3B.5 The Integrated Surface Database: Partnerships and Progress

Neal Lott*, Russell Vose, Stephen A. Del Greco, and Thomas F. Ross
NOAA National Climatic Data Center, Asheville, North Carolina

Steven Worley and Joey Comeaux
National Center for Atmospheric Research, Boulder, Colorado

ABSTRACT

The National Oceanic and Atmospheric Administration (NOAA) National Climatic Data Center (NCDC) is responsible for the stewardship of global climatological data to support a variety of applications in the private, research, and public sectors. This includes the development of datasets to be used as input for a planned global climatological reanalysis. In support of these activities and to better streamline operations, NCDC implemented a plan to integrate surface climatological data into a common format and data model and process these data through one processing system that capitalizes on network-independent standardized quality control algorithms and procedures. Federal Climate Complex (FCC) partners (US Air Force and Navy) collaborated closely with NCDC in development of the global Integrated Surface Database (ISD), providing large volumes of data as input sources. Now, through additional partnerships, ISD is being further expanded and enhanced.

1. BACKGROUND

The Integrated Surface Database (ISD) consists of global surface observations compiled from numerous sources, into a single common ASCII format and common data model. Its development began in 1998 as a joint activity within Asheville's FCC (NCDC, U.S. Air Force, U.S. Navy). ISD integrates data from over 100 original data sources, including numerous data formats which were key-entered from paper forms. The database of over 20,000 stations comprises over 1.7 billion surface observations, includes data from as early as the late 1800's (many stations beginning in 1948-1973 timeframe), is operationally updated with the latest data, and is used by numerous customers in many applications. These figures will be increasing substantially over the next several years.

The development of ISD (previously called ISH—Integrated Surface Hourly) has been an iterative process. This includes the development of the integrated format, collection of datasets to include in the initial ISD database, development of a data model to use in a relational database for customer servicing, quality control of the historical ISD, development of the end-to-end Integrated Surface Data Processing System (ISDPS), and development of online products.

* *Corresponding author address:* Neal Lott, NOAA National Climatic Data Center, 151 Patton Avenue, Asheville, NC 28801; e-mail: Neal.Lott@noaa.gov.

ISD Version 1 was released in 2001, with Version 2 (additional quality control applied) in 2003. Since 2003, there have been continued incremental improvements in automated quality control software, along with additional partnerships to further enhance the temporal and spatial coverage of the data.

The current ISD partnerships include:

- The Climate Data Modernization Program (CDMP), which provided for publications and forms as far back as the 1800's to be scanned, digitized, and integrated into ISD, and includes data processing at the Northeast Regional Climate Center (NERCC).
- The National Center for Atmospheric Research (NCAR), which is providing numerous datasets of global and national origin to be integrated into ISD.
- The FCC's 14th Weather Squadron (formerly the Air Force Combat Climatology Center) and US Navy Fleet Numerical Meteorological and Oceanographic Command Detachment (FNMOC Det), which provide continued flow and processing of global hourly, synoptic, and military station data.
- NOAA's National Weather Service, the Federal Aviation Administration, and NOAA's Climate Reference Network (CRN), which provide datastreams into ISD on a daily basis.

2. RECENT PROGRESS

The CDMP will continue to transform NOAA's paper and microform archives into an electronic database. Access to this database by business and economic sectors will allow these decision makers to use climate information when making decisions. The demand for rapid and complete access to the Nation's and world's climate data is a key driver of CDMP, which is managed through NCDC. This program was initiated by Congress to assist NOAA in modernizing and improving access to the Nation's climate data and information. Partnering with four private sector contractors, CDMP has placed online over 52 million weather and environmental images. These historic documents are now available to researchers and decision makers in both the public and private sector via the Internet.

During the next few years, CDMP will continue to make available keyed hourly data for several early climate locations that recorded hourly temperature data during the 19th century. These data will extend the

period of record for selected locations back into the early to mid 1800's for use in climate change and variability studies.

The number of CDMP historical hourly observations for 1928-1948 integrated into ISD is shown in Figure 1, and this now includes observations extending back to 1893 for U.S. locations. CDMP is also keying data for several other countries, and those data will be integrated into ISD.

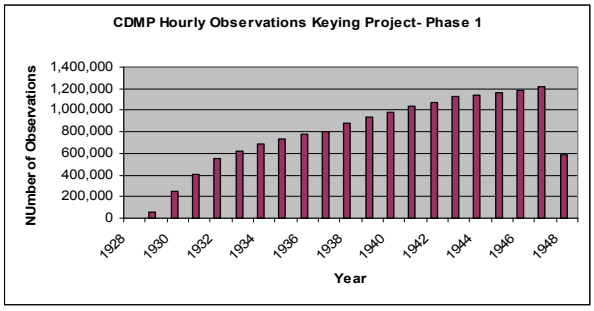


Figure 1. CDMP Hourly Data.

NCAR is now a key partner in the ISD effort, with numerous datasets from NCAR being made available for integration. Most recently, datasets from Brazil, Australia, Greenland, and Mexico were converted to ISD format, quality controlled, then merged and integrated into ISD. This (along with the CDMP effort) provided an additional 42 million surface observations covering a period of 108 years.

Through the partnership with NCAR, numerous additional datasets are being inventoried, evaluated, and prioritized for inclusion, to extend the spatial and temporal coverage of ISD. This will improve the usefulness of the data in reanalyses and climate change studies. Reanalysis requirements are among the main drivers in prioritizing datasets for integration.

Approximately fifty additional datasets have been identified for potential integration into ISD. Some are nation or region-specific, such as Russia and western Africa; some are mostly U.S.; and some are global in coverage. Also, through an effort recently undertaken with Environment Canada, we hope to eventually include the digital archives from Canada in ISD. It is important to note that ISD was a global database (Figure 3) even before the recent partnerships and additions; however, the additional data are enhancing the spatial coverage in various countries (via "new" stations), along with the temporal period of record for some of the stations.

The continued cooperation with our ISD partners within and outside of the FCC will provide continued daily flow of data into the system of services upon which our customers rely. Additionally, other datastreams such as CRN and the Surface Radiation Network (SURFRAD) are now being operationally integrated. Others such as the US Cooperative Network (over 8000 active stations) are planned. Since the Cooperative Network daily summary data are already online, they have a lower priority for historical integration into ISD, vs. datasets which are inaccessible or in unique native formats and more difficult to use.

3. DATA ACCESS AND PRODUCTS

ISD access and products include:

a) The NOAA Virtual Data System (NVDS) Climate Data Online (CDO) system: http://cdo.ncdc.noaa.gov (Figure 2) providing ASCII text output and 'printable' web forms.

b) NCDC's Geographic Information System (GIS) interface (Figure 3) -- click on "*search by map*" on NCDC's homepage: http://www.ncdc.noaa.gov or http://gis.ncdc.noaa.gov.

c) Data graphing for user-selected stations and periods of interest (Figure 4).

d) An ISD summary capability (Figure 5) within the CDO system and through the GIS interface. Fourteen different summary types can be generated, such as frequency distributions for cloud ceiling/visibility.

e) For U.S. stations – the Local Climatological Data (LCD) product (Figure 6): http://cdo.ncdc.noaa.gov/ulcd/ULCD.

For a web system which provides numerous details regarding ISD, its usage, links to all references, and links to all related products and services, see: http://www.ncdc.noaa.gov/oa/climate/isd/index.php.

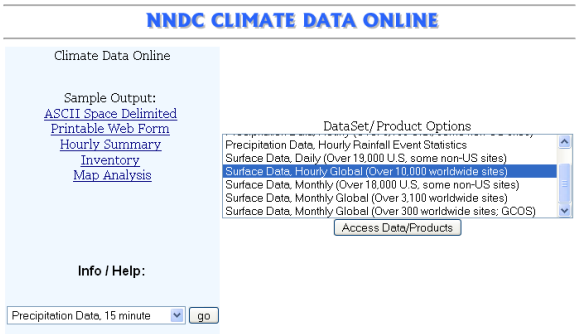


Figure 2. CDO Interface.

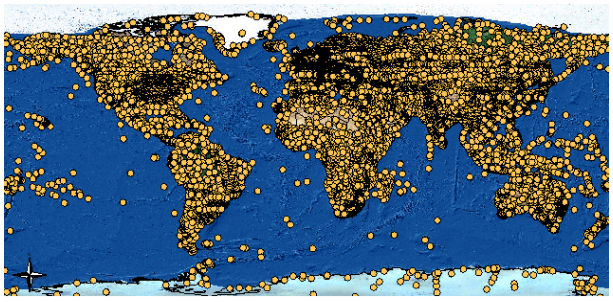


Figure 3. GIS Interface with ISD Global Map.

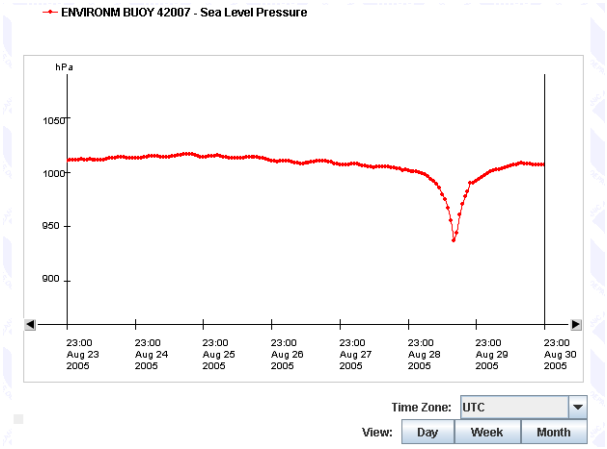
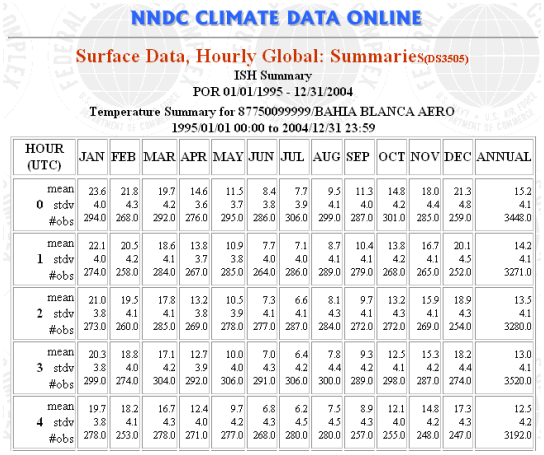


Figure 4. Graph of Sea Level Pressure During Passage of Hurricane Katrina.



NNDC CLIMATE DATA ONLINE

Surface Data, Hourly Global: Summaries(DS3505)

ISH Summary

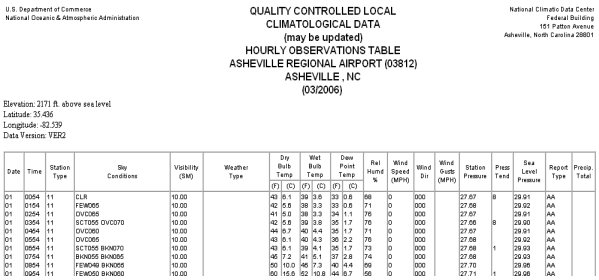
POR 01/01/1995 - 12/31/2004

Temperature Summary for 87750099999/BAHIA BLANCA AERO

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	23.6	21.8	19.7	14.6	11.5	8.4	7.7	9.5	11.3	14.8	18.0	21.3	15.2
	stdv	4.0	4.3	4.2	3.6	3.7	3.8	3.9	4.1	4.0	4.2	4.4	4.8	4.1
	#obs	294.0	268.0	292.0	276.0	295.0	286.0	306.0	299.0	287.0	301.0	285.0	259.0	3448.0
1	mean	22.1	20.5	18.6	13.8	10.9	7.7	7.1	8.7	10.4	13.8	16.7	20.1	14.2
	stdv	4.0	4.2	4.1	3.7	3.8	4.0	4.0	4.1	4.1	4.2	4.1	4.5	4.1
	#obs	274.0	258.0	284.0	267.0	285.0	264.0	286.0	289.0	279.0	268.0	265.0	252.0	3271.0
2	mean	21.0	19.5	17.8	13.2	10.5	7.3	6.6	8.1	9.7	13.2	15.9	18.9	13.5
	stdv	3.8	4.1	4.1	3.8	3.9	4.1	4.1	4.3	4.1	4.3	4.1	4.3	4.1
	#obs	273.0	260.0	285.0	269.0	278.0	277.0	287.0	284.0	272.0	272.0	269.0	254.0	3280.0
3	mean	20.3	18.8	17.1	12.7	10.0	7.0	6.4	7.8	9.3	12.5	15.3	18.2	13.0
	stdv	3.8	4.0	4.2	3.9	4.0	4.3	4.2	4.4	4.2	4.1	4.2	4.4	4.1
	#obs	299.0	274.0	304.0	292.0	306.0	291.0	306.0	300.0	289.0	298.0	287.0	274.0	3520.0
4	mean	19.7	18.2	16.7	12.4	9.7	6.8	6.2	7.5	8.9	12.1	14.8	17.3	12.5
	stdv	3.8	4.1	4.3	4.0	4.2	4.3	4.5	4.5	4.3	4.0	4.2	4.3	4.2
	#obs	278.0	253.0	278.0	271.0	277.0	268.0	280.0	280.0	257.0	255.0	248.0	247.0	3192.0

Figure 5. ISD Summary Sample.



U.S. Department of Commerce
National Oceanic & Atmospheric Administration

QUALITY CONTROLLED LOCAL CLIMATOLOGICAL DATA
(may be updated)
HOURLY OBSERVATIONS TABLE
ASHEVILLE REGIONAL AIRPORT (03812)
ASHEVILLE , NC
(03/2006)

National Climatic Data Center
Federal Building
151 Patton Avenue
Asheville, North Carolina 28801

Elevation: 2171 ft. above sea level
Latitude: 35.436
Longitude: -82.539
Data Version: VER2

Date	Time	Station Type	Sky Conditions	Visibility (SM)	Weather Type	Dry Bulb Temp (F)	Dry Bulb Temp (C)	Wet Bulb Temp (F)	Wet Bulb Temp (C)	Dew Point Temp (F)	Dew Point Temp (C)	Rel Humd %	Wind Speed (MPH)	Wind Dir	Wind Gusts (MPH)	Station Pressure	Press Tend	Sea Level Pressure	Report Type	Precip. Total
01	0054	11	CLR	10.00		43	6.1	39	3.6	33	0.8	69	0	000		27.67	0	29.91	AA	
01	0154	11	FEW065	10.00		42	5.6	38	3.3	33	0.6	71	0	000		27.68		29.92	AA	
01	0254	11	OVC065	10.00		41	5.0	38	3.3	34	1.1	76	0	000		27.67		29.91	AA	
01	0354	11	SCT055 OVC070	10.00		42	5.6	39	3.8	35	1.7	76	0	000		27.66	8	29.90	AA	
01	0454	11	OVC060	10.00		44	6.7	40	4.4	35	1.7	71	0	000		27.67		29.91	AA	
01	0554	11	OVC065	10.00		43	6.1	40	4.3	36	2.2	76	0	000		27.68		29.92	AA	
01	0654	11	SCT055 BKN070	10.00		43	6.1	39	4.1	35	1.7	73	0	000		27.68	1	29.93	AA	
01	0754	11	BKN055 BKN065	10.00		45	7.2	41	5.1	37	2.8	74	0	000		27.68		29.93	AA	
01	0854	11	FEW049 BKN055	10.00		50	10.0	46	7.3	40	4.4	69	0	000		27.70		29.95	AA	
01	0954	11	FEW050 BKN060	10.00		60	15.5	52	10.8	44	6.7	56	0	000		27.71	1	29.96	AA	

Figure 6. LCD Product Sample.

4. BENEFITS OF INTEGRATION

In conclusion, benefits of integrating various surface climatological datasets include:

- Reduction of subjectivity and inconsistencies among datasets that span multiple observing networks and platforms.
- Standardized QA/QC based on reporting time resolution (e.g., QC methodology for hourly temperature data independent of network).
- Products which are more easily developed, consistent, and improved by collective experience and expertise.
- Modular software which is much more extensible.
- Data documentation which is consistent and applies to the full period of record for the data.
- Simplified portability of data into reanalysis efforts and climate change studies.

5. ACKNOWLEDGEMENTS

Numerous individuals have contributed to the development and expansion of ISD and to its data access and product suite. In addition to the co-authors, this includes Xungang Yin and Brian May (STG Inc); Fred Smith, Pete Jones, and Richard Smith (GST Inc); Glen Reid and Scott Chapal (IMSG Inc); Tom Whitehurst, Mark Lackey, Rich Baldwin, Dee Dee Anders, Ron Ray, Dan Dellinger, Mike Urzen, Alan Hall, and Mark Seiderman (NCDC); Kathy Hawkins (formerly NCDC); Keith Eggleston (NERCC); and Dan Graybeal (formerly NERCC).

6. REFERENCES

Lott, Neal. The Quality Control of the Integrated Surface Hourly Database. 84th American Meteorological Society Annual Meeting, 2004, Seattle, WA.

Lott, Neal, R. Baldwin, and P. Jones. NCDC Technical Report 2001-01, The FCC Integrated Surface Hourly Database, A New Resource of Global Climate Data. [Asheville, N.C.]: National Climatic Data Center, 2001.

Del Greco, S.A., N. Lott, R. Ray, D. Dellinger, P. Jones, and F. Smith. Surface Data Processing and Integration at NOAA's National Climatic Data Center. 87th American Meteorological Society Annual Meeting, 2007, San Antonio, TX.

For additional references and information, see: http://www.ncdc.noaa.gov/oa/climate/isd/index.php