Symbology Modification for Climate Studies

Rich Baldwin (NCDC/NOAA)
Glen Reid (IMSG)

June 3, 2005

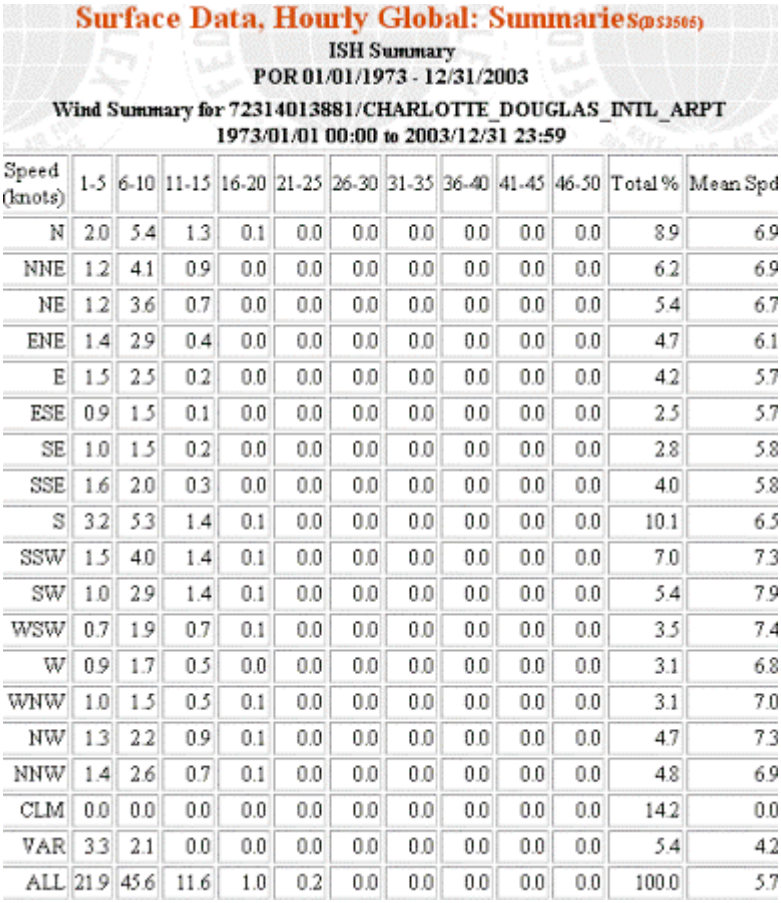
Abstract

The National Climatic Data Center archives worldwide weather data from the present date back to the 1700s. A key goal of this work is to ensure observational accuracy and explanations of isolated variances. Simple indicators of observational normals include climatic data summarizations and frequency distributions. These typically are statistical analyses of station data over 5-, 10-, 20-, 30-year or longer time periods. In a GIS map service, the results of these calculations can be represented by a given symbology set for different statistical criteria and observation type. Having the ability to modify the symbology "on-the-fly" is a useful tool in the analysis of station trends, accuracy, and regional or localized variances. A discussion of a priori processing, persistent database storage, and the technical details of climate analysis map services will be presented.

Introduction

Providing data access to world and US weather observations is a primary goal of the NCDC. With a data archive exceeding 1 petabytes and with the advent of GIS services, this objective is achieved with greater user success and ease. The global surface hourly dataset is one of the most popular available to NCDC users. This dataset totaling over 350 gigabytes is comprised of 40 different types of weather observations with 20,000 stations worldwide. NCDC and the US Navy have been developing value added products in the form of hourly summaries from many of these observations. An example of a climate summary would be a table of wind speed versus wind direction over a 30 year period (Figure 1.) These analyses can be used by the Federal Aviation Administration (FAA) in airport design and by the military in mission planning. Other uses of summaries include conditions for Instrument Flight Rules (IFR) and Visual Flight Rules (VFR), mean temperatures, mean relative-humidity, mean dew point, etc. Currently, there are 13 such summaries that are generated and stored persistently in a key/value database, Berkeley DB[1]. Table 1 gives an overview.

Analyses of these summaries on a global scale provide values used in symbology classification, which are stored in an Oracle/ArcSDE database. Retrieval and display of this information is done using ArcIMS and ArcSDE. A series of anomaly maps convey the stations' statistics for selected criteria, providing a broader understanding of trends, data quality, and completeness. A quick summary access function is also provided in the map service to allow users an option for comparative studies of a given station summary.



Surface Data, Hourly Global: Summaries(DS3505)

ISH Summary

POR 01/01/1973 - 12/31/2003

Wind Summary for 72314013881/CHARLOTTE_DOUGLAS_INTL_ARPT

1973/01/01 00:00 to 2003/12/31 23:59

Speed (knots)	1-5	6-10	11-15	16-20	21-25	26-30	31-35	36-40	41-45	46-50	Total %	Mean Spd
N	2.0	5.4	1.3	0.1	0.0	0.0	0.0	0.0	0.0	0.0	8.9	6.9
NNE	1.2	4.1	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.2	6.9
NE	1.2	3.6	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.4	6.7
ENE	1.4	2.9	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.7	6.1
E	1.5	2.5	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.2	5.7
ESE	0.9	1.5	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.5	5.7
SE	1.0	1.5	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.8	5.8
SSE	1.6	2.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.0	5.8
S	3.2	5.3	1.4	0.1	0.0	0.0	0.0	0.0	0.0	0.0	10.1	6.5
SSW	1.5	4.0	1.4	0.1	0.0	0.0	0.0	0.0	0.0	0.0	7.0	7.3
SW	1.0	2.9	1.4	0.1	0.0	0.0	0.0	0.0	0.0	0.0	5.4	7.9
WSW	0.7	1.9	0.7	0.1	0.0	0.0	0.0	0.0	0.0	0.0	3.5	7.4
W	0.9	1.7	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.1	6.8
WNW	1.0	1.5	0.5	0.1	0.0	0.0	0.0	0.0	0.0	0.0	3.1	7.0
NW	1.3	2.2	0.9	0.1	0.0	0.0	0.0	0.0	0.0	0.0	4.7	7.3
NNW	1.4	2.6	0.7	0.1	0.0	0.0	0.0	0.0	0.0	0.0	4.8	6.9
CLM	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	14.2	0.0
VAR	3.3	2.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.4	4.2
ALL	21.9	45.6	11.6	1.0	0.2	0.0	0.0	0.0	0.0	0.0	100.0	5.7

Figure 1. 30 year wind speed-direction summary for Charlotte airport.

Ceiling Height – Visibility
Flying Condition (ceiling height < 2500feet, visibility > 5miles, wind speed < 13 knots)
Flying Condition (ceiling height < 1500feet, visibility > 3miles, wind speed < 13 knots)
Flying Condition (ceiling height < 1500feet, visibility > 3miles, wind speed < 17 knots)
Flying Condition (ceiling height < 1500feet, visibility > 3miles, wind speed < 20 knots)
Wind Speed versus Direction
Dew Point Statistics
Temperature Statistics
Station Pressure Statistics
Sea Level Pressure Statistics
Present Weather Conditions
Sky Cover
Relative Humidity

Table1. Types of summaries for global station data recorded hourly

Process

For the climate analysis map services, four basic steps were completed for the automatic generation of symbology: 1) The global hourly database was built; 2) 10 year summaries for 8000+ stations were computed for the 13 different summary categories and stored in a key/value database; 3) The key/value database is interrogated for specific columns and rows for statistics of interest, and an SDE database is updated with this information; and 4) ArcIMS map services are modified to access different symbology classifications.

Development of the global hourly database has been an ongoing effort over many years. These data are a combination of numerous datasets from various sources: National Oceanic Atmospheric Administration (NOAA); US Air Force; US Navy; and World Meteorological Organization (WMO). The merging of the data involves a series of C, FORTRAN, and PL/SQL routines [2]. Currently, there are 15 years of hourly station data available online through the NCDC relational database, and greater than 100 years available via FTP.

There are over 8000 hourly stations worldwide which have a complete period of record over the last 10 years. These stations are used to compute summaries for all of the different types found in Table 1. These summaries are computed and stored in a Java collection class as a TreeMap. These objects can be persisted in the Berkeley DB database which is very well suited for rapid storage and access [3]. Each TreeMap object is persisted as a value with an appropriate key representing the type and date range of each summary. This results in a highly portable, easily accessible database which is fast and Open Source.

In every summary type there are conditions of particular interest to the climatologist. IFR and VFR conditions have been previously mentioned. As an example for ceiling height versus visibility, a pilot might come to an airport weather office wanting to know the VFR flying conditions (ceiling 1000ft with visibility at 3 miles) for any given destination. It is a simple task to identify these values in the Berkeley DB database and insert the values in ArcSDE. When the Berkeley DB database is re-cursed, the station identifier and percentage or average from the summary is stored in the ArcSDE database. When all stations have been processed, the map service can re-classify the symbology base on the user's selection. Since our 8000+ summary stations are a subset of the 12,000+ stations in the global hourly station layer, an ArcSDE view was initially created from the entire layer. A left outer join (a joining of 2 tables which retrieves all records from the first table and any records from the second table where the column values match) with the summary table was then performed on the view in Oracle. This procedure allows us to maintain the integrity of the ArcSDE layer for use in the ArcIMS configuration file (AXL). Using an ArcSDE view also gives us a seamless updating method, as changes to the summary table will automatically be reflected in the map.

ArcIMS references the ArcSDE view in the ImageServer configuration file. The summary symbolization is done through a custom ArcXML request to the spatial server.

The VALUEMAPRENDERER request renders features according to a value in a specified field. The request is built using a pre-defined field and upper and lower range values, all of which will depend on the type of summary classification to be performed. Color gradients are in accordance to the classification type and are also pre-defined. Both the range and color definitions are stored in a JavaScript array. Once built, the custom VALUEMAPRENDERER request is inserted into the LAYERLIST section (for the global surface hourly layer) of the larger GET_IMAGE request. The custom rendering thus overrides the default rendering for the layer.

Results

Users accessing the climate analysis map service can select the "Classify Summary" function option to obtain a drop down of available symbology classifications. Figure 2 shows the mean July temperature for stations over the last 10 years. Users can navigate through the map to identify areas or stations of interest. The function option "Quick Summary" gives access to complete station summaries (Figure 3 and 4). An additional example of ceiling height versus visibility is shown in Figure 5 and 6 for VFR flight conditions (1000ft ceiling height and 3 miles visibility).

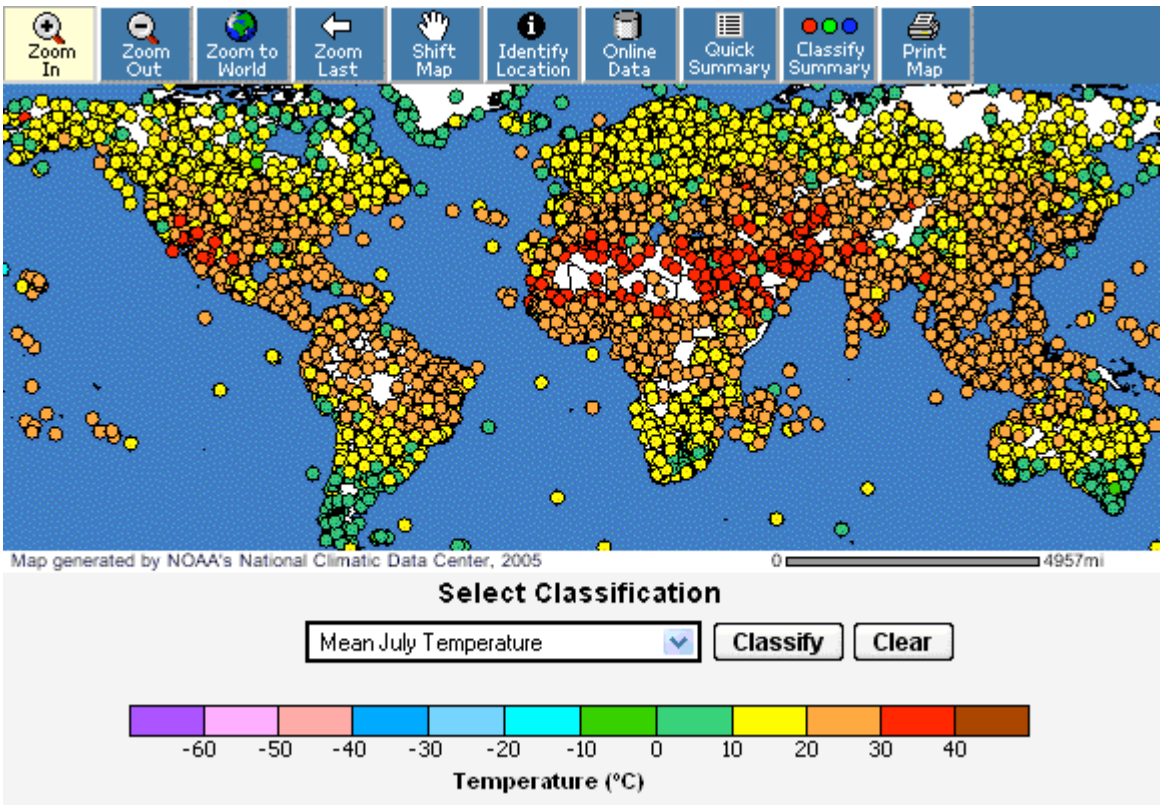


Figure 2. Station symbols reclassified to show the mean July temperatures over the last 10 years.

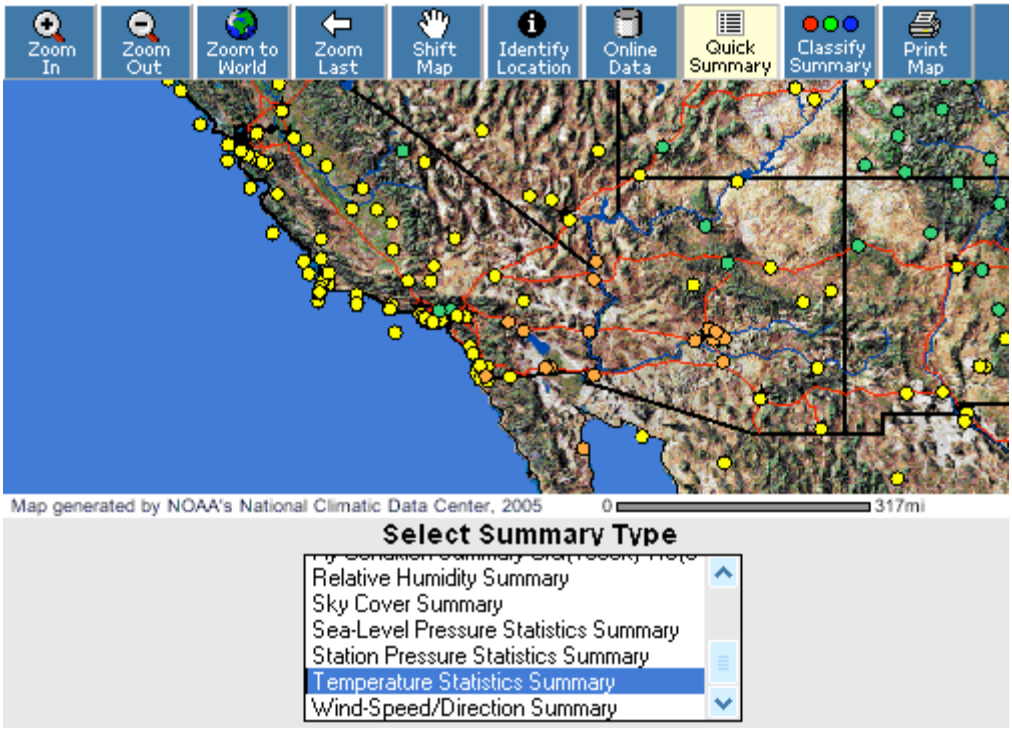
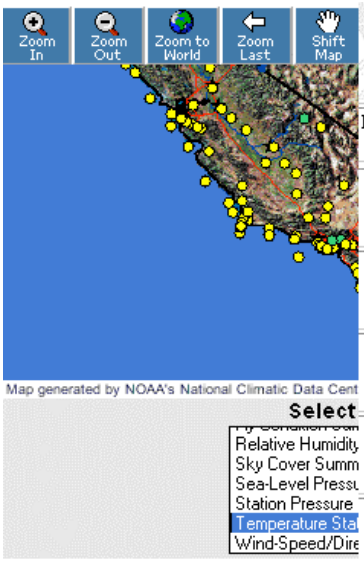


Figure 3. 10 year mean July temperature symbol classification for the Southwest US.



ISH Summary
POR 01/01/1995 - 12/31/2004
Temperature Summary for 72290993115/IMPERIAL BEACH REAM FIELD NAS;
1995/01/01 00:00 to 2004/12/31 23:59;

HOUR (UTC)		JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL
0	mean	15.7	15.9	16.3	16.9	18.0	19.1	20.9	21.8	21.7	19.6	17.8	16.0	18.3
	stdv	2.4	2.3	2.4	2.3	1.7	1.6	1.4	1.5	2.1	2.1	2.1	2.0	2.0
	#obs	183.0	173.0	206.0	207.0	200.0	198.0	194.0	207.0	196.0	203.0	173.0	172.0	2312.0
1	mean	14.9	15.3	15.8	16.5	17.5	18.6	20.4	21.4	21.0	18.8	16.7	14.8	17.6
	stdv	2.1	2.3	2.2	2.5	1.7	1.5	1.4	1.4	2.2	1.9	1.9	1.9	1.9
	#obs	188.0	173.0	199.0	198.0	195.0	187.0	198.0	204.0	170.0	182.0	174.0	171.0	2239.0
2	mean	13.9	14.4	14.8	15.7	16.9	18.0	19.8	20.7	20.2	18.0	15.7	13.6	16.8
	stdv	2.3	2.0	2.0	2.4	1.6	1.4	1.4	1.4	2.1	1.5	2.1	2.3	1.9
	#obs	178.0	174.0	207.0	189.0	185.0	184.0	178.0	189.0	162.0	172.0	160.0	156.0	2134.0
3	mean	12.8	13.6	14.2	14.9	16.3	17.5	19.1	20.0	19.8	17.5	14.6	12.3	16.0
	stdv	2.4	2.2	1.9	1.9	1.5	1.4	1.3	1.3	2.1	1.7	2.5	2.8	1.9
	#obs	164.0	167.0	193.0	194.0	192.0	187.0	179.0	181.0	161.0	183.0	154.0	149.0	2104.0
4	mean	12.0	12.9	13.9	14.5	16.0	17.2	18.7	19.7	19.5	16.7	13.8	11.6	15.5
	stdv	2.9	2.7	2.1	1.9	1.6	1.3	1.3	1.3	2.3	2.2	2.7	3.1	2.1
	#obs	162.0	162.0	190.0	188.0	192.0	178.0	176.0	187.0	163.0	169.0	148.0	150.0	2065.0

Figure 4. Map and temperature summary example

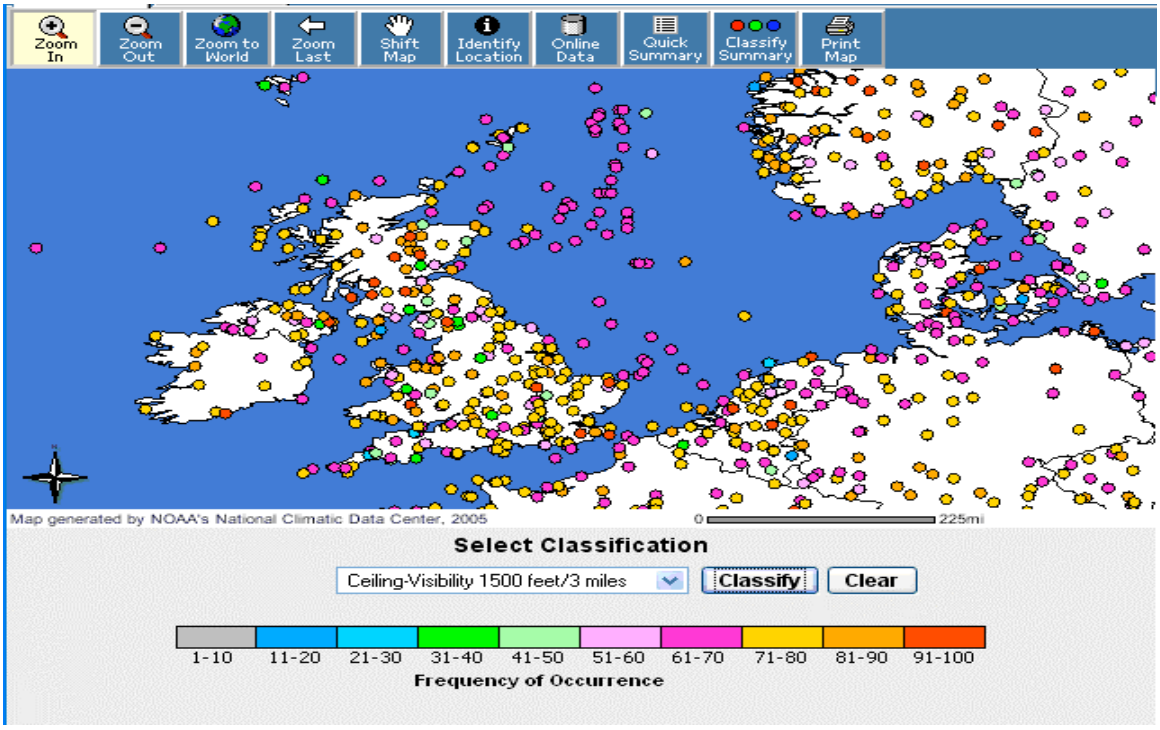
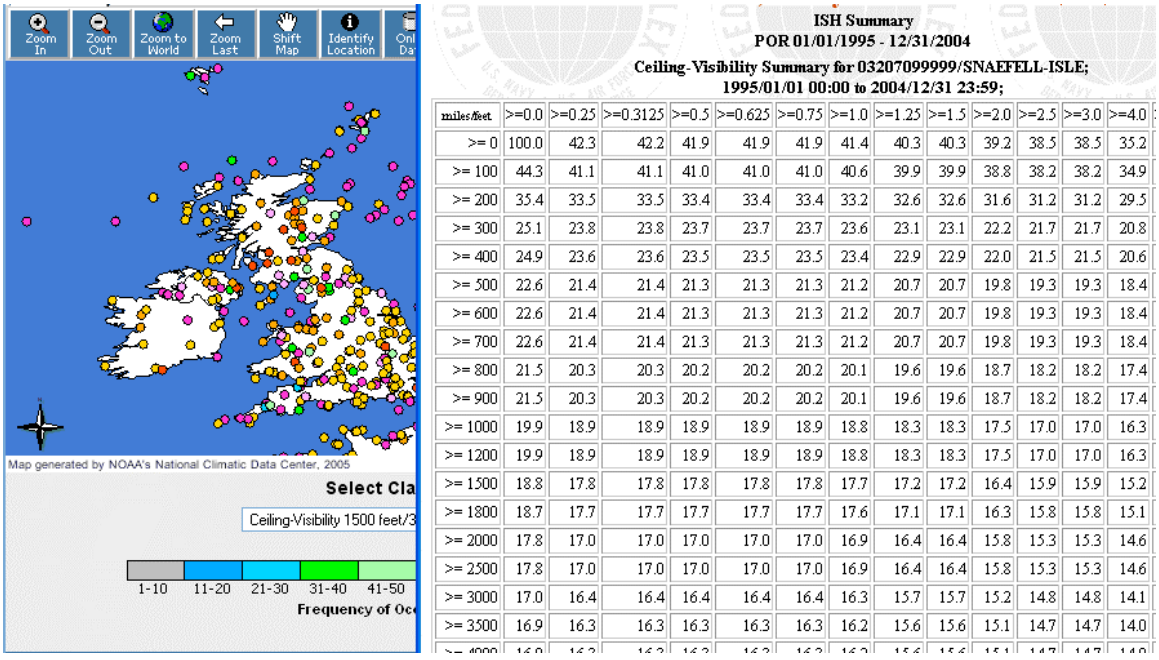


Figure 5. VFR ceiling-visibility classification for England and Northern Europe



ISH Summary
POR 01/01/1995 - 12/31/2004
Ceiling-Visibility Summary for 03207099999/SNAEFELL-ISLE;
1995/01/01 00:00 to 2004/12/31 23:59;

miles/feet	>=0.0	>=0.25	>=0.3125	>=0.5	>=0.625	>=0.75	>=1.0	>=1.25	>=1.5	>=2.0	>=2.5	>=3.0	>=4.0
>= 0	100.0	42.3	42.2	41.9	41.9	41.9	41.4	40.3	40.3	39.2	38.5	38.5	35.2
>= 100	44.3	41.1	41.1	41.0	41.0	41.0	40.6	39.9	39.9	38.8	38.2	38.2	34.9
>= 200	35.4	33.5	33.5	33.4	33.4	33.4	33.2	32.6	32.6	31.6	31.2	31.2	29.5
>= 300	25.1	23.8	23.8	23.7	23.7	23.7	23.6	23.1	23.1	22.2	21.7	21.7	20.8
>= 400	24.9	23.6	23.6	23.5	23.5	23.5	23.4	22.9	22.9	22.0	21.5	21.5	20.6
>= 500	22.6	21.4	21.4	21.3	21.3	21.3	21.2	20.7	20.7	19.8	19.3	19.3	18.4
>= 600	22.6	21.4	21.4	21.3	21.3	21.3	21.2	20.7	20.7	19.8	19.3	19.3	18.4
>= 700	22.6	21.4	21.4	21.3	21.3	21.3	21.2	20.7	20.7	19.8	19.3	19.3	18.4
>= 800	21.5	20.3	20.3	20.2	20.2	20.2	20.1	19.6	19.6	18.7	18.2	18.2	17.4
>= 900	21.5	20.3	20.3	20.2	20.2	20.2	20.1	19.6	19.6	18.7	18.2	18.2	17.4
>= 1000	19.9	18.9	18.9	18.9	18.9	18.9	18.8	18.3	18.3	17.5	17.0	17.0	16.3
>= 1200	19.9	18.9	18.9	18.9	18.9	18.9	18.8	18.3	18.3	17.5	17.0	17.0	16.3
>= 1500	18.8	17.8	17.8	17.8	17.8	17.8	17.7	17.2	17.2	16.4	15.9	15.9	15.2
>= 1800	18.7	17.7	17.7	17.7	17.7	17.7	17.6	17.1	17.1	16.3	15.8	15.8	15.1
>= 2000	17.8	17.0	17.0	17.0	17.0	17.0	16.9	16.4	16.4	15.8	15.3	15.3	14.6
>= 2500	17.8	17.0	17.0	17.0	17.0	17.0	16.9	16.4	16.4	15.8	15.3	15.3	14.6
>= 3000	17.0	16.4	16.4	16.4	16.4	16.4	16.3	15.7	15.7	15.2	14.8	14.8	14.1
>= 3500	16.9	16.3	16.3	16.3	16.3	16.3	16.2	15.6	15.6	15.1	14.7	14.7	14.0

Figure 6. Map and ceiling-visibility summary example

Conclusion

Summary symbology modification through classification allows studies of climate trend analysis, data quality control, and data completeness to be carried out using the climate analysis map service. The anomaly maps provide fast and detailed station information to the user.

Acknowledgements

Many thanks to NCDC, IMSG, TMC, and the US Navy staff for support of this effort.

End Notes

Climatic analysis map services will be available through http://www.ncdc.noaa.gov/

References

[1] Berkeley DB, http://www.sleepycat.com/

[2] N. Lott, R. Baldwin, P. Jones, The FCC Integrated Surface Hourly Database, A New Resource of Global Climate Data, US DOC/NOAA/NCDC Technical Report 2001-01, 2001.

[3] R. Baldwin, Views, Objects, and Persistence for Accessing a High Volume Global Data Set, Proceedings of the 20th IEEE/ 11th NASA Goddard Conference on Mass Storage Systems and Technologies, pp 77 – 81.

Author Information

Mr. Rich Baldwin
National Climatic Data Center, National Oceanic and Atmospheric Administration
151 Patton Ave., Asheville, NC 28801
828-271-4094, 828-271-4876 (fax)
Rich.Baldwin@noaa.gov

Mr. Glen Reid
IMSG
National Climatic Data Center, National Oceanic and Atmospheric Administration
151 Patton Ave., Asheville, NC 28801
828-350-2004, 828-271-4876 (fax)
Glen.Reid@noaa.gov