

ENGAGING THE PUBLIC IN CLIMATE SCIENCE

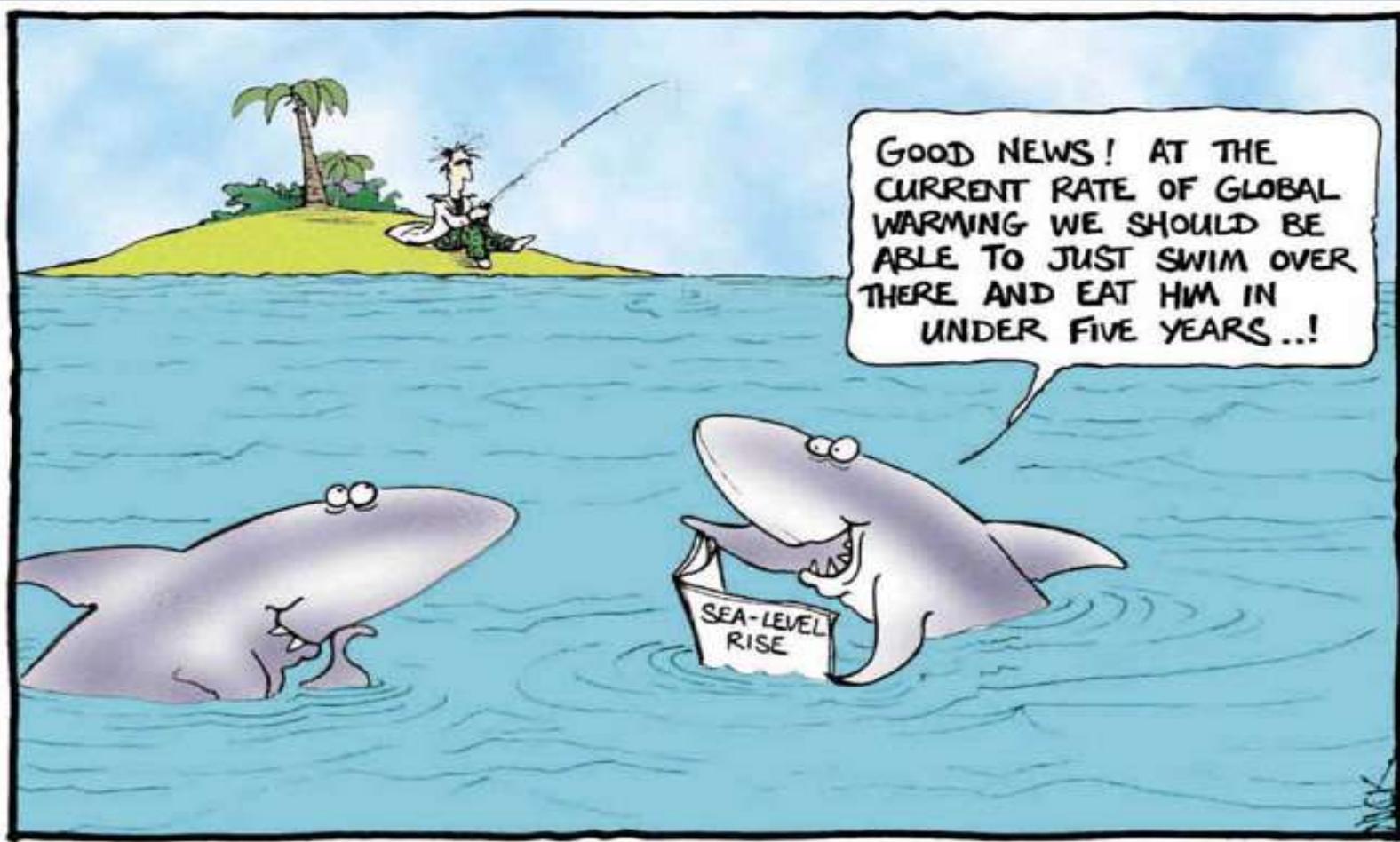
EXPLOITING CROWDSOURCING TO
DIGITIZE AND ANALYZE CLIMATE DATA



Scott Hausman
Deputy Director
National Climatic Data Center

OUTLINE

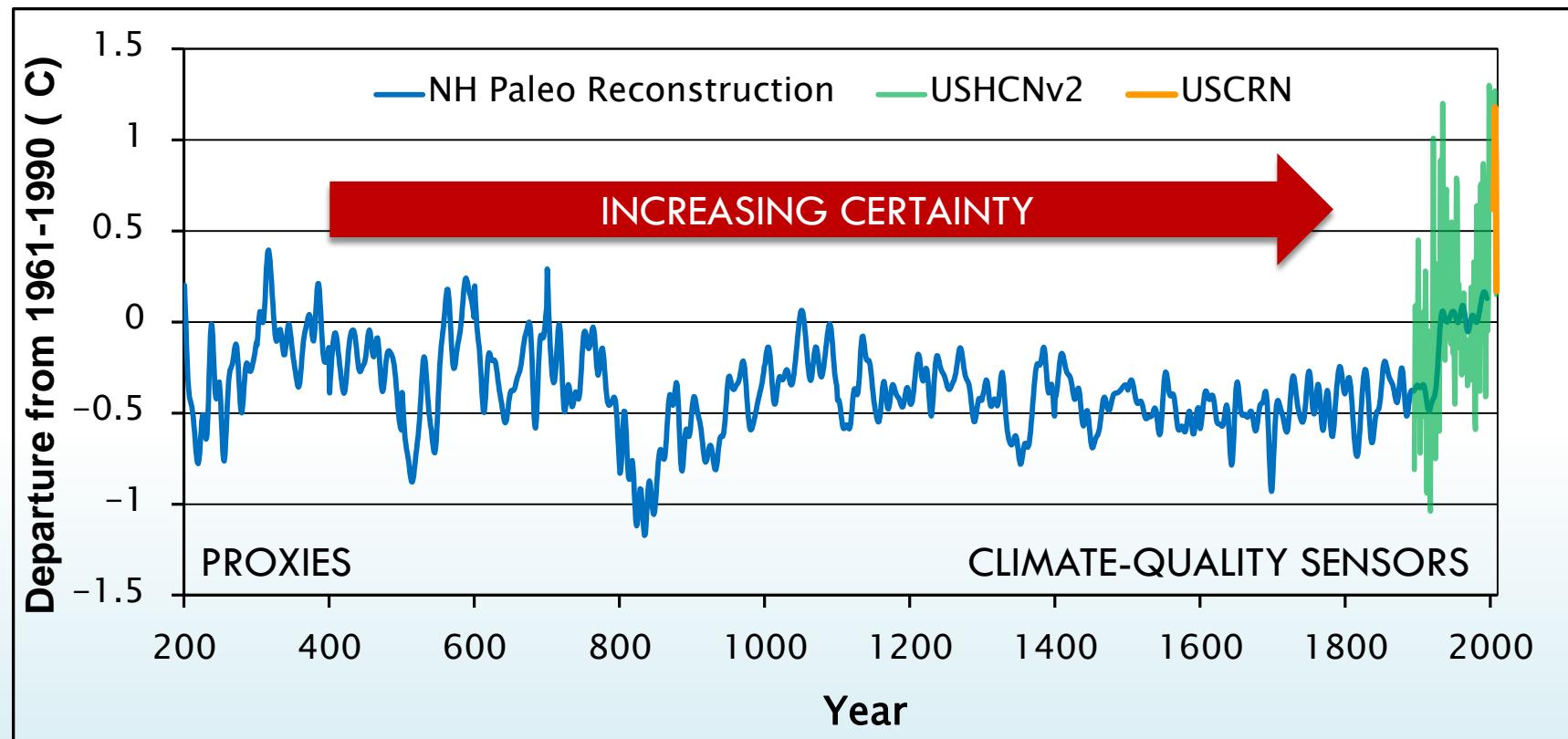
- Challenges of Monitoring Climate Change
- Recruiting Internet Citizen Scientists
- Partnering with Citizen Science Alliance
 - Project 1: Data Rescue for Surface Temperature Databank
 - Project 2: Tropical Cyclone Reanalysis



CHALLENGES OF MONITORING CLIMATE CHANGE

Reducing Uncertainty to Improve Public Understanding

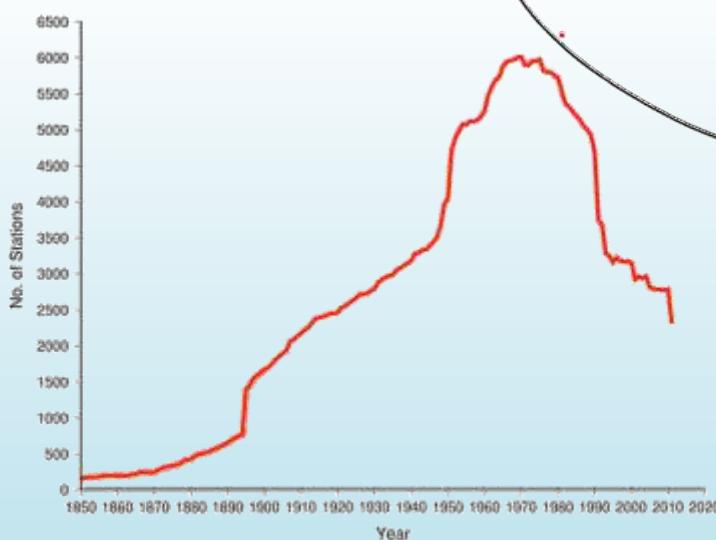
OBSERVATION QUALITY AND UNCERTAINTY



Three temperature time series are displayed relative to the 1961-1990 normals (°C). The paleoclimate reconstruction from Mann et al. (2008) is for Northern Hemisphere land, while U.S. Historical Climatology Network v. 2 dataset from Menne et al. (2009) and U.S. Climate Reference Network temperature departures are for the continental U.S. only. A major goal is to provide consistent time series across paleoclimate and instrumental networks.

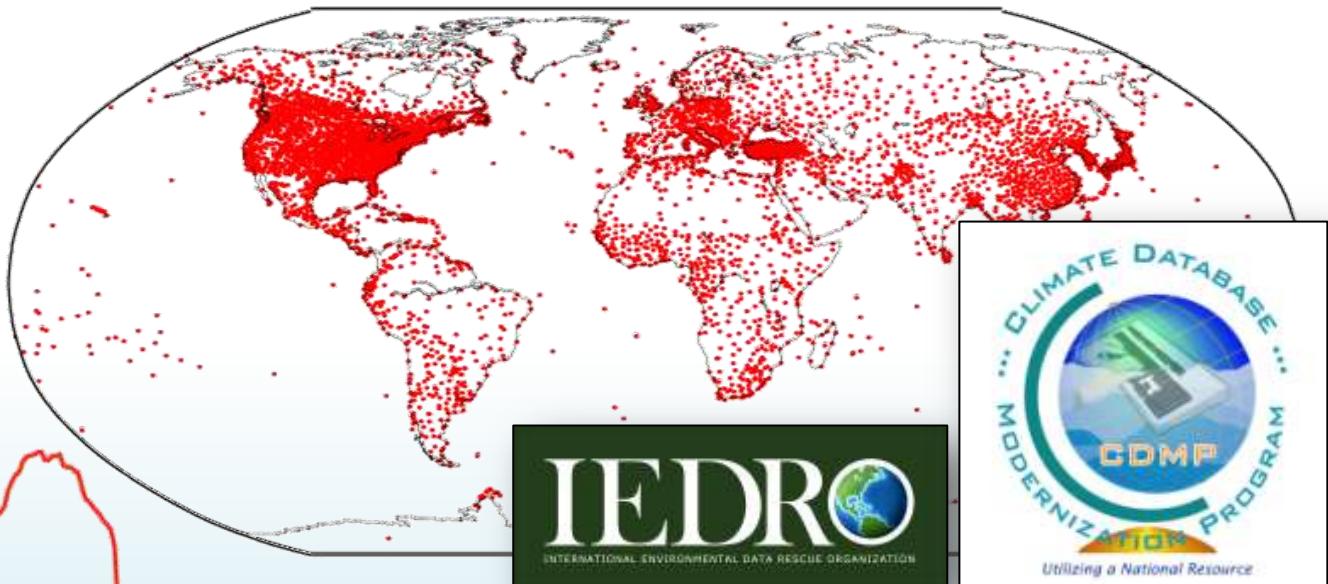
FILLING SPATIAL GAPS FOR REGIONAL CHANGE DETECTION

- GHCNv3 has about 2700 reliable stations
- Rapid increase during WWII
- Gradual decrease in 80s, 90s as stations close



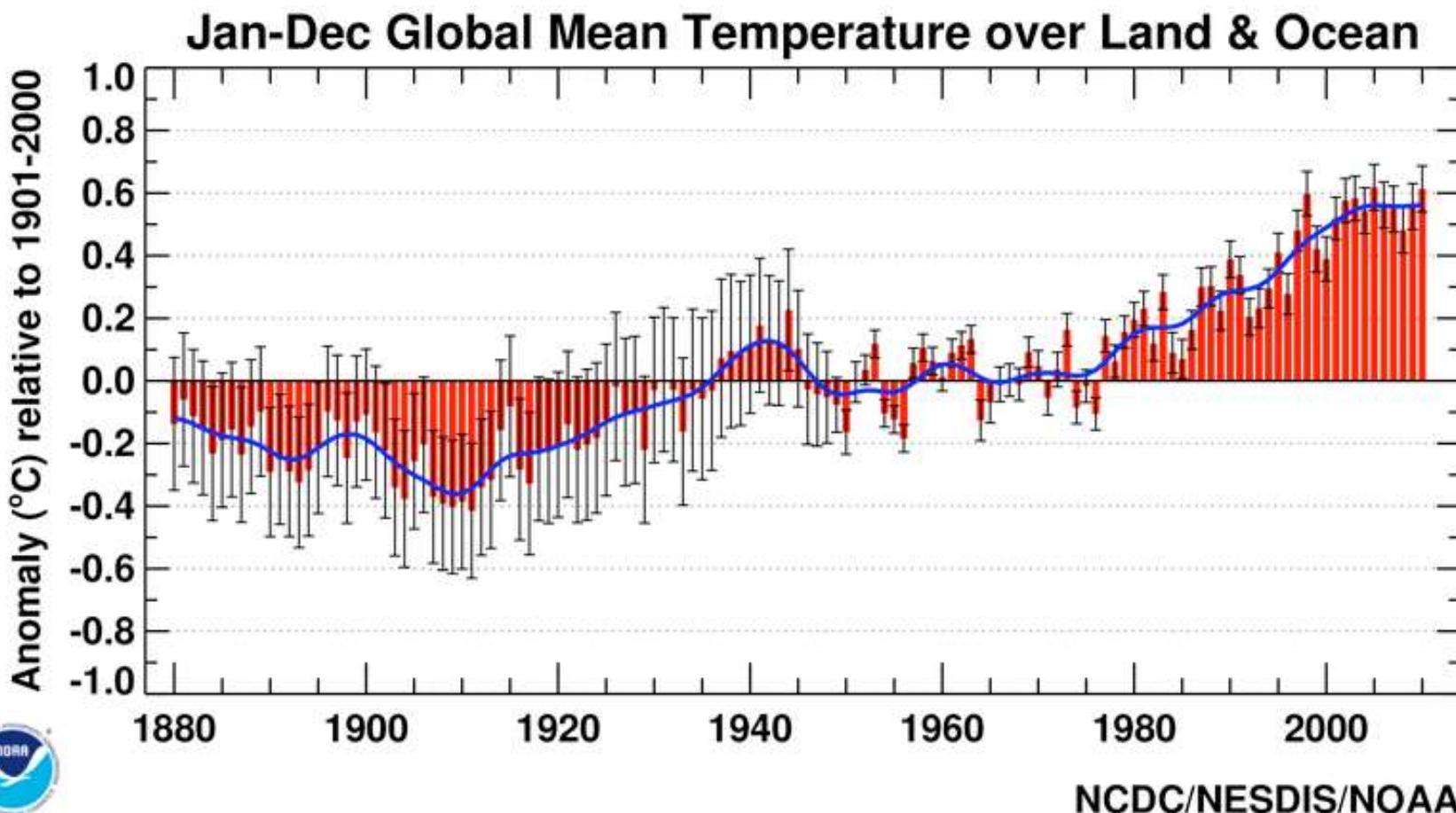
Global Historical Climatology Network (Monthly)

Stations Reporting Mean Temperature, 1961-1990



- Some data still not in digital form
- Rescuing data has been first priority
- Critical to understanding regional climate change.

REDUCING UNCERTAINTY IN GLOBAL TRENDS



More observations needed to reduce uncertainty in historical record.

MAKING DATA ACCESSIBLE FOR DISCOVERY & ANALYSIS

- NOAA On-line Foreign Data Library
 - Over 75 countries and former colonies
 - Covers from 1830 to 1970s
- Digitally imaged through Climate Database Modernization Program(CDMP)
 - Millions of images; more accessible
 - Data is still essentially
- Additional 2000+ boxes of international data in NCDC physical archives
 - Similar numbers of data held elsewhere
 - None of it is usable without digitization
- http://docs.lib.noaa.gov/rescue/data_rescue_home.html

DATE.	BAROMÈTRE (700-+).			THERMOMÈTRE.				HUMIDITÉ RELATIVE.			VENT.			SÉBULOSITÉ.			PLUIE.			REMARQUES.
	7 ^h	1 ^h	7 ^h	7 ^h	1 ^h	7 ^h	Min.	Max.	7 ^h	1 ^h	7 ^h	7 ^h	1 ^h	7 ^h	7 ^h	1 ^h	7 ^h	7 ^h	1 ^h	
1	22,4	19,4	20,6	21,6	31,1	28,0	18,4	35,0	89	36	56	Calm	ENE	5	Calm	6	0	4	0	0
2	23,0	21,5	20,4	22,6	32,4	24,6	19,0	34,4	86	31	51	Calm	NNE	3	NNE	6	0	6	0	0
3	20,9	19,5	19,2	21,4	31,0	27,3	18,6	36,8	87	36	53	NNE	1	3	ENE	3	2	6	6	0
4	20,1	19,8	21,7	20,0	38,0	25,6	18,2	29,8	91	47	57	Calm	NE	2	SNE	1	1	8	6	0,0
5	23,0	21,4	23,4	23,4	31,8	23,0	16,0	34,6	86	39	54	Calm	ENE	4	ENE	2	0	0	0	0

LIMITS TO OPTICAL CHARACTER RECOGNITION (OCR)

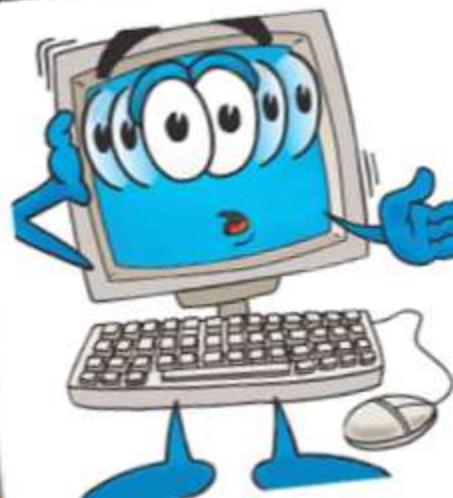
Time entries on this form are 75 th meridian time
To convert to G. c. t. add 5 hours
Height of barometer 16 feet (MSL)

12868

TYPE	TIME (LST)	CLOUDS (Thousands of feet) AND DEG	VISIBILITY (miles)	WEATHER AND OBSTRUCTIONS IN VISION	SEA LEVEL PRESS. (in.)	TEMP. (°F.)	DEW PT.
						(H)	(D)
R	0058	70(D)	10		207	74	7
R	0159	18(D)	10		193	71	76
R	0259	19(D)	10		193	72	76
R	0359	20(D)	10		193	72	70
S1	0459	1(D)	10		193	71	70
R	0459	1(D)	10		193	71	69
S2	0530	1(D)	10		196	70	68
R	0559	1(D)	10		203	69	67
S3	1630	40(D)	10		210	70	68
S4	0645	E35(D)	10				
R	0658	325(D)	10		213	75	71
S5	1722	1425(D)	10		217	71	71

Diary of the Weather, Fort Amst

Date	Barometer		Clouds		Cause of Weather	Wind	W.
	7	8	7	8			
1720	7	8	7	8	Altitude		
July 1	30	96	31	58	Fair	22	SW
2	71	96	31	58	W.	24	SW
3	30	86	71	58	Cloudy	5	SW
4	16	81	72	57	W.	24	SW
5	70	82	78	47	Fair	2	SW





“On the Internet, nobody knows you’re a dog.”

©The New Yorker Collection 1993 Peter Steiner
From cartoonbank.com. All rights reserved.

RECRUITING INTERNET CITIZEN SCIENTISTS

We need carbon-based computing for imagery analysis.

EXPLOITING CROWDSOURCING TO DIGITIZE AND ANALYZE CLIMATE
DATA

CITIZEN SCIENCE AND CROWDSOURCING

- Exploits the cognitive abilities of Human Computation!
- Novel mode of data collection:
 - Citizen Science = Volunteer Science = Participatory Science
 - e.g., VGI = Volunteer Geographic Information (Goodchild '07)
 - e.g., Galaxy Zoo @ <http://www.galaxyzoo.org/>
- Citizen science refers to the involvement of volunteer non-professionals in the research enterprise.
- The Citizen Science experience ...
 - must be engaging,
 - must work with real scientific data/information,
 - must not be busy-work (all clicks must count),
 - must address authentic science research questions that are beyond the capacity of science teams and enterprises, and
 - must involve the scientists.

Reference: Kirk Borne, Reference: Reaching Out with Eventful Astronomy, George Mason University

EXAMPLE: RECAPTCHA



reCAPTCHA IS A FREE
ANTI-BOT SERVICE THAT
HELPS DIGITIZE BOOKS.

steamboat train, from New
this morning ran off the track
New-London. Four cars plunged



→ LEARN HOW reCAPTCHA WORKS

USE reCAPTCHA ON YOUR SITE

- 🔒 STRONG SECURITY
- 🔊 ACCESSIBLE TO BLIND USERS
- 📊 30+ MILLION SERVED DAILY

NEW See how accurate reCAPTCHA is at
digitizing content!

[Blog](#) | [About Us](#) | [Contact](#) | [FAQs](#) | [Terms](#) | [Privacy](#)
© 2011 Google, all rights reserved.

Scanned type

This aged portion of society were distinguished from

OCR reads as

"niis aged pntkm at society were distinguished frow."

EXAMPLE: ZOONIVERSE



About the Zooniverse

The Zooniverse is home to the internet's largest, most popular and most successful citizen science projects. Our current projects [are here](#) but plenty more are on the way. If you're new to the Zooniverse, we suggest picking a project and diving in - the same account will get you into all of our projects, and you can keep track of what you've contributed by watching 'My Zooniverse'.

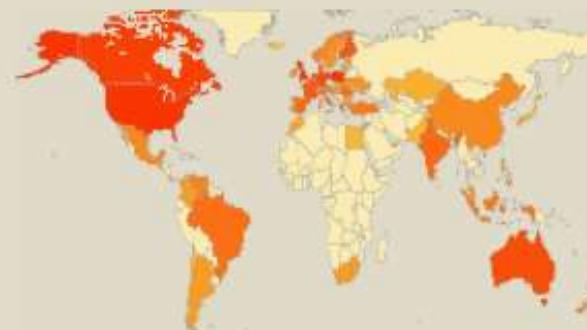
The Zooniverse and the suite of projects it contains is produced, maintained and developed by the [Citizen Science Alliance](#). The [member institutions](#) of the CSA work with many academic and other partners to harness the efforts and ability of volunteers to help solve data that confronts them.

The Zooniverse began with a single project in 2007. The Galaxy Zoo team had expected to be overwhelmed by the response to the project, but, buckling under the strain, they set about

Galaxy Zoo was important because it produced many unique scientific results, ranging from those using classifications that depended on human judgment to those that did not. This commitment to producing real results with your time - is at the heart of everything we do.

Zooniverse Activity

Total Volunteers: **430,804**



Recent Zooniverse activity

Live Projects

[planethunters.org](#)

[THE MILKY WAY PROJECT](#)

[MOON ZOO](#)

[GALAXY ZOO HUBBLE](#)

[oldWeather](#)

[SOLAR STORMWATCH](#)

[GALAXY ZOO - UNDERSTANDING COSMIC MERGERS](#)

[GALAXY ZOO - THE HUNT FOR SUPERNOVAE](#)

[ZOOSHOP](#)

WEAR THE BADGE
SPREAD THE WORD

[GO SHOPPING](#)



EXAMPLE: GALAZXYZOO.ORG

The screenshot shows the Galaxy Zoo website interface. At the top, the logo 'GALAXY ZOO' is displayed with a small image of a galaxy. A user icon with the number '2' is on the right. Below the logo is a navigation bar with links: Home, How To Take Part, My Galaxies, Contact Us, Profile, and Logout. The main content area on the left shows a large image of a galaxy with a bright central bulge and a smaller red star-like object in the lower left. Below the image are two buttons: 'INVERT GALAXY IMAGE' and 'ADD TO MY FAVOURITES'. The right side of the screen is a classification task titled 'Classify Galaxies'. It asks the question 'Is the galaxy simply smooth and rounded, with no sign of a disk?'. Three options are provided: 'Smooth' (a white circle), 'Features or disk' (a spiral galaxy icon), and 'Star or artifact' (a red 'X'). A 'Need help?' button with a question mark icon is at the bottom right of the classification box.

- ~260,000 participants (and growing)
- ~1 million galaxies have been labeled (classified)
- ~180 million classifications have been collected

EXAMPLE: OLDWEATHER.ORG

The screenshot shows the OldWeather.org homepage. At the top, there is a navigation bar with links: HOME, VESSELS, TUTORIAL, TRANSCRIBE, ABOUT, BLOG, FORUM, and a blue button labeled GET STARTED. Below the navigation bar are three icons: 1. Follow vessels (a ship icon), 2. Digitise pages (a person at a computer icon), and 3. Get promoted (a sailor with a raven icon). The main content area features a large image of a steamship. A red arrow points from the bottom of the ship's hull area to a small red icon of a ship. To the right, a box displays project statistics: **68 %** OF THE LOGS COMPLETED, 517217 PAGES DONE, and 118 SHIPS COMPLETE. The bottom of the page has a red banner with the text: Old Weather: Our Weather's Past, the Climate's Future. The banner also includes links for Google and Terms of Use.

oldWeather

HOME VESSELS TUTORIAL TRANSCRIBE ABOUT BLOG FORUM GET STARTED

1. Follow vessels 2. Digitise pages 3. Get promoted

Project Statistics

Old Weather transcriptions so far

68 % OF THE LOGS COMPLETED

517217 PAGES DONE

118 SHIPS COMPLETE

Old Weather: Our Weather's Past, the Climate's Future

EXAMPLE: OLDWEATHER

Cadet hausmans

0 weather reports on 0 pages contributed to this voyage. 30 weather reports more for promotion to Lieutenant

2. LOCATION

4. OTHER

5. FINISH

H.M.S. "Teutonic" Wednesday 14th day of April 1915

From	To	or At	Sea						
Patent Log	Distance Run	Standard Compass	Wind	Heights of Sun	Height of Barometer and Attached Thermometer	Temperature	Position	Latitude	Longitude
Enter date	Enter distance	Enter standard compass	Enter wind	Enter heights	Enter barometer and attached thermometer	Enter temperature	8.0 a.m.	60° 14' N	5° 59' W
							8.0 p.m.	63° 26' N	5° 52' W
							8.0 p.m.	63° 26' N	5° 52' W

Enter a weather reading

Date	Location	Weather	Event
1	Wind	Force	Code
2	WSW	5	COP
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			



"It's black, and it looks like a hole.
I'd say it's a black hole."

PARTNERING WITH CITIZEN SCIENCE ALLIANCE

Lessons learned from astronomy applied to climatology.

INTERNATIONAL CROWDSOURCING COLLABORATION

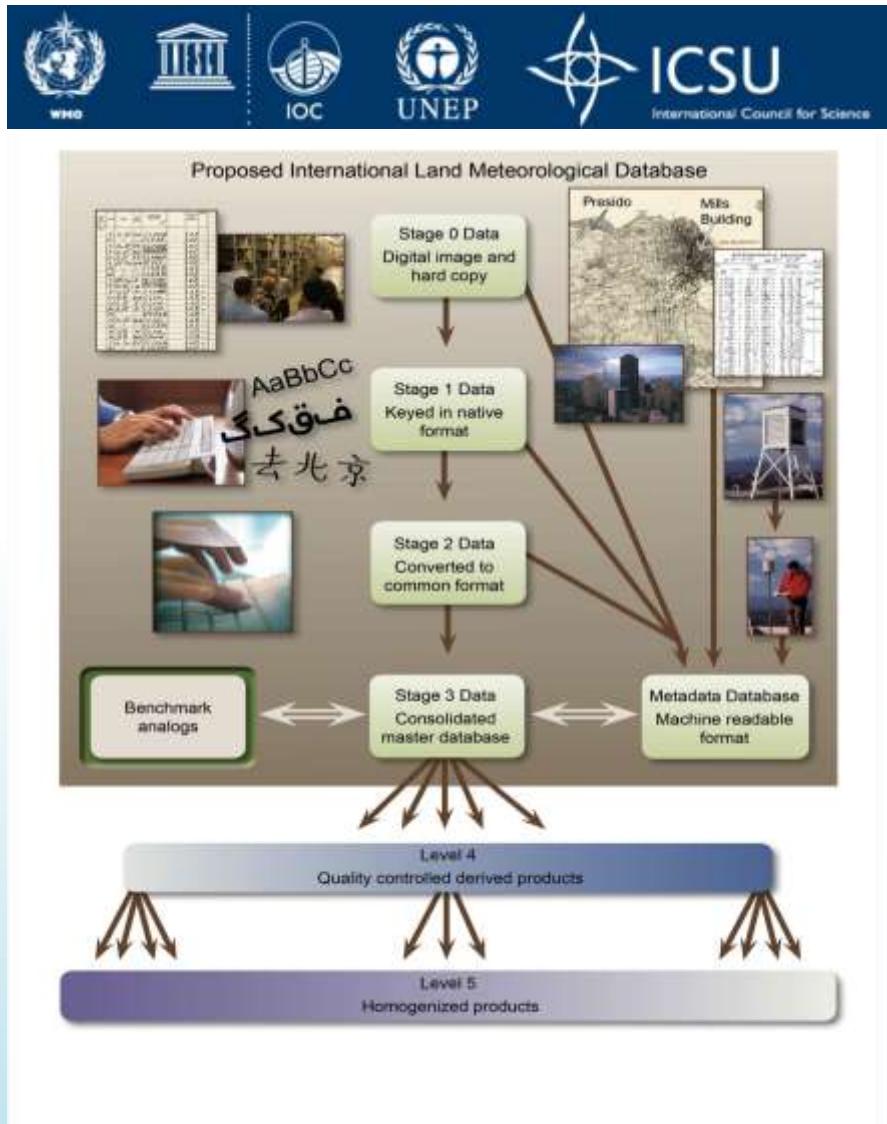
- Establishing a long-term partnership with Citizen Science Alliance (Zooniverse)
 - Effort led by scientists from the Cooperative Institute for Climate and Satellites in North Carolina (CICS-NC)
- In short-term, developing prototype capabilities
 - International exchange of scientific expertise and technology transfer
- Long-term goal of joining the CSA and developing climate crowdsourcing applications
- Two initial projects are proposed
 - Date Rescue for Surface Temperature Databank
 - Tropical Cyclone Reanalysis



PROJECT 1: DATA RESCUE FOR SURFACE TEMPERATURE DATABANK

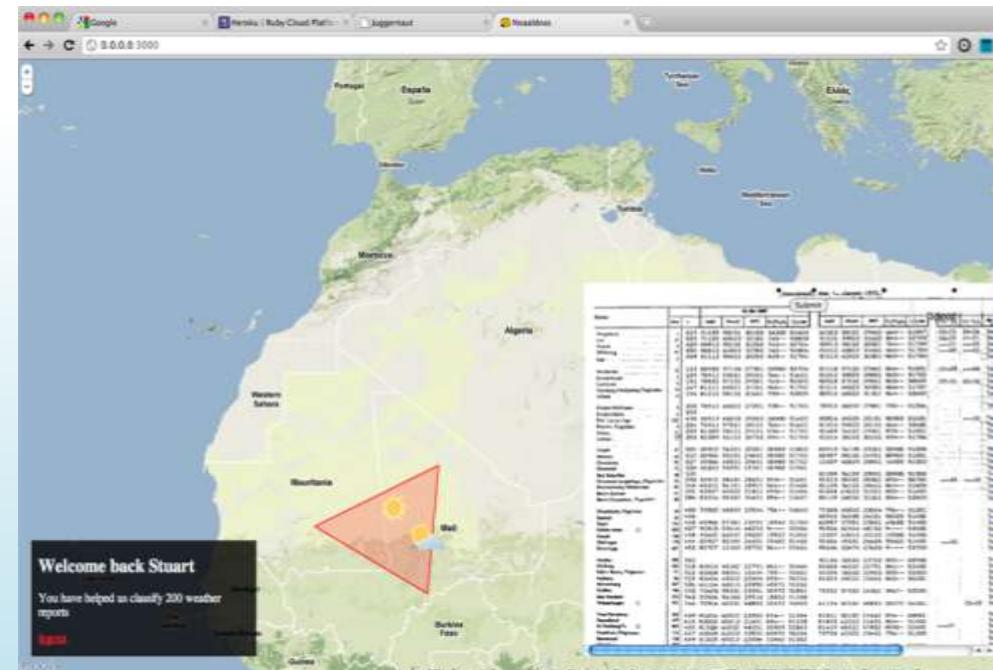
GLOBAL DATABANK

- Society expects openness and transparency in the understanding of the (un)certainty on how climate has changed and how it will continue to change
- UK Meteorological Office (UKMO) proposed a new International Analysis of Land-Surface Air Temperature Data
 - Endorsed by World Meteorological Organization (WMO) Commission for Climatology, February 2010
- *International Surface Temperature Initiative*
 - Established at workshop in Exeter, UK in September, 2010
 - NOAA/NCDC is leading in establishing and potentially hosting the *data bank*
 - Data rescue is a priority of the initiative
- <http://www.surfacetemperatures.org>



DATA DIGITIZATION THROUGH CROWDSOURCING

- Proposed method similar to oldWeather.org
 - Multiple redundant keying of historical images
 - Similar to ReCAPTCHA, use human to refine OCR results
- Digitized data placed in The Databank
 - Retain the redundant values as valuable metadata
 - Full provenance and version tracking – anyone can go back to the raw data value at any time
- Finding the right motivation is challenge
 - Citizen scientist must understand the impact they're making



PROJECT 2: TROPICAL CYCLONE REANALYSIS

TROPICAL CYCLONE REANALYSIS

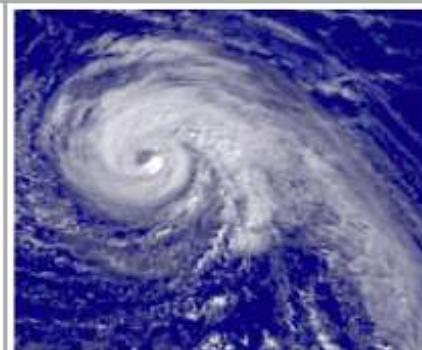
- Historical intensity records of tropical storms are based on regional methods, leading to basin-to-basin differences
 - Methods have also changed with time and even differ by forecaster
- Since the late-1970s we have archived satellite images
- Use the 'crowd' to create a consistently analyzed historical record across the globe.
- Dvorak (intensity) classification is well-suited because it can be easily translated to a GalaxyZoo-type analysis and is the worldwide standard



Tropical Storm Wilma at T3.0



Tropical Storm Dennis at T4.0



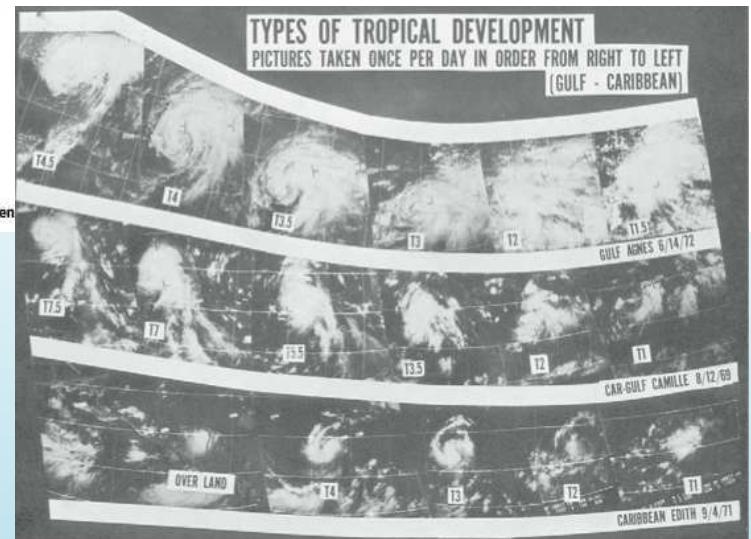
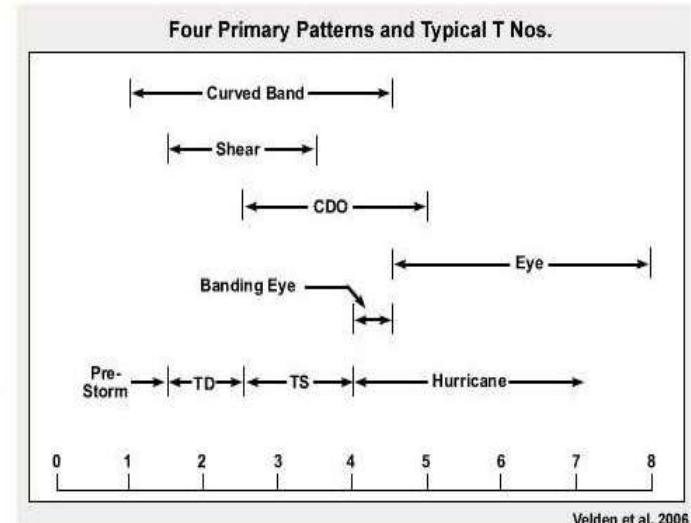
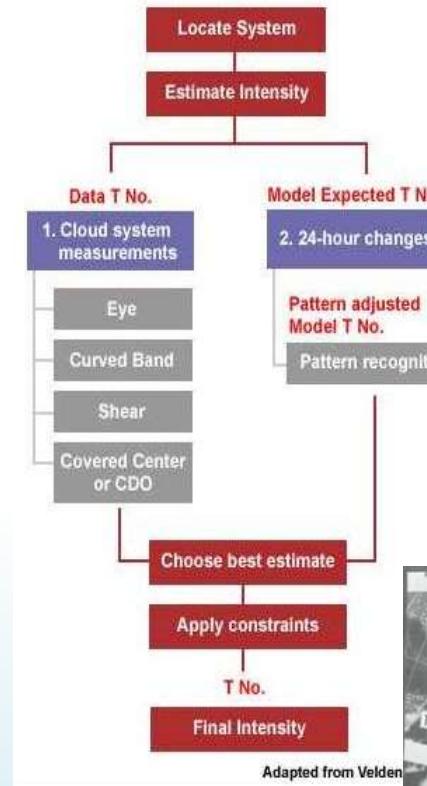
Hurricane Jeanne at T5.0



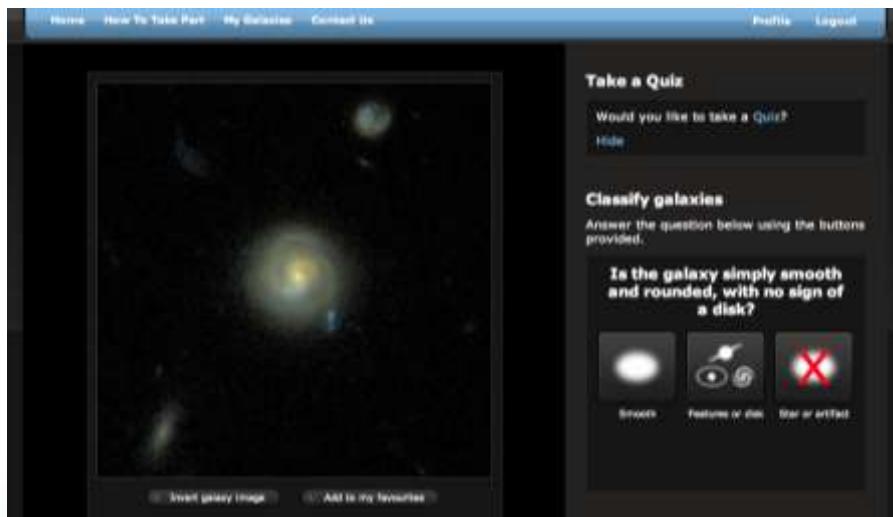
Hurricane Emily at T6.0

DVORAK TECHNIQUE

- Subjective estimate of tropical cyclone (TC) intensity based solely on visible and infrared satellite images
- TCs of similar intensity tend to have certain characteristic features, and as they strengthen, they tend to change in appearance in a predictable manner
- A "T-number" and a Current Intensity (CI) value are assigned to the storm
 - 1 - minimum intensity
 - 8 - maximum intensity
- Developed in 1973 by Vernon Dvorak



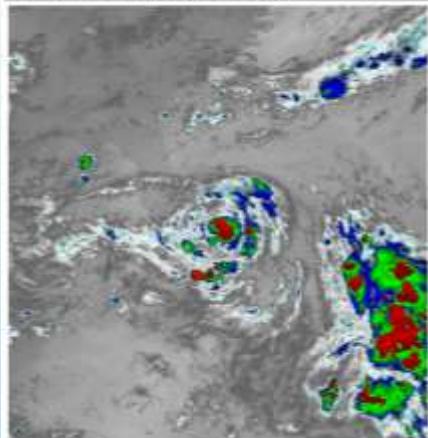
Tropical Cyclone Reanalysis Prototype



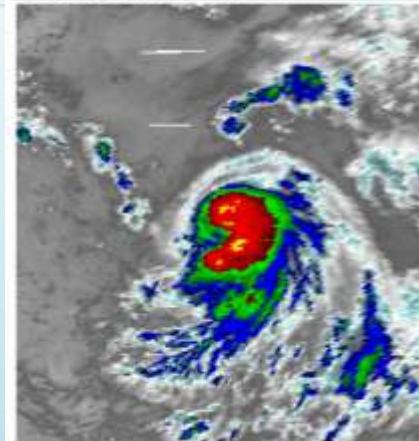
The Galaxy zoo community has provided over 250 million classifications through the galaxy zoo website. Hurricanes like galaxies are still best classified by humans, the citizen science community can produce an equally valuable dataset for meteorologists and climatologists to help them understand these extreme events

PROTOTYPE OF HURRICANE INTERFACE

Storm classification



Storm classification



SUMMARY

- Critical environmental data has been digitally “rescued” for long-term preservation, but essentially remains “lost” to scientific inquiry
- Crowdsourcing offers tremendous potential to not only leverage online “cranial capacity”, but also to engage the public in science
- NOAA is collaborating with Citizen Science Alliance, which is leading the way in crowdsourcing, to fill climate data voids and enable climate science

Scott Hausman

Deputy Director

NOAA's National Climatic Data Center (NCDC)

151 Patton Avenue, Room 557

Asheville, NC 28807-5002

 828-271-4848

 828-271-4246

 828-450-9188

Scott.Hausman@noaa.gov

www.ncdc.noaa.gov

