manning these stations. The weather market most often selects the first-order stations for pricing and management of US weather risks. These stations operate 24 hours a day and are maintained by National Weather Service (NWS) trained and certified staff.[5] These first-order stations are responsible for observing official climate data as well as providing synoptic data for forecasting purposes. Observations made at these stations span a wide number of variables including temperature, precipitation, surface pressure, humidity, wind speed and direction, cloud cover, snow depth, visibility and solar radiation. These observations are taken and distributed frequently, either on an hourly basis or at specific times of the day. These observations are also cleaned by NWS and reported each month in Local Climatological Data (LCD) reports.

Trained staff that are supervised by NWS personnel also maintain "second-order" stations. These stations record temperature, precipitation and many, but not necessarily all, of the other variables measured at first order stations. These stations provide unofficial climate data as well as some synoptic data. However, no official monthly LCD reports are provided for these stations.

Europe

Similarly, in Europe, there are multiple meteorological networks. The UK has also classified its stations in a tier system based on weather observations quality. The UK system is comparable to the first-order and second-order US system classification, but it is not officially classified in such a way. Many of these stations provide both climate and synoptic data. Stations reporting climate data are located at a combination of commercial airports, Royal Air Force bases and observatories.

In Germany, there are several networks, each one identified by the type of weather observations (for example, rainfall, climate, synoptic, etc).[6] The rainfall network contains up to 4,000 stations while the climate network is comprised of approximately 200 stations.

Japan and Australia

In Japan, there are two types of network, the AMEDAS (Automated Meteorological Data Acquisition System) and SYNOP.[7] Although there are more AMEDAS stations than SYNOP stations (more than 1,000 compared to less than 150), the weather market prefers using SYNOP stations because most of these stations are manned (about 120) and daily temperatures are obtained from continuous readings. Additional supplemental data is available through provincial offices, but the data quality can be questionable. Data reports are sometimes missing and the data is not strictly quality-controlled. Also, the stations may not provide continuous 24-hour reporting.

In Australia, both climate and synoptic data are observed at many of the same stations within the different meteorological networks.[8] The stations used by the weather market are mostly comprised of stations located at airports and manned by government officials.

The Climate Reference Network has been established in Australia to provide data for climate change research. These stations are currently of little use to the weather industry due to their remote locations and limited history. In addition, cooperative volunteers, usually residing in agricultural areas of the country, provide a major supplemental rainfall network of approximately 7,000 stations. This rainfall network may have quality problems, as it is not very well controlled.

Special networks

As well as climate and synoptic networks, NMSs sponsor an array of specialised observation networks. These networks serve the purpose of providing greater spatial coverage and/or providing specialised observations. These networks vary significantly in accuracy, reliability and availability. The US NWS operates a network of cooperative observers.[9] This network of approximately 12,000 active stations serves to provide a greater spatial coverage of meteorological observations. A subset of these stations with a few first- and second-order stations are part of the 1,221-station US Historical Climatology Network (HCN), a database compiled by NCDC and used in analysing US climate (Easterling *et al*., 1999). Many of these stations record only the maximum and minimum temperatures and precipitation, usually on a daily basis. Data for these stations should be used with caution; there are several known problems with the cooperative networks. The observers are volunteers and not professionals, the equipment is basic, data is often missing or erroneous, real-time data availability is poor and the times at which observations are made may vary ("time of observation bias").

PANEL 1

SYNOPTIC VS CLIMATE DATA

One of the issues encountered by the weather market is the definition of climate and synoptic data. In order to ease confusion and set a standard, the Weather Risk Management Association (WRMA[10]) has defined synoptic data – as "data that are collected in real-time at various stations around the globe and provided through the Global Telecommunication System (GTS[11]). The data, minimum and maximum temperature and precipitation – are normally provided four times daily at 00.00 Greenwich Mean Time (GMT), 06.00 GMT, 12.00 GMT, and 18.00 GMT. There is usually a 12-hour minimum and 12-hour maximum temperature that is recorded, but the time which the 12-hour values represent depends on the local time in the country of measurement". Climate data is defined as, "data that are quality controlled by the respective NMSs where the data is collected. The 'climate data' are the 'official' station data of that country". In a review of these two types of data, the WRMA has recommended that climate data is the most appropriate data to be used for the weather derivatives industry.

Historical synoptic data tends to have a shorter record than climate data; the longest history of synoptic data only extends back to 1974. Synoptic records tend to be less accurate because all temperatures were rounded to whole degrees Celsius before 1982 and quality control procedures may not have been thorough. Synoptic data may also contain a large percentage of missing values.

Climate data is the official historical data for a country. This status dictates a dataset with a long period of record (typically extending back to at least 1961), substantial quality control, fewer missing values, greater resolution and a clearly defined definition

Differences between synoptic and climate data, Berlin Templehof, January 1, 1983 to December 31, 2000

Differences in the range	No. of Tmin differences (%)	No. of Tmax differences (%)
>10°C	0.02	0.97
5.0 – 10°C	0.41	1.31
2.0 – 5.0°C	3.45	3.11
1.0 – 2.0°C	4.78	5.74
0.5 – 1.0°C	4.50	6.39
0.0 – 0.5°C	6.96	9.42
any difference	20.12	26.95

Reprinted from: Jewson, S., and D. Whitehead, 2001, "In Praise of Climate Data", *Environmental Finance*, 3:2, pp. 22–3

of the data. The climate of a given day, defined by the relevant NMS, will typically be based upon the calendar day (ie, the average of what the climate has been on that date in the past), whereas the synoptic observations are reported according to WMO schedules. WMO observation schedules are created to provide data for running forecast models, while climate data serves the purpose of providing climatology for a given location.

Differences between ongoing synoptic data and climate data are most apparent in the terms of availability, accuracy and reliability. By its nature, synoptic data is available within minutes to hours of an observation in order to provide data for forecasts, whereas climate data can be delayed by hours to months. These differences in delivery time are the result of varying levels of quality control (QC). Without substantial QC, a high percentage of missing, as well as possibly incorrect data values may be passed to the user. Synoptic data compromises accuracy for availability.

Due to differences in observation time, when analysing the actual time-series, the differences between synoptic data and climate data can be quite large. For example, the table illustrates these differences between the two data sets for Berlin Templehof from January 1, 1984–December 31, 2001 (synoptic data before 1984 was not available). German synoptic data is observed daily in two 12-hour blocks; maximum temperature is observed from 06.00 GMT to 18.00 GMT, while minimum temperature is observed from 18.00 GMT to 06.00 GMT. German climate data is observed over a period of midnight–midnight. The differences between synoptic data and climate data are greater than 1°C occurring approximately 5% of the time.

As the weather derivatives market continues to mature, NMSs are realising the need to provide climate data in a more timely manner. In order to meet this need many NMSs are now producing preliminary climate data from which the final official climate values are derived. Preliminary climate data is composed of observations that undergo a limited QC procedure, but are not thoroughly edited.

STATION HISTORY AND METADATA

Clear advantages exist for organisations that conduct a thorough station history search prior to cleaning and enhancing data. The advantages of acquiring a good quality station history are that it provides the analyst with precise knowledge of when potential discontinuities occurred and what the physical causes of the discontinuities are. Historical information gathered on stations is generally referred to as "metadata".

The availability of future (ongoing) metadata is also very critical for making informed decisions concerning the selection of stations for weather contracts. The historical pricing of a trade will be irrelevant if a station moves during the period of the contract and a sudden shift in daily observations occurs. For the most part, NMSs do not have systems in place to inform users of future station changes. This task requires knowledge of the organisation of the individual meteorological offices, a working relationship with individuals responsible for station changes, and the ability to overcome translation difficulties. Private vendors have compiled and are providing continuously updated information on station history to users.

Many discontinuities do not result from a change in the instrument or relocation of the instrument itself, but from physical changes to the instrument's surrounding environment. These changes can be as big as the construction of buildings nearby to the instrument or installation of a new parking lot, or as seemingly insignificant as the removal of grass from under the sensor. The reporting processes set up by the various NMSs do not specify these surrounding environment changes to be documented; this creates shortcomings in the metadata recording. The private sector has also developed products to overcome this shortcoming. This is accomplished by employing a rigorous metadata research approach including in-depth interviews with airport managers and local meteorological office personnel.

US metadata

The US currently has the most complete metadata in the world available from NCDC via the Internet; information concerning station movement (latitude, longitude and elevation), observers (NOAA, NWS, contractors, etc) and administrative issues are available.[12] However, NCDC metadata is not complete and significant metadata may be missing from the NCDC archives due to local offices not reporting all of the changes and errors inputting the data in the archive.[13]

International metadata

Official metadata is less available in other countries than it is in the US. Metadata is often available in limited quantities, in hard copy and difficult to access. For example, metadata is available for the UK but it is time consuming to access because governmental pre-authorisations are required. Although a modernisation programme is underway by the UK Met Office, currently accessing the full metadata in the UK requires visiting and searching of hard copy materials to one of three locations (depending upon the station in question): Bracknell, Edinburgh or Belfast, in order to obtain the information desired. A similar process is required for acquiring Dutch metadata. A limited amount of German metadata is available in electronic format, but a search of hard copy data is required to obtain all German metadata.

AVAILABILITY AND COST

The unavailability and cost of meteorological data was one of the greatest limiting factors in the global expansion of the weather derivatives industry. Although weather data in the US are inexpensive and very easy to access, many national meteorological services are poorly equipped to meet the needs of the industry and pricing was often set at levels that make the use of the data uneconomical for weather risk valuation.

As the importance of weather data beyond forecasting has become increasingly evident, NMSs have reacted in diverging ways. Some countries, such as the Netherlands,[14] have moved to upgrade systems and implement a free data policy, while countries such as Finland[15] have acted to assert greater control over their data by limiting access. Data usage rights vary from restrictive end-user agreements to free distribution of all products. Data costs can vary from insignificant amounts to thousands of Euros per station for historical data.

Data usage rights and data cost have become increasingly heated issues among the various NMSs. In order to handle these issues in Europe, ECOMET (an economic weather group comprised of representatives of 20 European member countries) was established in 1995.[16] ECOMET's stated objective is "to preserve the free and unrestricted exchange of meteorological information between the NMSs for their operational functions within the framework of WMO regulations and to ensure the widest availability of basic meteorological data and products for commercial applications". In theory, any ECOMET member should act as a one-stop shop for European meteorological data. In practice, this is simply not the case; globally, in order to obtain historical and ongoing climate data, it is necessary to contact each country individually or to work with a private sector entity to secure the data. Each country places different restrictions pertaining to data usage, redistribution and the cost of data. Table 1 illustrates current data availability for selected countries.

Data cleaning

Unfortunately, "raw" weather data obtained from NMSs may be missing or be incorrectly reported. The sources of missing weather data may be that the instrument was broken and the data was never recorded, that there was a break in the transmission of the weather data or that the weather data was recorded and archived, but subsequently lost. Errors in weather data frequently are caused by poorly calibrated instrumentation, but also may be caused by errors in recording the data or while digitising older hard-copy records. In either case, this data must be "cleaned" before accurate valuation analyses can be performed.

ERRORS IN WEATHER DATA REPORTS

Weather data is reported in various formats, METAR and SYNOPTIC, being two common formats.[17] The METAR format is the standard international format for

Table 1. Weather data availability for selected countries

Country	Meteorological Organisation	Historical Data (availability)	Ongoing Data (availability)	Metadata (availability)	Ease of Access
Australia	Australian, Bureau of Meteorology	Yes	Yes	Limited	Easy
Belgium	Royal Meteorological Institute of Belgium	Yes	Yes	Limited	Very difficult
Brazil	Instituto Nacional de Meteorologica (INMET)	Yes	Yes	Limited	Difficult
Canada	Environment Canada	Yes	Yes	Limited	Easy
Finland	Finnish Meteorological Institute	Yes	Yes	No	Very difficult
France	Météo-France	Yes	Yes	Limited	Difficult
Germany	Deutscher Wetterdienst	Yes	Yes	Limited	Moderate
Italy	Italian Air Force*	Yes	Limited	No	Very difficult
Japan	Japan Meteorological Agency	Yes	Yes	Yes	Moderate
Mexico	Servicio Meterológico Nacional	Yes	Yes	Limited	Difficult
Netherlands	Royal Netherlands Meteorological Institute	Yes	Yes	Limited	Moderate
Norway	Norwegian Meteorological Institute	Yes	Yes	Limited	Moderate
Spain	Instituto Nacional De Meteorologica (INM)	Yes	Limited	Limited	Moderate
United Kingdom	The Met Office	Yes	Yes	Limited	Moderate
United States	National Weather Service	Yes	Yes	Yes	Very easy

* The Italian Air Force is the recognised National Meteorological Service of Italy by the World Meteorological Organization (WMO)

reporting hourly values in metric units. Supplementary daily maximum temperature and minimum temperature values are reported at midnight Local Standard Time in the US, providing the earliest information to the weather derivatives market of the previous day values. SYNOPTIC format is the international standard format for reporting values in metric units at standard synoptic hours as discussed in the previous section. A typical reporting error would be to leave out a character in recording the data. Subsequently, all values would be misinterpreted.

QC procedures applied to the data itself will detect other types of errors. Errors include a daily minimum temperature greater than the daily maximum temperature and other unreasonable values in a given day, conflicting values from different sources and inconsistencies in consecutive day values.

The frequency, type and magnitude of the errors are dependent on the country, historical time period, network and type of weather data. For example, China and Japan have a rigorous QC system for their real-time synoptic reporting system and few errors are found, while data from other countries could contain missing and erroneous values. Error frequencies in the US and European countries are low, typically less than 1%, with a wide range from almost no errors to more than 2%, but some large errors do occur.

CLEANING OF WEATHER DATA

Weather data cleaning consists of two processes: the replacement of missing values and the replacement of erroneous values. These processes should be performed simultaneously to obtain the best result.

The replacement of one missing daily value is fairly easy. However, the problem becomes much more complicated if there are blocks of daily missing values. Such cases are not uncommon, particularly several decades ago. The problem of data cleaning then becomes a problem of replacing values by interpolations between observations across several stations (spatial interpolation) and interpolations between observations over time (temporal interpolation).

Spatial interpolation

There are many potential approaches to spatial interpolation of random meteorological inputs. Some of the first work in interpolation in the meteorological field was by Cressman (1959). Cressman's interpolation method corrected a grid-point value linearly combining the residual values between a predicted and an observed value. Barnes (1964) expanded on the work of Cressman, using a linear combination of the observations themselves to generate a "first-guess" field. This field was then used as the input to a second pass of the data. The Barnes approach is still used today in the creation of the objective analysis fields for inputs into numerical meteorological models. The Cressman and Barnes techniques are designed for interpolating a set of grid-point values from a random set of inputs, but can be adapted to interpolate any non-standard grid-point value and to incorporate additional variables, such as point-to-point correlations.

With the advent of Geographic Information System (GIS) technology, there has been an increase in the investigation of interpolation methods to create non-linear surfaces, from which any geo-referenced data point can be derived. This technology is applicable to the problem of replacing missing weather data because the spatial surface of many weather variables approximates non-linear surfaces.[18]

Theoretically, a functional fit, such as the Kriging method, is ideal for interpolation because the weights calculated are chosen such that the function derived provides optimal interpolation. The Kriging approach also has the advantage that it makes it possible to determine the distance at which a value no longer makes a useful contribution to the interpolation. Unfortunately, when the Kriging approach is employed, weather data often does not provide a recognisable optimal functional fit, especially for variables such as rainfall, where the point-to-distance correlation declines rapidly.

One of the simplest forms of interpolation when distance is the only consideration is an inverse-distance or inverse-distance squared weighted averaging. This method works well for fairly homogeneous data such as daily temperature when there is a fairly uniform distribution of data around the target location. The method does not work well for more complex solutions, such as daily rainfall.

Multiple linear regression approaches can also be applied when there is more than one input determining the shape of the data field. For example, such an approach is preferred to inverse-distance weighted methods when interpolating rainfall as a function of distance, elevation and month of the year because this approach often provides a better "fit" of the data and is easier to calculate.

New methods of spatial interpolation of maximum and minimum temperature values employ artificial neural network (ANN) techniques (Snell *et al.*, 2000). While these approaches show superior statistical results compared to simple techniques in "downscaling" gridded forecast temperature values to specific locations, they have not been thoroughly tested for the application of unequally spaced input data nor have they been thoroughly tested for weather variables other than temperature.

Temporal interpolation

Temporal interpolation methods using statistical probability distribution functions have been employed. Gullet *et al.* (1992) and Mekis and Hogg (1999) employ such techniques in rehabilitating daily Canadian rainfall time-series. Hierarchical polynomial regression techniques have also been employed for filling in missing temperature and wind data for agricultural simulation models (Acock and Pachepsky, 2000). Such methodologies may be satisfactory for historical cleaning of small gaps of missing data, but provide little value in real-time data cleaning required for the near-term settlement of weather derivatives. Spatial interpolation methods are thus preferred for real-time cleaning.

COMPLICATIONS IN WEATHER DATA CLEANING

The actual application of data cleaning techniques is complicated by the fact that countries report with different daily timing conventions. For example, in the US, a calendar day is midnight–midnight Local Standard Time, the Canadian day is 06.00–06.00 GMT and the UK reports on a 09.00–09.00 GMT day, with the minimum temperature reported actually being for the previous day. The reporting times are also sometimes different for different parameters. For example, the US reports daily maximum/minimum temperatures at midnight local time and reports 24-hour rainfall at 12.00 GMT. These inconsistencies in reporting require that specific data cleaning techniques be developed for each weather variable and each country.

Another complicating factor in cleaning weather data is that it can be unwise to assume that the closest points correlate higher than ones further away. Microclimate issues often produce inadequate spatial correlation results when using standard geographic spatial cleaning techniques. Optimal techniques, such as a modified Barnes technique, evaluate historical correlation and distances between stations when developing a spatial cleaning approach.

A further complicating factor is that each weather variable has its own unique spatial correlation characteristics. For example, rainfall correlation with distance is uniquely different to temperature or humidity's correlation with distance. The distance correlations may also change with season.

Given all of these requirements for interpolation of weather data, a weather data interpolation model should address the problem of estimating values at non-standard spaced points, using both distance and weather variable correlation to derive the weights.

Impact of instrumentation changes/relocations

Unbiased weather data is crucial in developing models for use in pricing weather risk trades. The private sector undertook projects to address discontinuities in historical temperature data in the US and other countries. Climatologists have long recognised that most climate datasets contain inhomogeneities or discontinuities introduced by non-climatic factors, such as changes of instrumentation, changes of station physical location, changes in surrounding, changes in operation procedure, human errors and changes of operators. Because these non-climate "signals" are "noises" for climate research, climatologists have long attempted to correct these inhomogeneities (Conrad and Pollak, 1950). The task is complicated by occasional malfunctions of instruments (a fan breaking causing rapid excessive warming, for example) or "drifts" out of calibration (for example, dirt and grime can cause a slow warming/cooling trend or a "calibration drift" of the instrument). In many cases, these incidents last for a few weeks or a few months until the next maintenance. These malfunctions are usually not critical but they create additional "noise" in the detection of discontinuities.

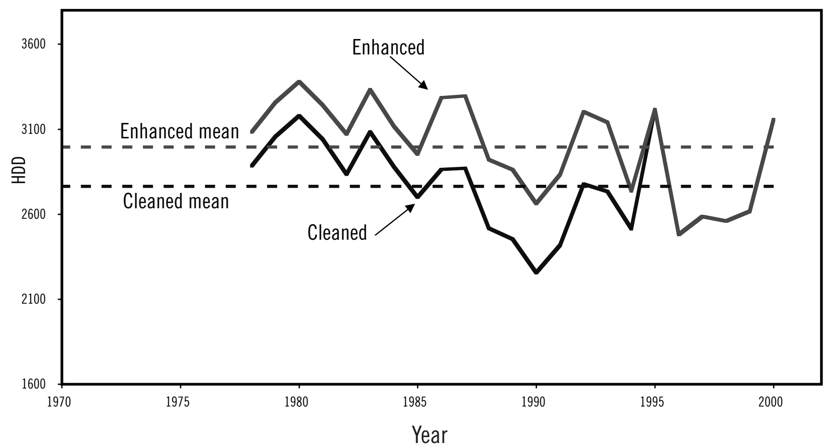
As an illustration, Figure 1 shows the annual HDD indexes for a hypothetical situation where a cooling discontinuity is assumed to occur in 1995, causing an upward shift in the observed HDDs prior to 1995. The bottom series is calculated from the cleaned data and the upper series is estimated after discontinuities have been accounted for; both the cleaned and enhanced means have been superimposed. The purpose of enhancing data is to remove the discontinuities and to adjust the weather series to present day observations. Estimates of the index mean and standard deviation will be biased if the discontinuities are not removed. As a result, the trending analysis will also be biased.

The impact of discontinuities on valuation of weather contracts can be large. For example, a 10-year burn rate for a November–March HDD weather contract can change by as much as 60 degree-days if a discontinuity of 0.8°F is present in the data in 1995 (0.8°F is the average magnitude of discontinuities found by Schrumpf and McKee, 1996, when analysing temperature differences between observations using the recently installed Automated Surface Observing Systems (ASOS) with the previous thermometers).[19] This bias can lead to a large spread in valuation assuming the typical nominal value to be US$5,000 per degree-day.

UNITED STATES

EarthSat and RMS found in a study of 200 first-order US stations (Clemmons and VanderMarck, 2000) that:

1. Cleaned and enhanced annual HDD series



1. almost all the stations have at least one discontinuity over the past 50 years;
2. more than 50% of these stations have at least one discontinuity over the past 10 years; and
3. some stations have discontinuities larger than 2.5°F in either the minimum or maximum temperature time-series.

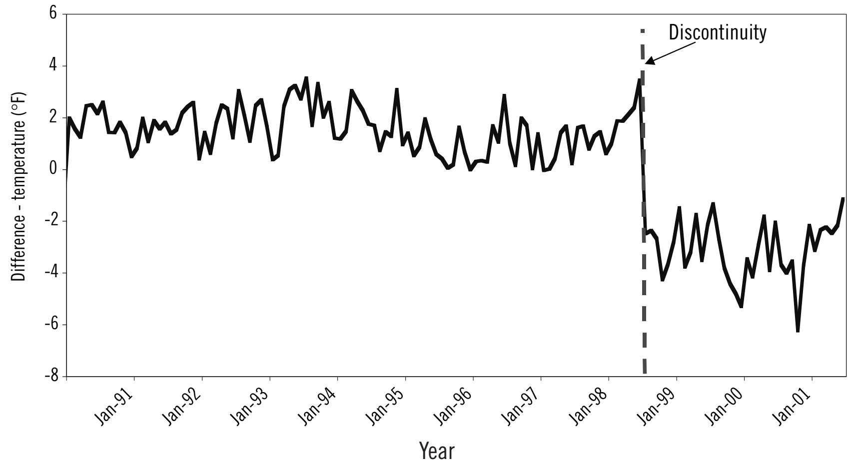
These conclusions are not surprising upon reviewing the history of these stations. First order US stations experienced an average of 13 reported events since 1950 that were potential causes for discontinuities. Equally important is that the rates of such events did not decrease much over the past 10 years. An average of three reported events per station were found to be causes for potential discontinuities since 1992.

Several national changes of instrumentation occurred in the US over the past 40 years. The HO63 hygrothermometer temperature readings gradually replaced maximum/minimum thermometers as the official maximum/minimum recorder between the mid 1960s and the early 1980s, although the HO63s were often initially used only for recording hourly temperatures. A new generation of instruments, the HO83, installed between 1982 and 1986 is generally regarded to have a warm bias. McKee *et al.* (2000) reported that the HO83s had, on average, a warm bias of +0.57°F compared to a calibrated field standard.

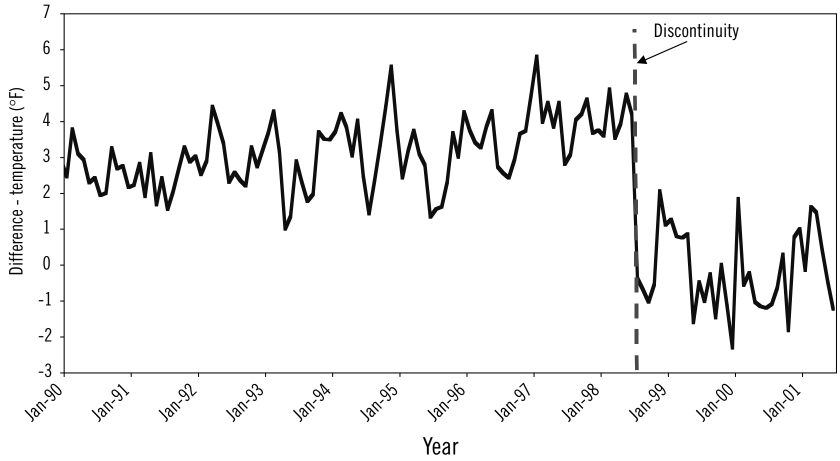
A large programme took place for installing the ASOS type of instrumentation in the 1990s for approximately 850 stations (both first order NWS and second-order Airway stations). At many airports, coincident with the installation of ASOS, the instruments were generally relocated from areas near the Weather Service Office to the end of runway, generally in grassy areas. For example, the instrument in Charlotte, North Carolina was relocated in 1998 from an area near concrete between the Weather Service Office and Butler Aviation Terminal to a low grassy area about 500 feet perpendicular to the end of one of the main runways. An extreme discontinuity, greater than 2°F, was caused by this relocation, which is clearly seen in Figures 2 and 3, showing the monthly temperature difference series between Charlotte and two of its neighbouring stations.[20]

Studies have demonstrated that the installation of the ASOS and relocation of the instrumentation resulted in a cooling of the records in many locations (Schrumpf and McKee, 1996 and McKee *et al.*, 2000). Schrumpf and McKee report, in a sample study of 76 stations (not all first-order stations), that the average temperature difference between ASOS and HO83 instruments is a cooling of –0.79°F with a range of –2.56°F to +0.61°F.

2. Monthly difference temperature series between Charlotte, NC and Greer Greenville-Spartanburg, SC



3. Monthly difference temperature series between Charlotte, NC and Greensboro Piedmont Triad International Airport, NC



A recent survey among ASOS managers on 200 first-order stations shows that, between February 2001 and February 2002, 10 stations experienced events that have potential for discontinuities (mostly station relocations due to airport construction and expansion) and that the rate of pending station changes over the next 12 months will be only slightly less than the rate of experienced events over the past 12 months.

OTHER COUNTRIES

Studies by international researchers (Peterson *et al.*, 1998 and Tuomenvirta, 1998) have also reported the existence of discontinuities in international weather data. Discontinuities are also caused by events such as instrument changes and station enclosure changes (for example, the recent installation of the semi-automatic meteorological observation stations, SAMOS in the UK).

In Japan, a review of metadata indicates that very few relocations occurred until the early 1990s, after which time the number of station moves has been increasing.

In Australia, the general policy of the Bureau of Meteorology is to build a new site, usually at an airport. Then the old site, usually a city site, is discontinued. New identification numbers are assigned at the new sites. If there is a need to provide a continuous record for the two sites, a comparison of rainfall and temperature history is performed in an attempt to remove any in homogeneity caused by the move. In addition, occasional instrumentation relocations at airports do occur, mainly to accommodate airport construction.

Methods for treating discontinuities

There is a potential that any changes in instrumentation, surroundings and relocation of instrumentation may create a serious discontinuity in the data series. An ideal treatment of weather data series containing discontinuities would be to reconstruct the historical observations as if the present instrumentation and surrounding stations conditions were uniform over the historical observation period. Such a task is very difficult because the impact of discontinuities on measured weather variables is "diluted" by the impact of climate changes. A good solution is to use the best features of several methods for identification and quantification of discontinuities in weather series.

Methods should be applied to each weather variable independently. For example, climatic and non-climatic factors affect the minimum and maximum temperature time-series differently. Three questions need to be answered when preparing to remove discontinuities from series:

1. What are the potential dates of discontinuity?
2. Did a discontinuity actually occur?
3. What is the magnitude of the discontinuity?

The first question may be answered by information available on the station (metadata) or by performing supplementary statistical tests on the series if the metadata is incomplete, as is the norm. The second question may be answered by a combination of subjective and objective methodologies. The third question is best answered with mathematical and statistical models.

POTENTIAL DATES OF DISCONTINUITIES

Metadata are invaluable in helping determine the date of discontinuities. Two general classes of events are reported in the metadata: events associated with a station move or instrumentation change (called "true" events) and those described as administrative changes. An example of an administrative change would be the renaming of an airport. Although administrative changes should not affect weather data, there are a limited number of cases where true events occurred coincidentally at these dates. Consequently, methods for identification and quantification of discontinuities first analyse data behaviour before and after true events. A follow-up analysis of data behaviour before and after administrative events is then conducted with stricter acceptance criteria than for true events.

In an ideal situation, true and administrative events reported in the metadata will explain all discontinuities. Unfortunately, this is not the case. Not all events are reported and tests need to be performed for dates other than those listed in the metadata (called "blind dates"). One approach is to perform tests at each month for identification of potential discontinuities using more stringent acceptance criteria than those used for true or administrative events.

DETECTING DISCONTINUITIES

Several papers summarise methods for treating discontinuities (for example, Vincent and Gullet, 1999, Mekis and Hogg, 1999, Alexandersson and Moberg, 1997, Peterson *et al.*, 1998 and Sneyers, 1990). These methods either try to detect and/or to assess the magnitude of the discontinuities by using only data from the station itself (single station) or by developing a reference series from neighbouring stations (multiple stations) thereby removing the effects of regional climate signals.

Detecting discontinuities using one station

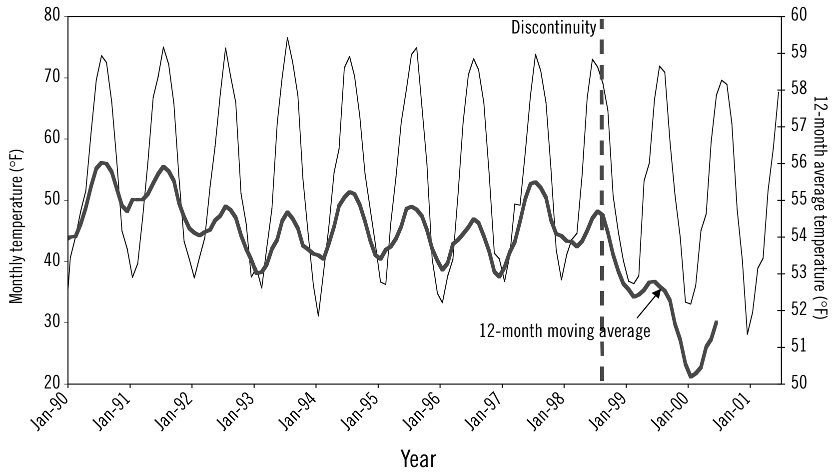
The series itself can be used for detecting large discontinuities after processing the data. Techniques such as those described by Zurbenko *et al.* (1996) and Rhoades and Alinger (1993), where filters are applied to the series, or those reported by DeGaetano and Allen (1999), where quantiles of the distribution of the series are plotted over different periods of time, can be used to detect large discontinuities greater than 2°F. For example, Figure 4 shows a monthly temperature series for Charlotte, North Carolina, around 1998, the year in which a large discontinuity was detected. Although the discontinuity is not easily detected by visual inspection of the monthly series, it can be more easily seen on the superimposed 12-month moving average.

Detecting discontinuities using multiple stations

The preferred solution for detecting non-climatic signals in a climatological series (the "target") is to develop a second series (the "reference"). The reference is generally a weighted aggregate of time-series from neighbouring stations of the target station. The difference (temperature) or the ratio (precipitation) between the target and the reference series attempts to reduce, if not to eliminate, most of the climatological signals.

The use of a reference in the detection of discontinuities in a target assumes that

4. Monthly temperature series for Charlotte, NC



the reference series include the regional climate trends and fluctuations present in the data of the target, and that the reference series do not contain discontinuities themselves during the time period of analysis of the discontinuity in the target.

In practice, it is extremely rare to find neighbouring stations that have exactly the same regional and local climatological signals as the target station and, at the same time, do not contain any discontinuities. Enough stations are needed so that a major discontinuity in one of the neighbouring stations' time-series is not significant in the aggregated reference series and will not prevent the identification of discontinuities in the target station. However, if too many distant (or less correlated) neighbouring stations are used, the resulting reference may not reflect adequately the true climatic fluctuations of the target station.

A simple distance test is not sufficient because the closest stations might correlate poorly while further stations might be subject to similar climate conditions as the target station. Therefore, the first step is to select stations that are strongly correlated with the target station. However, two stations with a strong correlation may mean that the correlation is due to remote teleconnections. Consequently, the analyst also should use some distance function or regional climate check in the selection of the reference stations.

For the above reasons, references are developed by weighting values from a set of neighbouring stations. Each weight is a function of a distance and of a correlation factor.

Once the reference is developed, the major issues in detecting and quantifying discontinuities are:

- ❏ How can a signal be distinguished from the noise with confidence?
- ❏ What time periods should be tested before and after a suspected discontinuity?
- ❏ What if the apparent signal varies over time?
- ❏ What are the trends in the underlying data and how should they be addressed?
- ❏ What are the smallest values that can be detected for a discontinuity?
- ❏ What should be done to detect discontinuities that are close to each other in time?
- ❏ What should be done to handle discontinuities in the neighbouring stations?
- ❏ What should be done to address the variability in the shifts over different time periods?

METHODS FOR DETECTING DISCONTINUITIES AND ADJUSTING WEATHER DATA
The above issues are addressed by using both subjective and objective methods; subjective methods have a quantitative component but are characterised by a final

judgment by an experienced meteorologist, whereas objective methods are fully automated procedures without human intervention.

Subjective methods

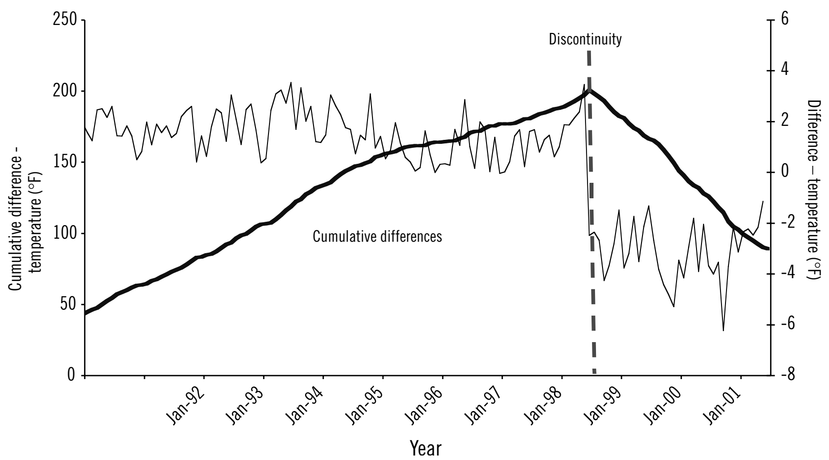
Subjective judgement by experienced meteorologists is important for modifying the weights given to various neighbouring stations based on many factors that cannot all be systematically programmed. For example, a visual review of a difference series between a target station and its neighbouring stations may give an indication of the quality of neighbouring stations' data and of the quality of metadata. This knowledge can lead to a subjective change in parameters of the tests to use in determining and quantifying discontinuities found with objective methods.

An advantage of subjective techniques is that they provide the meteorologist with simple visual tools for the quick identification of periods of time during which one or more discontinuities could exist. Acceptance or rejection of these discontinuities is decided after reviewing the results obtained from objective analyses. One of the first subjective techniques developed was the "double mass analysis" (Kohler, 1949). It plots cumulative difference temperature values over time between a target and its reference. For example, Figure 5 shows a difference time-series between Charlotte and one of its neighbouring stations and, superimposed, the cumulative differences. If the observed temperatures behaved similarly in both stations, then the cumulative difference between the two stations increases uniformly over time. Graphically, the cumulative temperature differences over time would be almost linear, which is observed up to the date of the discontinuity.

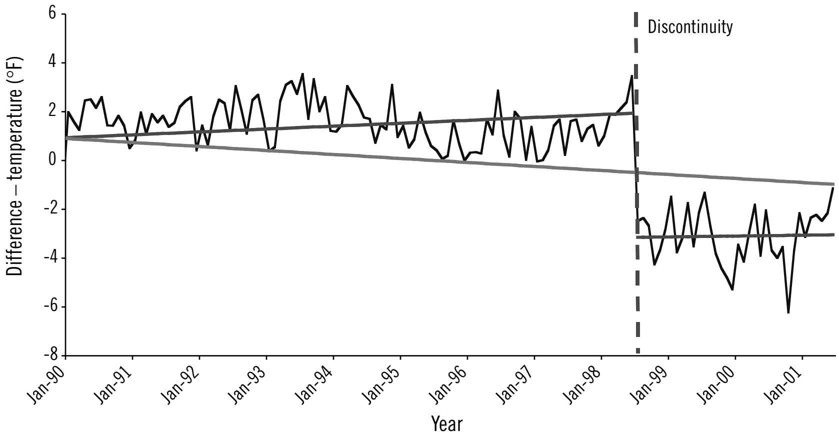
Graphical analysis of the residuals between observed values of the target series and predicted values of the same target series, obtained from linear regression analyses between the target and its references, is useful for detecting discontinuities because the graph of the residuals against time will show an abrupt change if such discontinuities exist. Also, a common assumption is that the residuals are independent random variables with a normal distribution of zero mean and constant standard deviation. The quantile–quantile (q–q) plot is one graphical technique for determining if the residual data sets come from a normal distribution.

Discontinuities can also be visually detected by plotting regression lines on the series before and after a given date. Figure 6 shows the difference series between Charlotte and one of its neighbouring stations and, superimposed, the two regression lines before and after the 1998 discontinuity and the regression line over both periods. Should there have been no discontinuities, the two regression lines would have theoretically been the same as the regression line over the entire series.

5. Double mass analysis applied to Charlotte, NC and Greer Greenville-Spartanburg, SC



6. Monthly difference temperature series between Charlotte, NC and Greer Greenville-Spartanburg, SC



Visual inspection of these plots can usually identify with success large discontinuities (greater than 1.5°F) within a series. Visual inspection also provides "hints" for much smaller discontinuities, which will be further analysed using objective methods.

Objective methods

Most 'stand-alone' objective methods can detect single large discontinuities (1.5°F or larger) with more than 95% (99% for the best methods) accuracy in climatological series. The success rate decreases to about 50% for detecting discontinuities of the order of 0.5°F (Easterling and Peterson, 1995). It is therefore important not to rely solely upon objective methods when detecting discontinuities.

Typical objective methods are mathematical tools that detect the time and magnitude of discontinuities by testing the assumed statistical properties of the difference (or ratio) of assumed truly homogeneous target and reference series. Most of the statistical tests are formulated to automatically detect a discontinuity in the difference (or ratio) series at a given date because the parameters of the distribution of the series (either in the time or frequency domains) have significantly changed statistically before and after this date. Tests are based on the null hypothesis that the series is homogeneous (for example no changes in the means of the difference series before and after a date) and on the alternative hypothesis that the series becomes non-homogeneous at a given date. These types of tests are used in methods developed by Potter (1981), Karl and Williams (1987), Alexandersson (1986) and DeGaetano and Allen (1999).

Another common group of tests using the same approaches have developed statistical significance tests based on either the residuals of linear regression analyses, the parameters of the regressions or the autocorrelation of the residuals (Bois, 1970, Vincent, 1998, Solow, 1987, Easterling and Peterson, 1995).

Other tests use test statistics based on the significance of the rate of changes of the filtered difference (or ratio) series (for example, Zurbenko *et al.*, 1996, Rhoades and Alinger, 1993) and those based on non-parametric tests (Pettitt, 1979 and Sneyers, 1990).

More recent methods (Caussinus and Mestre, 1996 and Szentimrey, 2001) simultaneously generate reference series and detect discontinuities. These methods first identify multiple dates where discontinuities are possible for a set of series in the same climate region. Each of these series can then be used as a homogeneous reference to the other series for the time intervals where no discontinuities are found. When a discontinuity in a series is consistently found with the references derived from the other series, the discontinuity is attributed to this series.

Some discontinuities can be easily detected because they produce a large permanent change in the difference (or ratio) series but other apparent discontinuities are short-term shifts resulting from either changing microclimate conditions or shifts resulting from temporary instrument malfunctions. For these reasons, series need to be analysed over a sufficient time period before and after a potential discontinuity to ascertain a real discontinuity. Typically, several years are required for reliable quantification of discontinuities. However, it is possible to detect and to make initial magnitude estimates of discontinuities within 15 months of the discontinuity date, and even earlier if large discontinuities are present and reviewed by trained meteorologists.

Ideally, no discontinuities should be present in neighbouring stations around the time where a discontinuity has been found in the target stations. If this happens, discontinuities in the target will not be detected or will be falsely detected. Also, true discontinuities will be "smoothed" across several stations. These problems can be prevented by analysing difference (or ratio) series of the target stations with individual neighbouring stations and by estimating the magnitude of discontinuities without using these neighbouring stations.

Because it is important to detect with a high rate of success all discontinuities for pricing a weather transaction, any homogenisation or enhancement methodology must rely on a combination of both subjective and objective methods. An optimal method is to select a sample of found discontinuities of all magnitudes corroborated with metadata and feedback from station operators. This sample can then be used for developing statistical criteria for objective methods and criteria for best applying subjective methods. This process is time-consuming and departs from usual homogenisation methods employed by meteorologists studying climate data but it is needed for enhancing weather data used in valuation of weather transaction.

Quantification of discontinuities

Once the dates of discontinuities have been identified, it would be ideal if "laboratory" experiments were available for assessing the magnitude (shift) of these discontinuities. For example, comparison of overlapping time-series between the old and new instrumentation could detect discontinuities when an instrument change or relocation occurs. Unfortunately, these comparisons are only performed at a very limited number of stations and usually over a very short period of time.

The magnitude of the discontinuity is generally calculated as the differences of the means of the difference (or ratio) series before and after the dates at which discontinuities have been identified. The means are estimated over several years because the difference or ratio series are not really stationary in time (impacts of residuals of climatic and non-climatic signals still corrupt the series). Effects of discontinuities on temperature series might be season-dependent. However, historical detailed meteorological conditions at any time and the physical reasons for the existence of discontinuities are often unknown. Thus, it is often best to apply a constant shift through the year at the risk of creating biases otherwise.

After the discontinuities have been applied, it is appropriate to rerun the methodology to see if any discontinuities may have not been totally resolved in the first iteration of the methodology.

VALIDATION OF METHODOLOGIES FOR DETECTING AND QUANTIFYING DISCONTINUITIES

An excellent way to test the validation of the methods is to perform pair-wise tests, using the same methodology used in detecting discontinuities to systematically compare the target station with several of its neighbouring stations. Both subjective and objective tests need to be performed individually on the target series and the most correlated neighbouring stations to insure the discontinuities are true and that their magnitudes are not corrupted. When possible, interviews with station operators should be conducted for verification of the discontinuities, especially within the past 10 years. It may, however, be less possible to get such feedback for discontinuities older than 10 years.

Finally, homogenisation of any given target station should not drastically change the regional climatology. A rank test is usually performed on the monthly station climatology to make certain the station climatology is not being changed. If the station climatology is being changed, eg, if new monthly temperatures extremes are being created by the homogenisation, extra tests are then needed for validation.

Conclusion

A robust weather derivatives market requires timely, inexpensive, high quality weather data. The main provider of data for each country is the NMS, who has the ultimate responsibility for providing the weather data. Weather data are generally classified as either "climate" or "synoptic". Few countries, such as Australia, do not distinguish between climate data and synoptic data. In most countries, however, there is a significant distinction between these two data sets. Climate data is the official data provided by the NMSs. Synoptic data is the data provided at regular reporting times for input into global weather models.

Weather data provided by the NMSs often has missing values and errors, and is not adjusted for changes in station locations during the period of record. The private weather industry plays a vital role in the weather risk market by providing methodologies for cleaning and enhancing weather data. Weather data cleaning is performed to remove missing values and to replace erroneous values. Weather data enhancement is performed to adjust the weather data time-series for changes in station location.

There are many methods for data cleaning and data enhancement. Data cleaning is achieved by performing a series of quality control checks on the data to identify missing values and to "flag" suspect values. These suspect or missing values can then be replaced using spatial or temporal interpolation techniques. Either method may be used for historical data cleaning but spatial interpolation techniques are preferred for real-time weather data cleaning because they do not require data before and after the replacement date for the interpolation. There are several methods of spatial data interpolation. The model that should solve the problem of estimating non-standard spacing of points would use both distance and point-to-point correlations to derive the interpolation weights.

Data enhancement is performed to determine when discontinuities occurred and to determine the magnitude of the discontinuities. Metadata is very important in evaluating the potential dates of discontinuities. Either single station methodologies or multiple station methodologies may be employed to enhance the time-series. Multiple station methodologies using a reference time-series are often preferred because they remove the effect of regional climatic trends. Validation of any methodology should be conducted using pair-wise tests and operator interviews.

A new technology area yet to be fully explored by the weather derivative industry for weather data is remote sensing technologies. Current radar and satellite technologies can diagnose surface temperature using infrared sensors and rainfall using microwave sensors after proper calibrations (Brown *et al.*, 2001). This could be a solution for areas without instrumentation or where there are reliability issues in the data or moral hazard problems.

1 *The WMO's website can be found at: http://www.wmo.ch/index-en.html.*

2 *The UKMO's website con be found at: http://www.met-office.gov.uk.*

3 *The NCDC's website can be found at: http://lwf.ncdc.noaa.gov/oa/ncdc.html. Weather data can also be obtained from Regional Climate Centers, see http://www.noaa.gov/regions.html.*
4 *The NOAA's website can be found at: http://www.noaa.gov.*
5 *The NWS's website can be found at: http://www.nws.noaa.gov.*
6 *More information can be found on the Deutscher Wetterdienst website. URL: http://www.dwd.de.*
7 *More information can be found on the Japan Meteorological Agency's webiste. URL: http://www.kishou.go.jp/english/index.html.*
8 *More information can be found on the Bureau of Meteorology website. URL: http://www.bom.gov.au.*
9 *http://www.nws.noaa.gov/om/coop/what-is-coop.html.*
10 *The WRMA website can be found at: http://www.wrma.org.*
11 *Information on Global Telecommunication Systems can be found on the WMO's website. URL: http://www.wmo.ch/web/www/TEM/gts.html.*
12 *http://lwf.ncdc.noaa.gov/oa/climate/stationlocator.html.*
13 *For example, when Earth Satellite Corporation (EarthSat) and Risk Management Solutions (RMS) expanded the set of metadata entries currently available in NCDC website log, this research yielded 60% more entries than are available in the NCDC website logs.*
14 *More information can be found on the KNMI website. URL: http://www.knmi.nl.*
15 *More information can be found on the Finnish Meteorological Institute website. URL: http://www.fmi.fi.*
16 *The ECOMET website can be found at: http://www.meteo.be/ECOMET.*
17 *For example, the National Weather Service METAR/TAF information can be accessed at: http://205.156.54.206/oso/oso1/oso12/metar.htm. European METAR information can be accessed at: http://blinder.lfv.se/met/metar.europe.htm.*
18 *Davis (1973) performed some of the original work in this area, but his application was directed at geological formations. Burroughs et al. (1998) was one of the original textbooks written, in which the use of interpolation methods was discussed in a GIS format. Burroughs discusses some of the key methods of "optimal" interpolation, including the Kriging method attributed to the original work of mathematicians Krige (1966) and Matheron (1967).*
19 *As part of a recent modernisation programme, NWS has been installing a standardised instrument package called the Automated Surface Observing Systems (ASOS). Information on these weather stations is available at: http://www.nws.gov/asos/.*
20 *Note that similar comparisons were made with other neighboring stations confirming the discontinuity belongs to Charlotte.*

BIBLIOGRAPHY

Acock, M. C., and Y. A. Pachepsky, 2000, "Estimating Missing Weather Data for Agricultural Simulations using Group Method of Data Handling", *Journal of Applied Meteorology*, 7, pp.1176–84.

Alexandersson, H., 1986, "A Homogeneity Test Applied to Precipitation Data", *International Journal of Climatology*, 6, pp. 661–75.

Alexandersson, H., and A. Moberg, 1997, "Homogenization of Swedish Temperature Data. Part I: Homogeneity Test for Linear Trends", *International Journal of Climatology*, 17, pp. 25–34.

Barnes, S. L., 1964, "A Technique for Maximizing Details in Numerical Weather Map Analysis", *Journal of Applied Meteorology*, 3, pp. 396–409.

Bois, P., 1970, Une methode de controle de series chronologiques utilisees en climatologie et en hydrologie, Laboratory of Fluid Mechanics, University of Grenoble I, p. 49.

Brown, P. E., P. J. Diggle, M. E. Lord and P. C. Young, 2001, "Space-Time Calibration of Radar-Rainfall Data", *Journal of the Royal Statistical Society*, 50, pp. 221–42.

Burroughs, P. A., and R. A. McDonnell, 1998, *Principles of Geographical Information Systems for Land Resource Assessment (Spatial Information Systems and Geostatistics)*, Oxford Press.

Caussinus, H., and O. Mestre, 1996, "New Mathematical Tools and Methodologies for Relative Homogeneity Testing", Proceedings of the First Seminar for Homogenization of Surface Climatological Data, Budapest, pp. 63–82.

Clemmons, L. and P. VanderMarck, 2000, "Raining Stats and Logs", Future and Options World, December 2000, pp. 34–7.

Conrad, V., and C. Pollak, 1950, *Methods in Climatology*, Second Edition, (Harvard University Press) p. 459.

Cressman, G. P., 1959, "An Operational Objective Analysis System", *Mondial Weather Review*, 87, pp. 367–74.

Davis, J. C., 1973, *Statistics and Data Analysis in Geology*, (John Wiley & Sons).

DeGaetano, A. T., and R. J. Allen, 1999, "Single-Station Test to Adjust in Homogeneities in Daily Extreme Temperature Series", Eleventh Conference on Applied Climatology, Dallas, Texas, American Meteorological Society, Boston, pp. 193–5.

Easterling, D. R., and T. C., Peterson, 1995, "A New Method for Detecting Undocumented Discontinuities in Climatological Time Series", *International Journal of Climatology*, 15, pp. 369–77.

Easterling, D. R., T. R. Katl, J. H. Lawrimore and S. A. Del Greco, 1999, "United States Historical Climatology Network Daily Temperature, Precipitation, and Snow Data for 1871–1997", ORNL/CDIAC-118, NDP-070, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee.

Gullett, D. W., W. R. Skinner, and L. Vincent, 1992, "Development of a Historical Canadian Climate Database for Temperature and Other Climate Elements", *Climatological Bulletin*, 26, pp. 125–31.

Karl, T. R., and C. N. Williams, 1987, "An Approach to Adjusting Climatological Time Series for Discontinuous Inhomogeneities", *Journal of Climate Applications*, Meteorology, 26, pp. 1744–63.

Krige, D. G., 1966, "Two-Dimensional Weighted Moving Average Trend Surfaces for Ore Evaluation", *Journal of the South African Institute of Mining and Metallurgy*, 67, pp. 13–79.

Kohler, M. A., 1949, "Double Mass Analysis for Testing the Consistency of Records and for Making Adjustments", *Bulletin of the American Meteorological Society*, 30, pp. 188–9.

Matheron G., 1967, "Kriging, or Polynomial Interpolation Procedures?" *Canadian Mining and Metallurgy Bulletin*, 70, pp. 240–4.

McKee, T. B., N. J. Doesken, C. A. Davey and R. P. Pielke Snr., 2000, "Climate Data Continuity with ASOS – Report for Period April 1996–June 2000", Climatology Report No.00:3, Colorado Climate Center, Colorado State University.

Mekis, E., and W. D. Hogg, 1999, "Rehabilitation and Analysis of Canadian Daily Precipitation Time Series", *Atmosphere–Ocean*, 37(1), pp. 53–85.

Peterson, T. C., D. R. Easterling, T. R. Karl, P. Groisman, N. Nicholls, N. Plummer, S. Torok, I. Auer, R. Boehm, D. Gullett, L. Vincent, R. Heino, H. Tuomenvirta, O. Mestre, T. Szentimrey, J. Salinger, E. J. Forland, I. Hanssen-Bauer, H. Alexandersson, P. Jones and D. Parker, 1998, "Homogeneity Adjustments of in situ Atmospheric Climate Data: A Review", *International Journal of Climatology*, 18, pp. 1493–517.

Pettitt, A. N., 1979, "A Non-Parametric Approach to the Change-Point Problem", *Journal of Applicational Statistics*, 28, pp. 126–35.

Potter, K. W, 1981, "Illustration of a New Test for Detecting a Shift in Mean in Precipitation Series" *Mondial Weather Review*, 109, pp. 2040–5.

Rhoades, D. A., and M. J. Salinger, 1993, "Adjustment of Temperature and Rainfall Records for Site Changes", *International Journal of Climatology*, 13, pp. 899–913.

Schrumpf, A. D., and T. B. McKee, 1996, "Temperature Data Continuity with the Automated Surface Observing System", *Atmospheric Science Paper*, 616, Department of Atmospheric Science, Colorado Sate University.

Snell, S. E, S, Gopal and R. K. Kaukman, 2000, "Spatial Interpolation of Surface Air Temperatures Using Artificial Neural Networks: Evaluating their Use in Downscaling GCM's", *Journal of Climate*, 5, pp. 886–95.

Sneyers, R., 1990, "On the Statistical Analysis of Series of Observations", Technical Note, 143, World Meteorological Organization, Geneva, Switzerland.

Solow, A., 1987, "Testing for Climatic Change: An Application of the Two-Phase Regression Model" *Journal of Climate Application and Meteorology*, 26, pp. 1401–05.

Szentimrey, T., 2001, *Multiple Analysis of Series for Homogenization (MASH v 2.0),* (Budapest: Hungarian Meteorological Service), p.38.

Tuomenvirta, H, 1998, "The Influence of Adjustments on Climatological Time Series", Proceedings of the Second Seminar for Homogenization of Surface Climatological Data, Budapest, pp.73–86.

Vincent, L. A., 1998, "A Technique for the Identification of Inhomogeneities in Canadian Temperature Series", *Journal of Climate*, 11, pp. 1094–104.

Vincent, L. A., and D. W. Gullett, 1999, "Canadian Historical and Homogeneous Temperature Datasets for Climate Change Analyses", *International Journal of Climatology*, 19, pp. 1375–88.

Zurbenko, I., Porter, P. S., Rao, S. T., Ku, J. Y., Gui, R. and R. E. Eskridge, 1996, "Detecting Discontinuities in Time Series of Upper Air Data: Development and Demonstration of an Adaptive Filter Technique", *Journal of Climate*, 9, pp. 3548–3560.