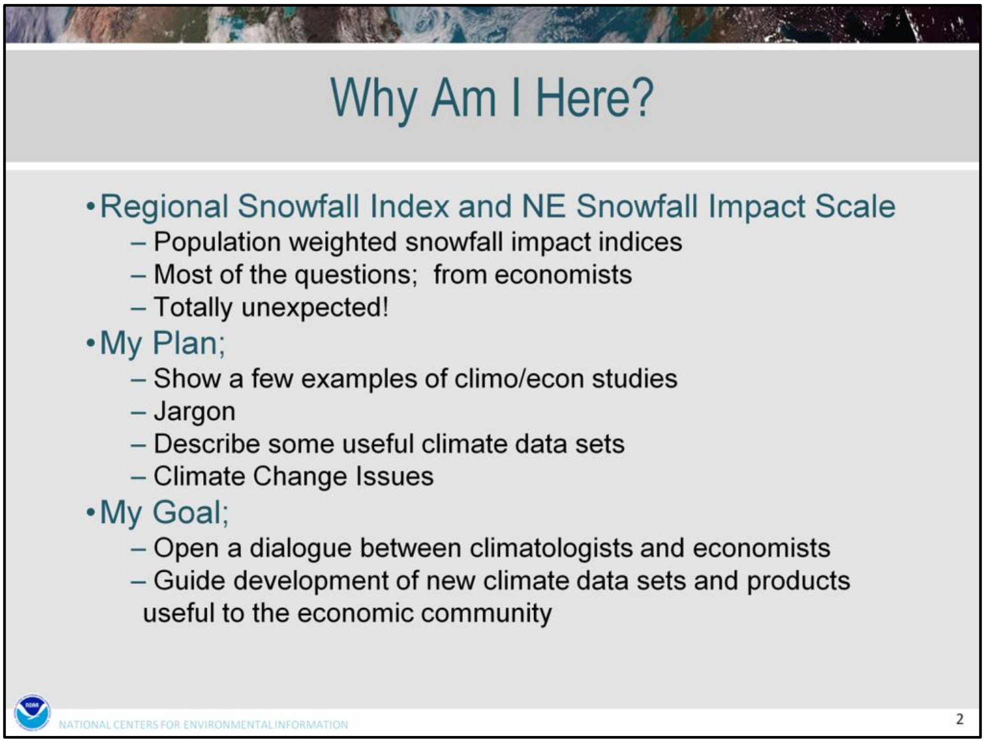
Weather and Climate Data Sets Useful for Economic Modeling

Michael Squires
NOAA's National Centers for Environmental Information

2016 American Economic Association Annual Meeting
Weather and United States Economic Activity
3 January 2016

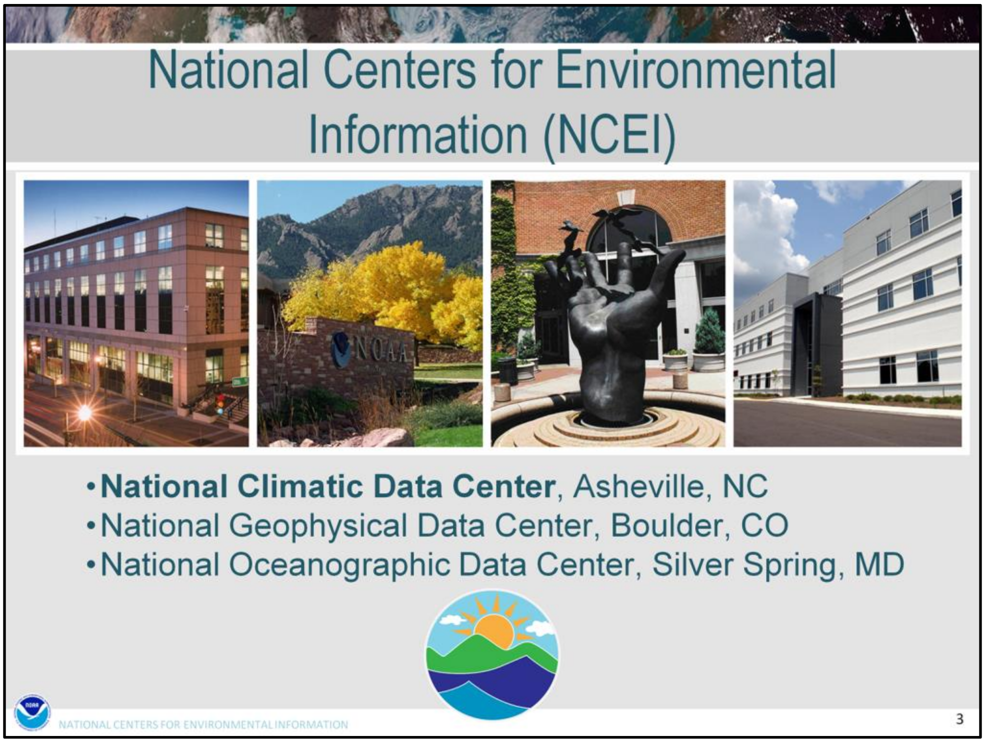
NOAA Satellite and Information Service | National Centers for Environmental Information



So why is a climatologist speaking at a meeting for economists? It begins with the Regional Snowfall Index (RSI) and the Northeast Snowfall Impact Scale (NESIS). The National Centers for Environmental Information (NCEI) has been producing both of these products operationally for several years. One would expect that most of the questions received about these products would come from the emergency management community or the media. But most of the questions have come from economists including the President's Council of Economic Advisors, Federal Reserve Board/Banks, and numerous financial firms. In the process of answering these questions I have become connected with several economist which led to me being invited here.

My plan is to show a few examples from the economic literature that relate to weather and climate. I also want to discuss some of the jargon both disciplines use to help facilitate communication. A large part of the presentation will list and summarize data sets and products that I believe could be useful to economists. I will also describe some climate change issues and discuss some implications for economists.

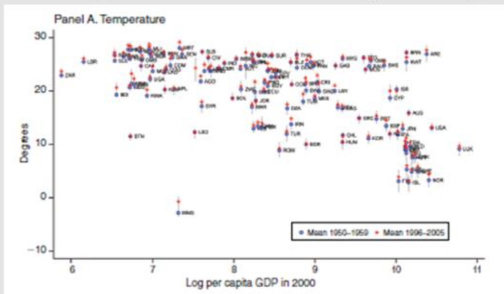
My primary goal is to open a dialogue between the economic community and NCEI. We are striving to produced user inspired data sets and products that are authoritative and transparent in their methodology. We have not really had much interaction with the economic community before, so I see this as a great engagement opportunity.



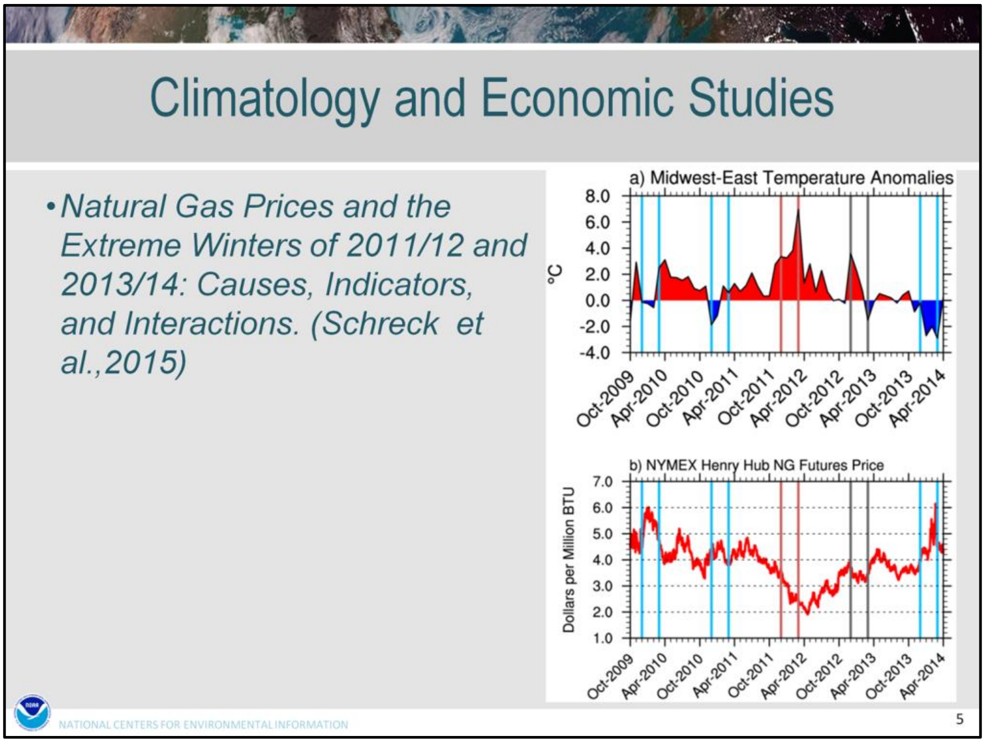
Before I begin, I need to say a few words about NCEI. We were formerly known as the National Climatic Data Center (NCDC). However, last year the National Climatic Data Center, National Geophysical data Center, and the National Oceanographic Data Center merged into NCEI. The headquarters for the NCEI is in Asheville, NC. Much of NCDC is now part of the Center for Weather and Climate (CWC), also located in Asheville.

Climatology and Economic Studies

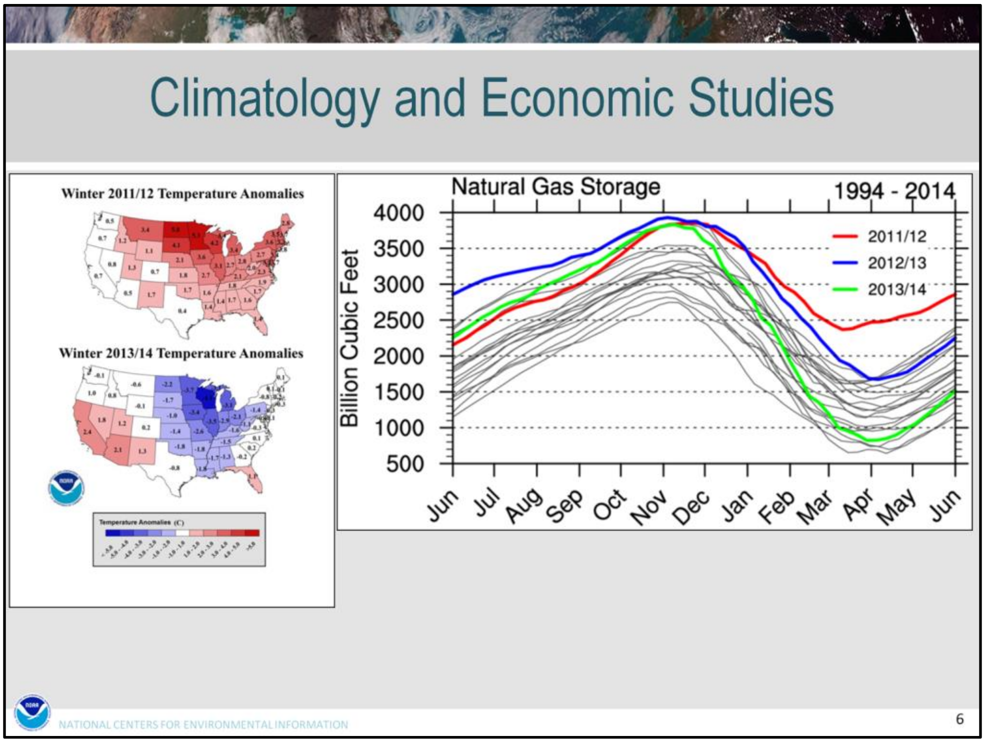
- *Temperature Shocks and Economic Growth: Evidence from the Last Half Century (Dell et al., 2012)*
  - Higher temperatures
    - reduce economic growth in poor countries
    - may reduce growth rates
    - reduce agricultural/industrial output and political stability



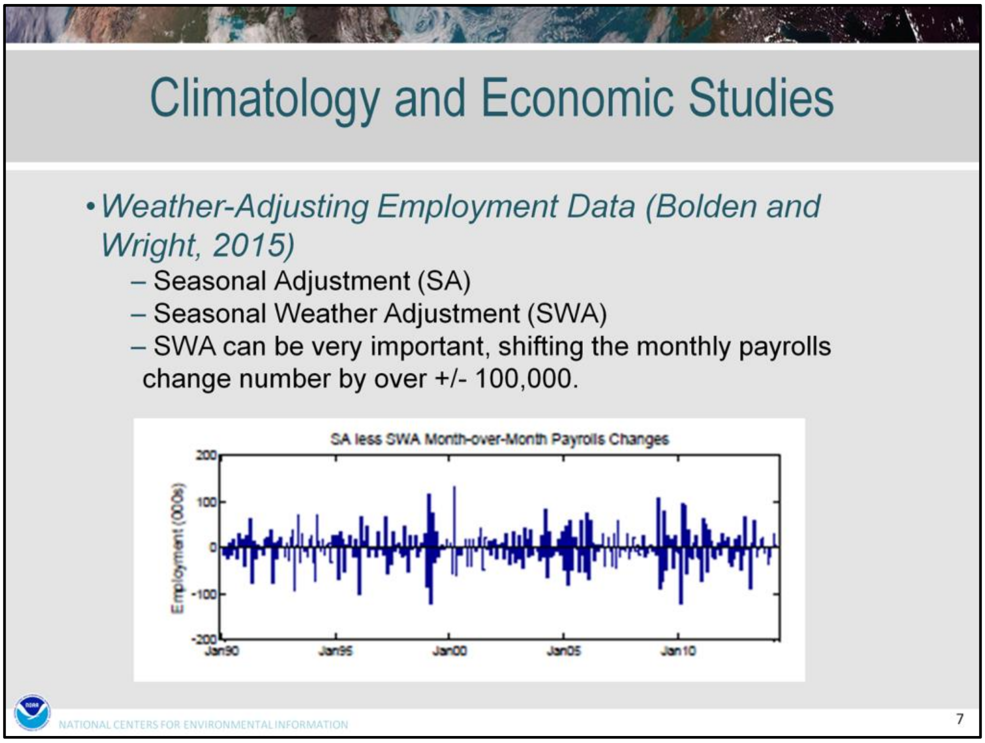
The first example from the climate-economic literature is about temperature shocks and economic growth. It concludes that higher temperatures lead to; reduced economic growth, may reduce growth rates, reductions in agricultural and industrial output, as well as reduced political stability (Dell et al., 2012).



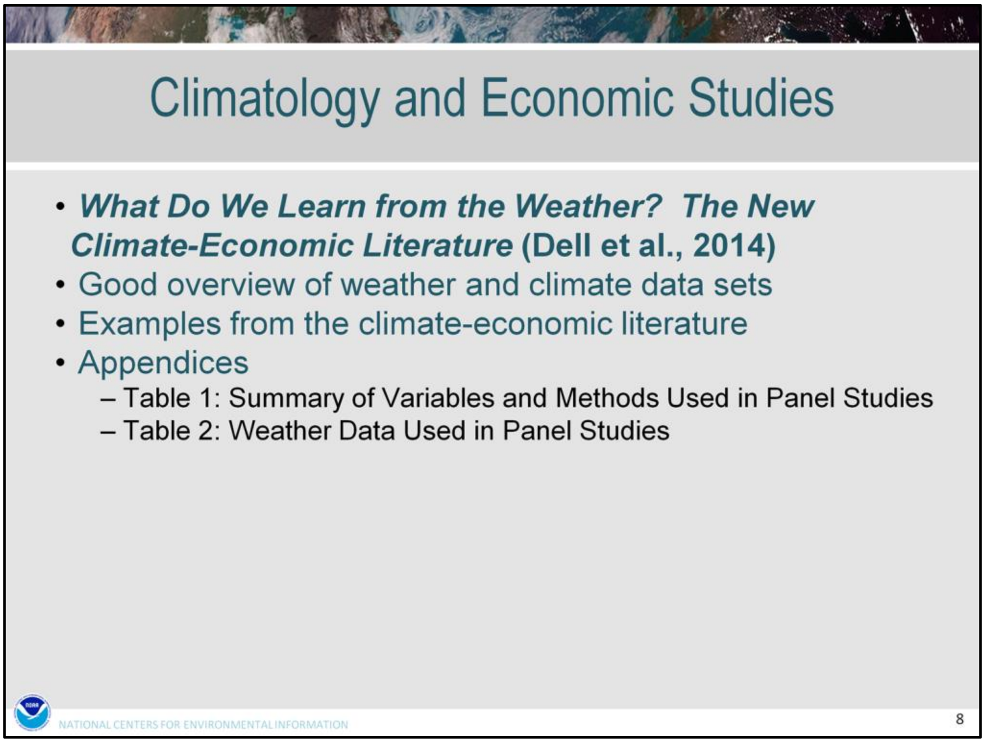
The next example shows how the extremely warm winter of 2011/12 and the extremely cold winter of 2013/14 affected natural gas prices and storage amounts (Schreck et al., 2015). The top panel shows the temperature anomalies for the northeastern quadrant of the United States. The vertical lines indicate the winter seasons that were warm (red), neutral (grey), and cold (blue). Note that the 2011/12 winter was very warm and the 2013/14 winter was very cold. The lower panel shows the natural gas futures price corresponding to the same time period. During the warm 2011/12 winter the prices fell to very low levels. However, during the cold 2013/14 winter, the price of natural gas spiked to its highest level in five years.



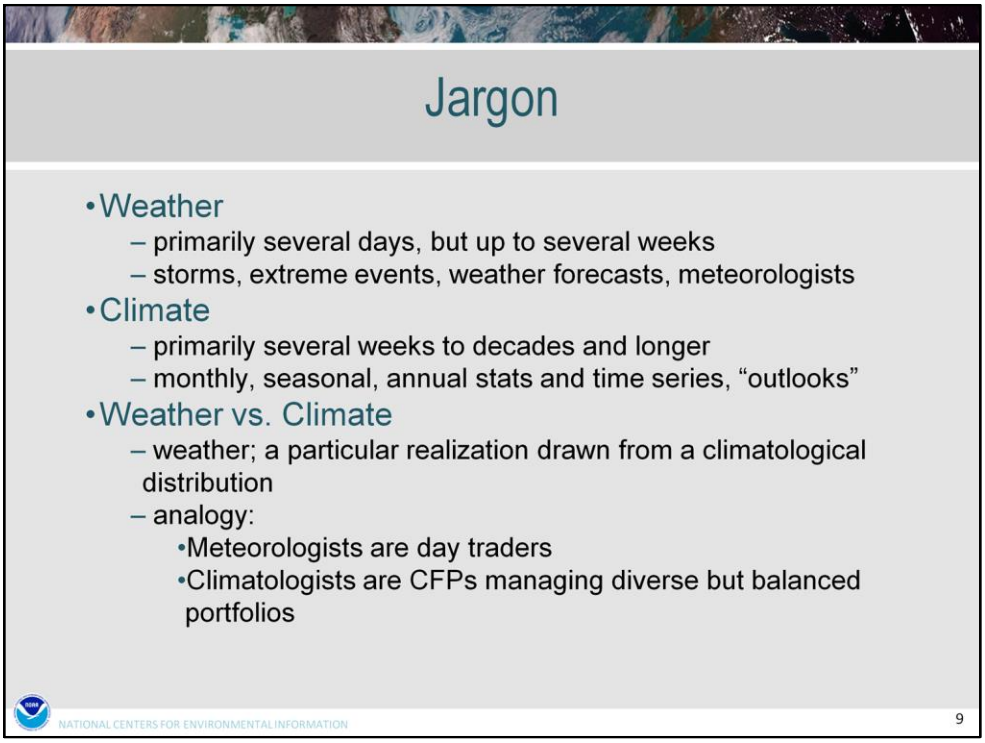
The charts on the left show the state level anomalies with these two winter seasons. The chart on the right shows natural gas storage levels from June through June for the 1994-20114 period. During the warm winter of 2011/12 (red line), natural gas had low demand so by spring and summer, supplies were at historic highs. The next winter was near normal, so while supplies dipped, they still remained a relatively high levels (blue line). Going into the extremely cold winter of 2013/14 (green line), natural gas levels were near record highs but fell to near record low amounts because of the very cold winter.



This final example shows a proposed methodology to incorporate "weather" adjustments to the current employment statistics. This is in addition to the seasonal adjustments that are routinely made. Including the weather affects can be very important, shifting the monthly payrolls number over +/- 100,000 (Bolden and Wright, 2015).



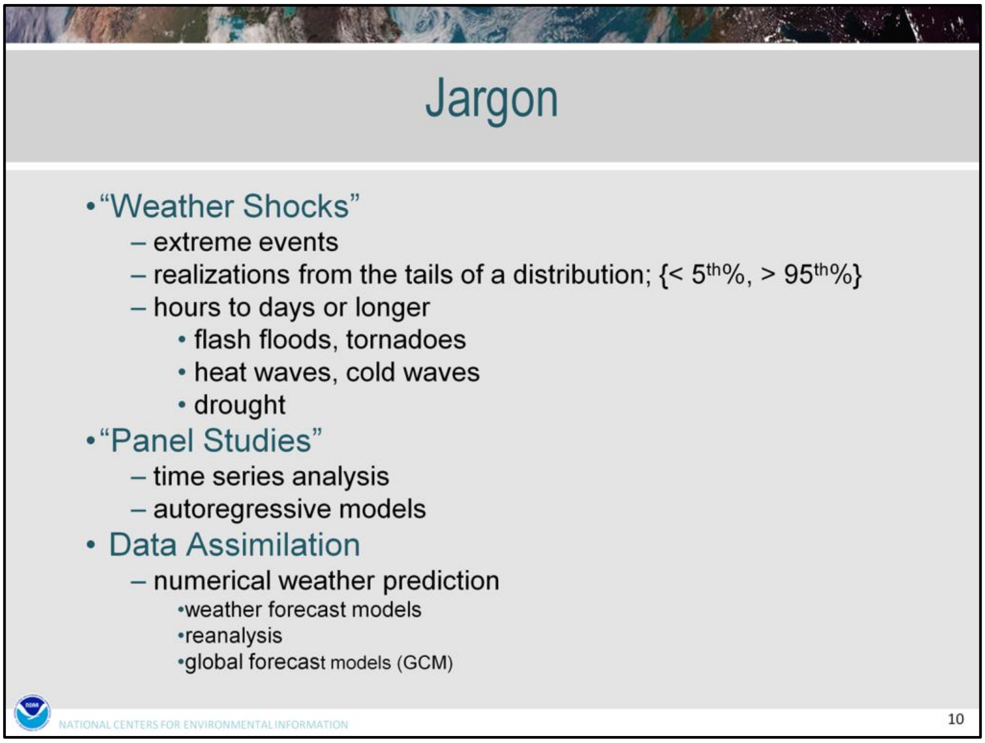
In "*What Do We Learn from the Weather? The New Climate-Economic Literature* " (Dell et al., 2014), the authors give an overview of different types of weather data available and reviews the new climate-economic literature. Recent work in this area is summarized including data sets used, methodologies, and findings. The authors also consider applications of climate change projections and how economic impacts of future climate change could be estimated.



My primary goal is to open a dialogue between NCEI and the economic community. So it is important to understand words that have specific meanings to our respective professions. This may seem obvious, but it’s best to be clear about things when communicating about technical topics. First let’s talk about weather vs. climate. Weather happens on time scales from minutes through days, but depending on the context perhaps up to several weeks. Examples include tornadoes, cold fronts, and heat waves. Meteorologists work with weather data; for example a forecaster at the National Weather Service.

Climate operates on times scales of months to years to decades or longer. Climatology typically involves developing and analyzing long term data sets of meteorological elements such as temperature or precipitation. Every 10 years, NCEI creates climatologies based on the last 30 years that are known as “Climate Normals”. Climate “forecasts” are know as “outlooks” because they have much less skill than a weather forecast. Outlooks are expressed in probabilistic terms. For example; a 40% chance of above normal precipitation for a particular region. In the United States, climate outlooks have the most skill during strong El Niños and La Niñas.

From a statistical frame of reference, a weather event is a particular realization drawn from some climatological distribution. A less formal description and analogy; meteorologists are day traders while climatologists are CFPs managing a diverse but well balanced portfolio.

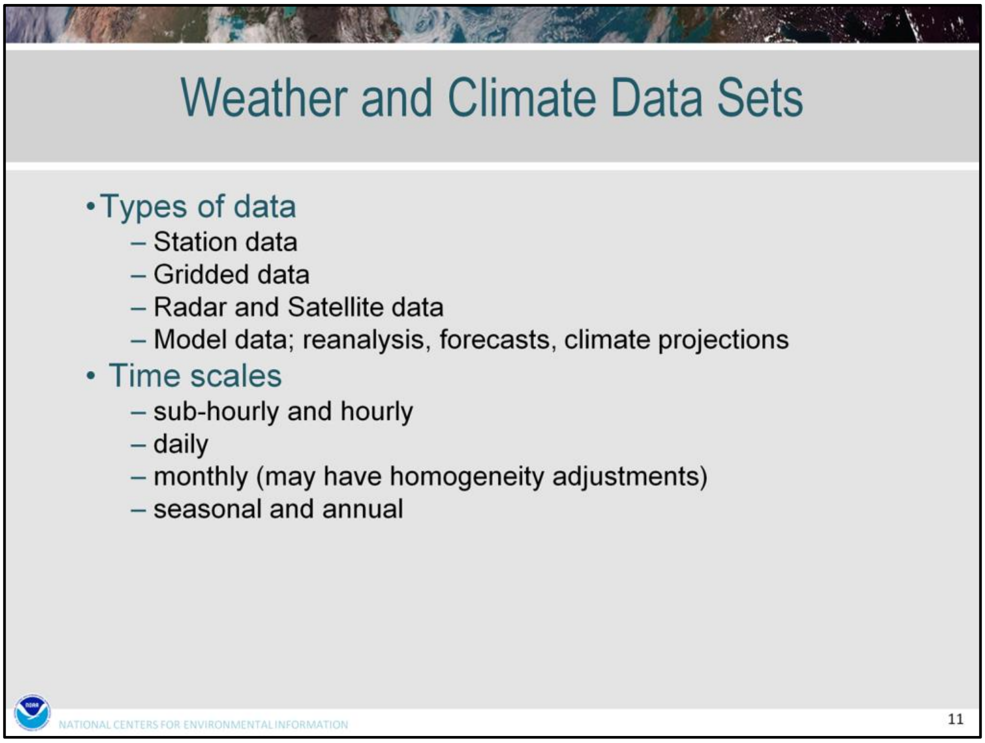


In the emerging climate-economy literature, economists often uses the terms “weather shocks” or “temperature shocks”. These typically refer to what meteorologists and climatologists call extreme events. Typical examples include record or near record maximum or minimum temperature, flash floods , heat waves, or severe drought. An extreme event often refers to a large storm. It could be rain, snow, or any type of meteorological event. In some settings “extreme” may be linked to a specific definition. For example, a daily rainfall event may be defined as extreme if its value exceeds the 95^{th} percentile of daily rainfall totals. Essentially, weather shocks are realizations from the tails of some climatological distribution. The shock (extreme event) could be “two-sided”. For example, the lower 10% or upper 10% of the daily minimum temperature distribution. So besides looking at record or near record low minimum temperatures, near record “high minimum temperatures” are also analyzed. The latter is important in climate change issues.

Many of the papers in the climate-economy literature use “panel studies”. However this term is seldom seen in the weather-climate literature. Panel studies refer to time series analysis which are studies where the data values are ordered chronologically (sorted from oldest to newest). These studies typically use regression where at least one of the independent predictors is time. Autoregressive models are also used in panel studies.

Some of the climate literature has used the term “data assimilation” to mean numerical weather prediction (NWP). NWP models are physics based models based on “first principles” and most commonly used to make weather forecasts. More on NWP in the next section.

The point here is that there is no right or wrong way of expressing a particular type of analysis. But it is important that economist and climatologists are able to clearly understand concepts when they are communicating with each other.

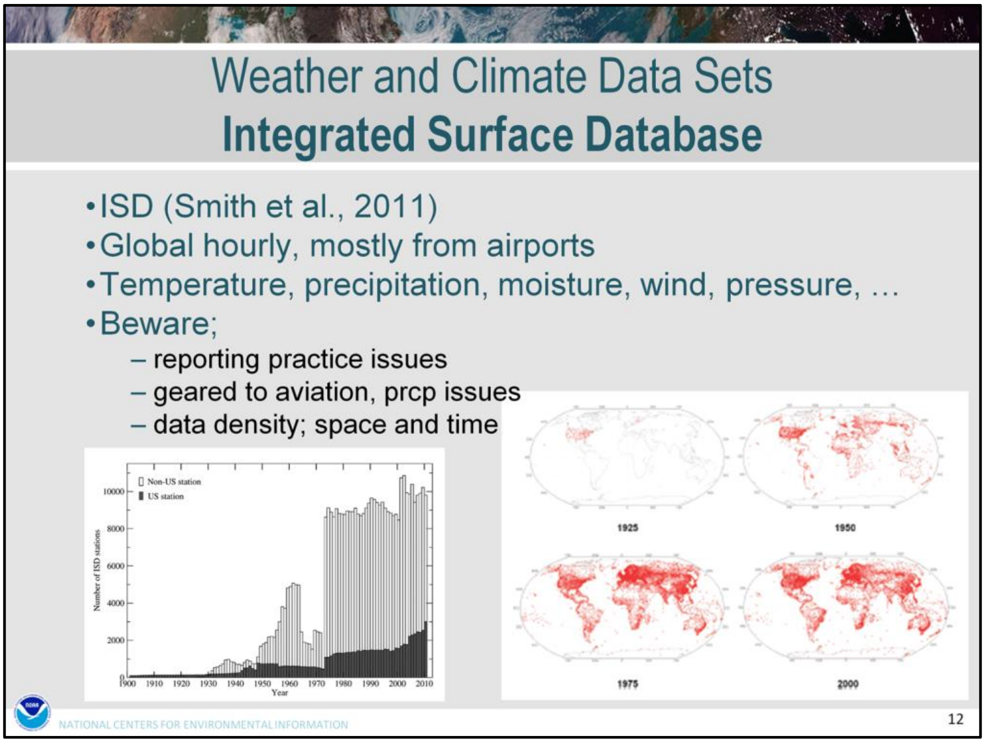


There are many different types of weather and climate data available. The most basic data is observed at weather stations. Many of these weather stations are located at airports and their primary purpose is to support aviation. The observations are taken hourly and under certain weather conditions (low visibility or low clouds) are taken sub-hourly. There are other observations that are taken once a day and only include elements such as maximum and minimum temp and total precipitation. In the field of geographical information systems (GIS), station data is considered “point data”.

Gridded data is typically station data that is interpolated to some type of regular grid. More about interpolation later. The advantage of gridded data is that it is spatially continuous. Grid data is also known as raster data in the GIS community. Remotely sensed data such as radar and satellite data are typically provided on grids. Although radar and satellite data are spatially continuous, the user of such data should be aware of attributes of the data such as resolution, update time, and any calibration issues.

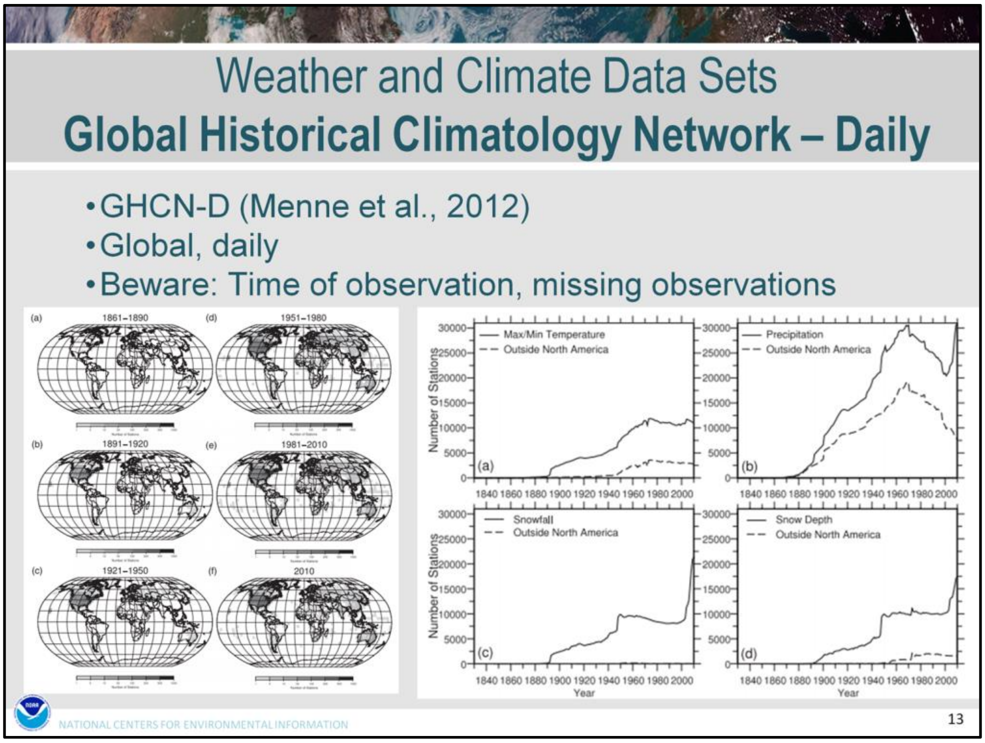
Within the weather and climate community, model data implies output data from a numerical weather prediction (NWP) model. Most forecast today are largely based on NWP. NWP uses input data from weather stations, including upper air observations made with instruments attached to balloons. The NWP models for one to five day forecasts are run twice a day at 00:00 and 12:00 UTC (7:00 pm and 7:00 am EST). NWP models are also run in “hindcast” or reanalysis mode on historic data to provide 3-dimensional datasets that are continuous in space and time as well as physically consistent. These reanalysis models are often run on a global scale. These global models are also run for decades into the future to produce climate projections. The use of global climate models is explained more in a later section.

As discussed above, climate data sets are produced at several time scales. Daily temperature and precipitation data is aggregated to monthly mean temperature and monthly total precipitation, respectively. Monthly temperature is often adjusted to eliminate inhomogeneity issues. A long time series of temperature data can be degraded by small station moves, changes in instrumentation, land use changes, or change of observation time. Monthly data can be aggregated into seasonal or annual data, all of which can be analyzed as time series (panels). All of these time series can be summarized with descriptive statistics to produce climatologies.

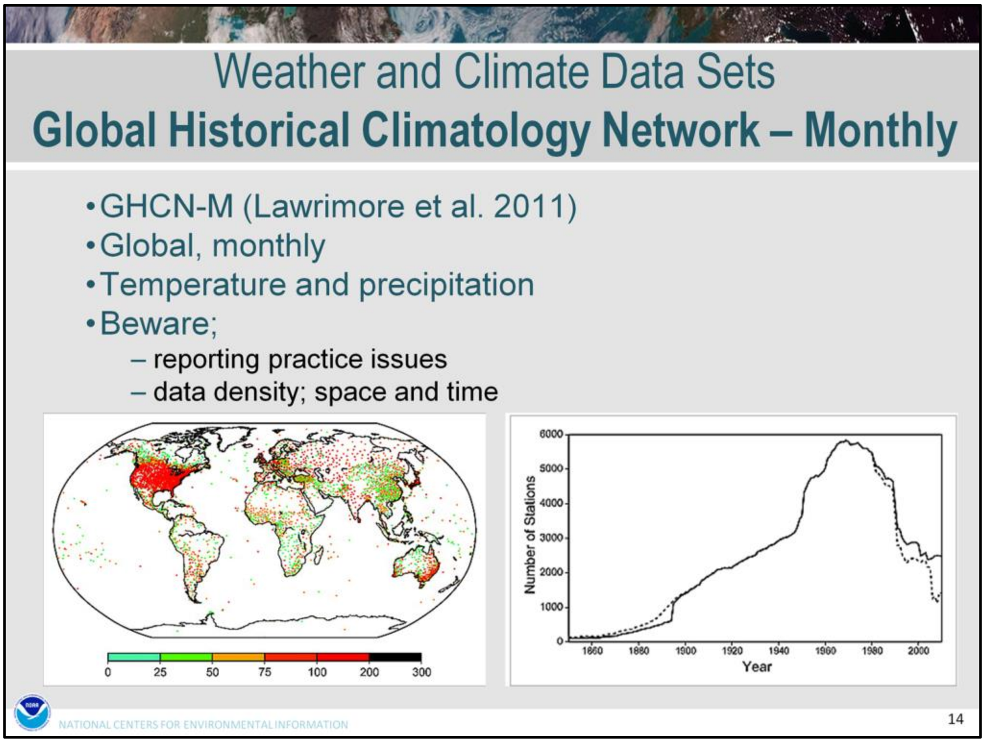


The Integrated Surface Database (ISD) contains over 2 billion surface weather observations from more than 20,000 stations worldwide 1900–present. The figure on the left depicts the approximate number of stations per year, which generally increase through time. One notable exception is the decline in reporting stations during the late 1960s through early 1970s due to the transition from keying of data to digital transmission/receipt of data. Some stations have more than 50 years of continuous reporting during the latter half of the time period; however, many stations have breaks in the period of record (e.g., 40 years of data may be spread over a 70-year period).

The figure on the right shows the spatial distribution of reporting ISD stations in 1925, 1950, 1975, and 2000. Since 1950, spatial coverage has been quite reasonable over North America, Europe, Australia, and parts of Asia, with noteworthy gaps in Africa and South America until the early 1970s, when the Global Telecommunications System came into existence. At present there are more than 11,000 active stations that are updated daily in the database (i.e., near real-time data that are ingested each day).

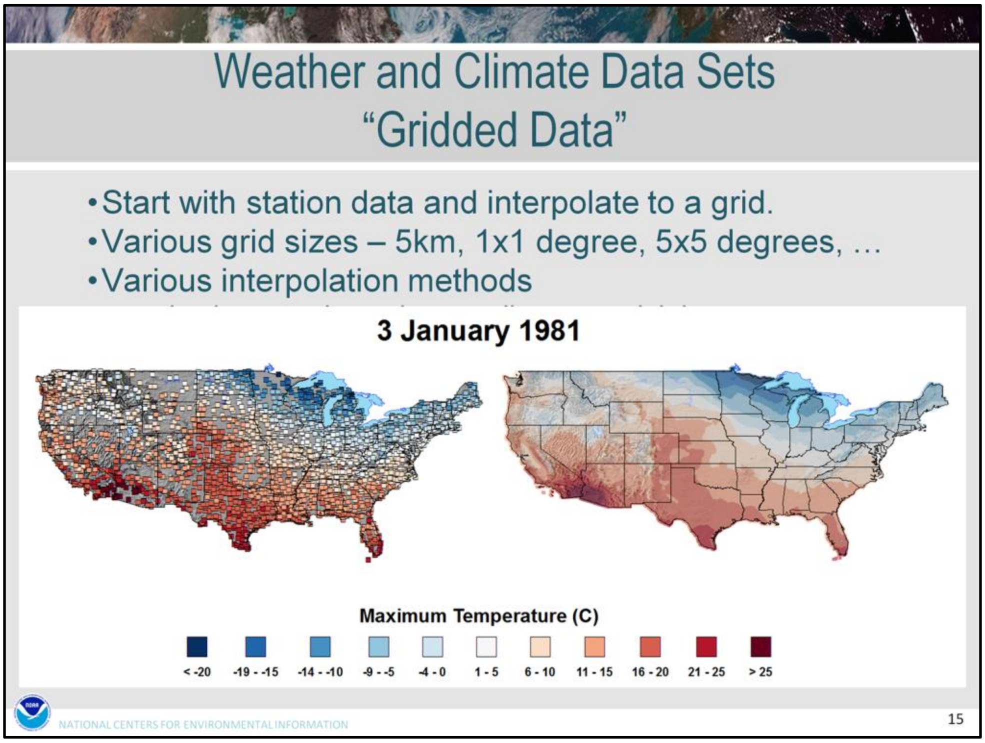


The Global Historical Climatology Network – Daily (GHCN-D) was developed for a wide variety of applications, including climate analysis and monitoring studies that require data at a daily time resolution (e.g., assessments of the frequency of heavy rainfall, heat wave duration, etc.). The dataset contains records from over 80,000 stations in 180 countries and territories, and its processing system produces the official archive for U.S. daily data. Variables commonly include maximum and minimum temperature, total daily precipitation, snowfall, and snow depth; however, about two-thirds of the stations report precipitation only. Quality assurance checks are routinely applied to the full dataset, but the data are not homogenized to account for artifacts associated with the various eras in reporting practice at any particular station (i.e., for changes in systematic bias). Daily updates are provided for many of the station records in GHCN-Daily. The dataset is also regularly reconstructed, usually once per week, from its 201 data source components, ensuring that the dataset is broadly synchronized with its growing list of constituent sources. The daily updates and weekly reprocessed versions of GHCN-Daily are assigned a unique version number, and the most recent dataset version is provided on the GHCN-Daily website for free public access. Each version of the dataset is also archived at the NOAA/NCEI in perpetuity for future retrieval.

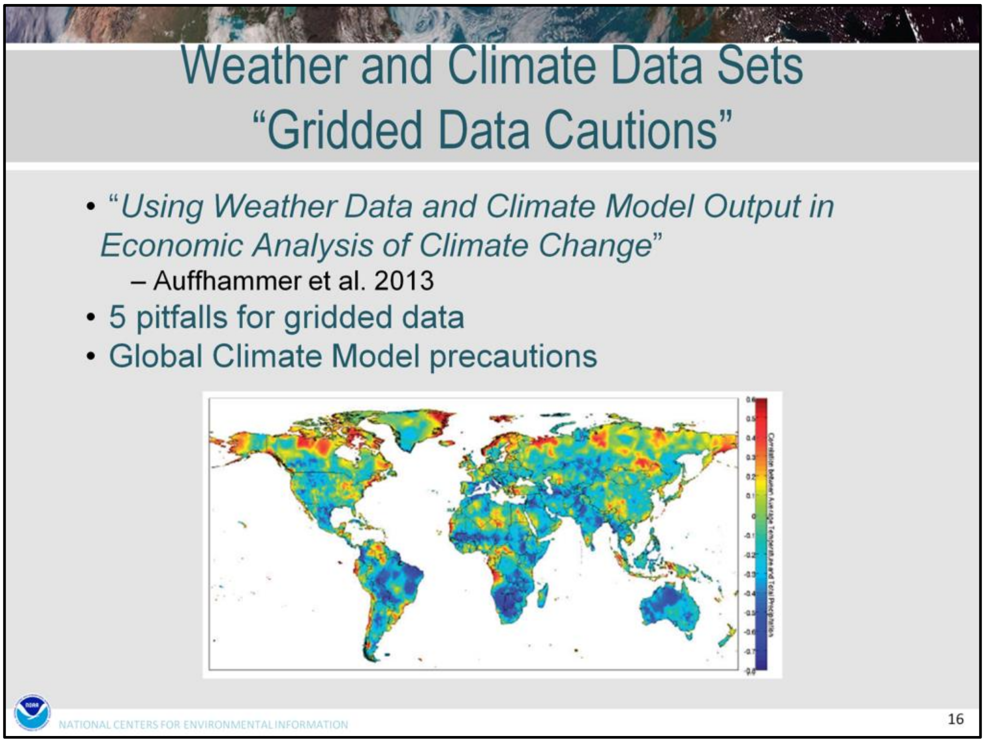


Since the early 1990s the Global Historical Climatology Network-Monthly (GHCN-M) data set has been an internationally recognized source of data for the study of observed variability and change in land surface temperature. It provides monthly mean temperature and precipitation data for 7280 stations from 226 countries and territories, ongoing monthly updates of more than 2000 stations to support monitoring of current and evolving climate conditions, and homogeneity adjustments to remove non-climatic influences that can bias the observed temperature record. The release of version 3 monthly mean temperature data marks the first major revision to this data set in over ten years. It introduces a number of improvements and changes that include consolidating “duplicate” series, updating records from recent decades, and the **use of new approaches to homogenization** and quality assurance. Although the underlying structure of the data set is significantly different than version 2, conclusions regarding the rate of warming in global land surface temperature are largely unchanged.

The figure on the left shows the locations of the 7,280 temperature stations in GHCN-M. The color indicates the number of years available at each station. The figure on the right indicates the number of stations available by year. The solid line is GHCN-M Version 3 and the dashed line is Version 2. The decrease in station counts after 1980 are due to station closures.



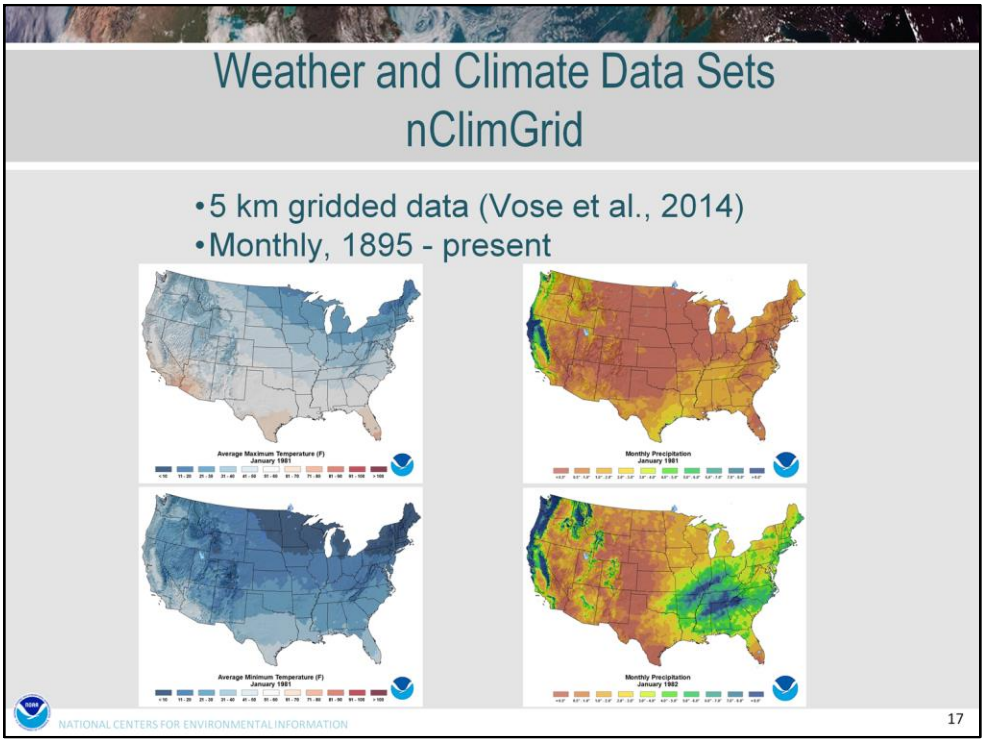
It is often useful to interpolate station data at irregularly spaced points to a uniform grid of regularly spaced points. The gridpoints may be aggregated by various geographic entities such as counties or states to estimate means or totals for those regions. This grid may also be combined with grids of other variables, such as population for example, to calculate an index. There are various methods to interpolate station data to grids. A very simple method is just to average the points that are within a grid cell. More typical is weighting by the inverse of distance; points that are close to a grid point are weighted more, points further away are weighted less. This is known as inverse distance weighting (IDW). A more sophisticated method is to use n-dimensional thin-plate splines. This method allows one to use multiple predictors in addition to the weather variable being interpolated. In the maps shown here, the 3 January 1981 maximum temperatures at about 2,500 stations are being interpolated to about half of a million 5 km grid points using a spline technique. Besides the temperature values, 5 km grids of elevation and a coastal influence are also used in the spline interpolation. The result is a much more accurate and realistic map, especially in the mountains where there is a relative shortage of observations. These data sets can be very large since one may be increasing the amount of data points by orders of magnitude. For example in the map above, the ASCII point data set is about 100 kb, while the ASCII version of the grid is approximately 10,000 kb.



In “Using Weather Data and Climate Model Output in Economic Analysis of Climate Change” (Auffhammer et al. 2013), the authors provide guidelines for using gridded weather data sets. The gridded data sets could be of the past or projections into the future from climate models. The 5 pitfalls listed for historic gridded data are;

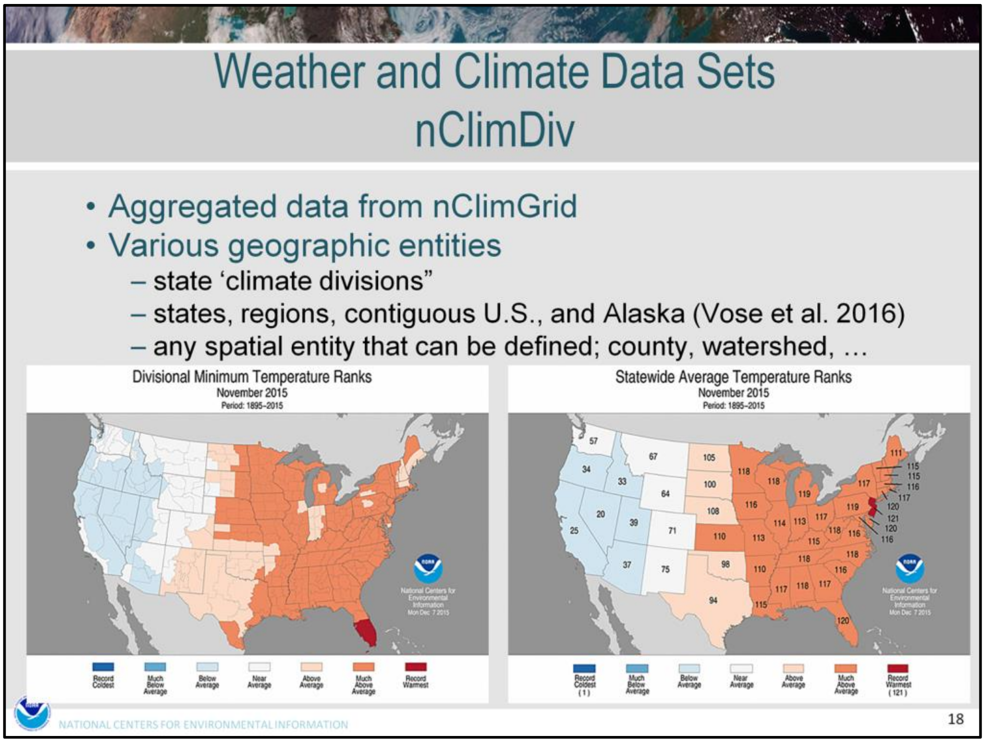
1. *choice of data set*
2. *averaging daily station level data across space*
3. *correlation of weather variables*. The figure above shows the correlation between annual average temperature and annual average total precipitation. The correlation can be negative or positive, depending on the region.
4. *spatial correlation*
5. *endogenous weather coverage*

There are numerous global climate models (GCM) and there are some differences in their projections. The authors suggest using a collection of GCMs and reporting the range of results. Another issue with GCMs is their relatively large grid size – up to 2.5 degrees or about 120 miles. A grid cell this size could include several different climate zones. This leads to “aggregation bias” . If projected climate data is needed at smaller scales, statistical or dynamical downscaling methods must be used.



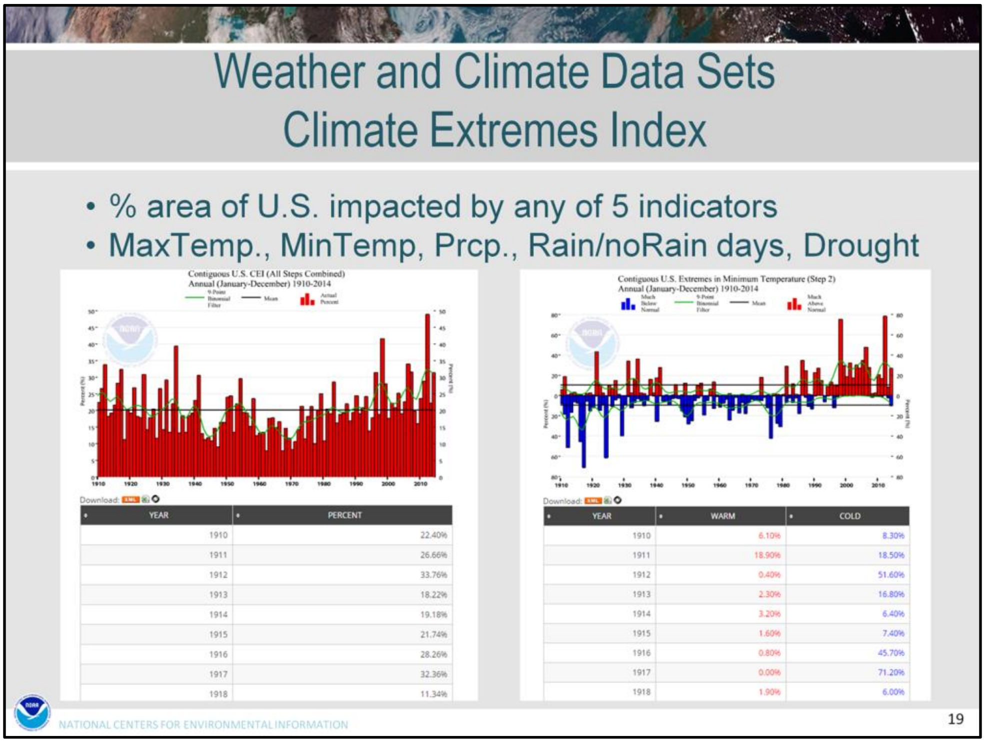
The nClimGrid data set is a 5km gridded data set of monthly averages of maximum and minimum temperature as well as total precipitation from 1895 through the present. The data set is updated monthly and is largely based on GHCN-M. The interpolation uses splines as described in the previous section. The temperature maps use elevation, coastal proximity, and climatological temperature inversion as additional independent predictors. Temperature normally decreases with height and proximity to the coast (especially the west coast). During the winter, some areas are subject to long term temperature inversions, so a climatological inversion grid is also used. This is especially important in alpine valleys. The precipitation uses elevation and slope and aspect as additional predictors. Precipitation typically increases with height and values are also larger on windward slopes of prominent mountains.

The maps on the left are average maximum and minimum temperature from January 1981. The top map on the right is the total precipitation for January 1981. It is easy to see the affect of terrain on the resulting temperature and precipitation fields. The bottom map on the right is total precipitation for January 1982 when a strong El Nino was in progress.



The nClimDiv data set is built upon the nClimGrid data set just described. The values at the 5 km gridpoints have been aggregated up to the state climate division level. Each state has up to a maximum of 10 climate divisions. In the east the climate divisions they are typically collections of counties. In the west, they are typically based on watersheds. In some areas, the climate divisions within a state may correspond to agricultural areas. But they are widely used by local and state governments and others for various weather and climate purposes. So for each month since 1895, one can view the average monthly temperature or total precipitation for each of the 344 climate divisions in the contiguous U.S.. This data is also available as ASCII files so one could compute trends or use this data in panel studies.

The map on the left shows the state divisional November minimum temperature in terms of ranks based on the 1895-2015 period. The map on the right shows the statewide November average temperature ranks based on the 1895-2015 period. The statewide values are obtained by aggregating the divisional values to the state level.

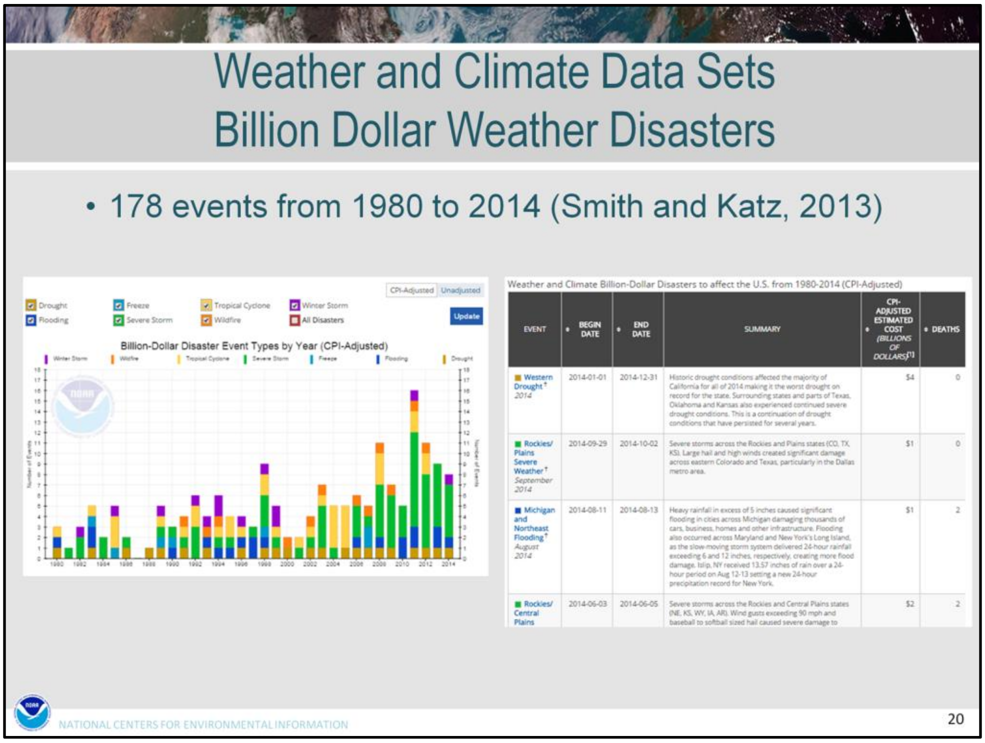


The Climate Extremes Index (CEI) is based on a set of climate extremes indicators that measure the fraction of the area of the United States experiencing extremes in monthly:

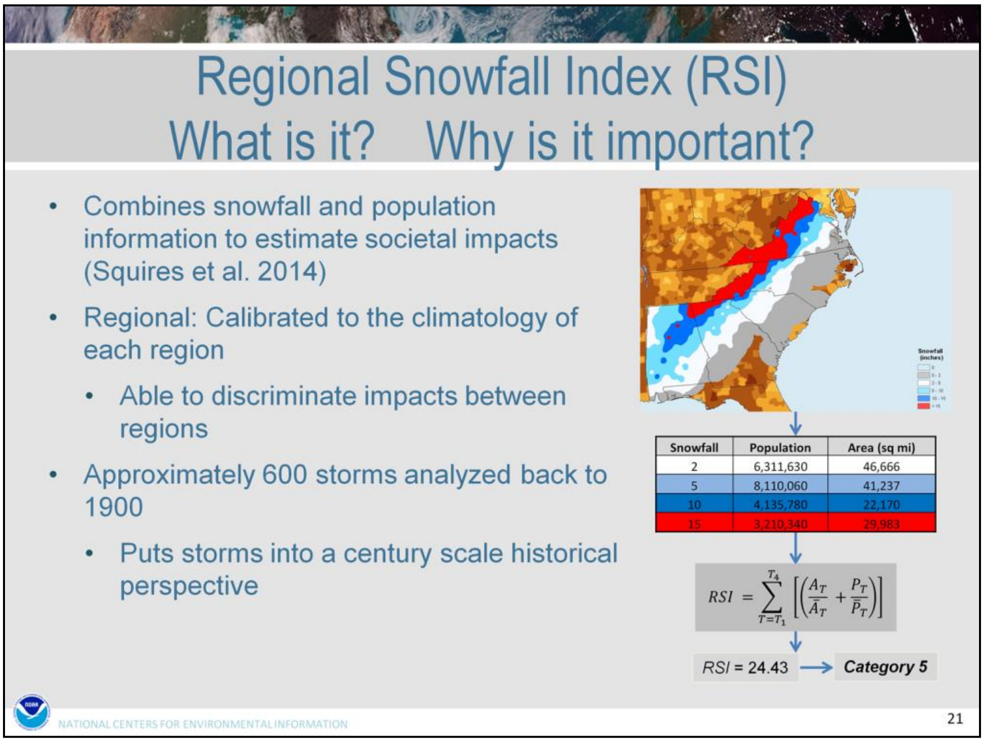
1. maximum temperature (upper and lower 10%)
2. minimum temperature (upper and lower 10%)
3. daily precipitation (upper 10%, this value is doubled so it is scaled with the other components.
4. number of days without precipitation (lower 10%) or with precipitation (upper 10%)
5. drought (lower 10%) or moisture surplus (upper 10%)

CEI is updated for the period from 1910 to the present in near–real time and is calculated for eight separate seasons; spring, summer, fall, winter, cold season, warm season, annual, and year to date.

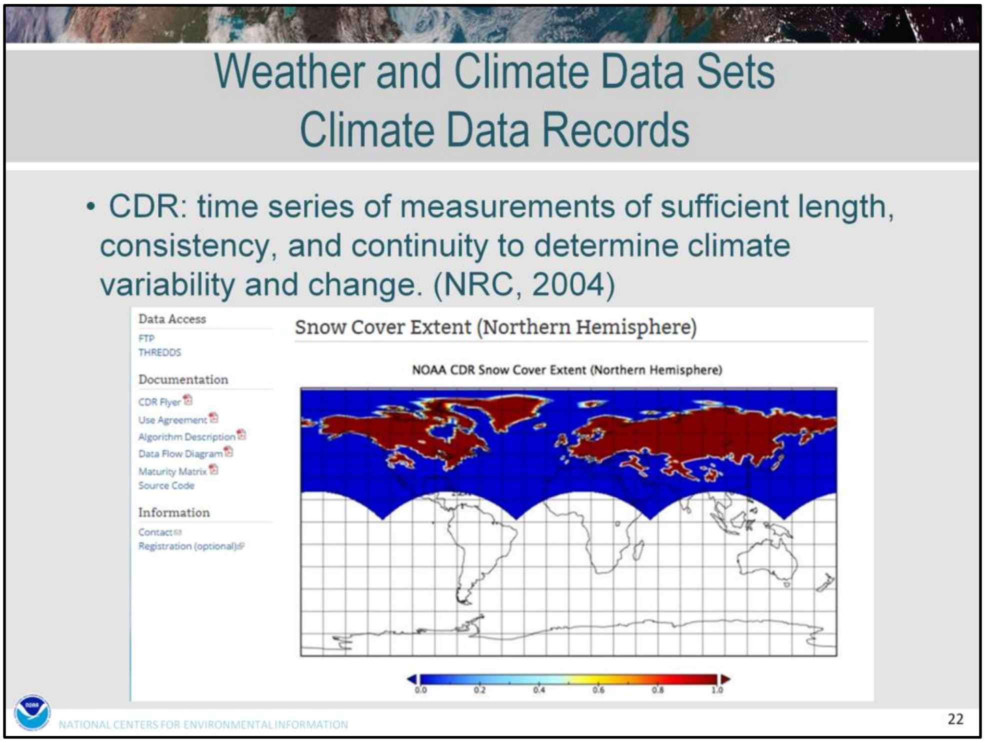
The chart on the left shows the annual time series of CEI for all five components . From CEI definition above, the expected value for any "season" is 20%. The CEI for all 5 components has been increasing for the last half century. The chart on the right shows time annual time series of CEI for the lower and upper portions of the minimum temperature component. This corresponds to "extreme cold minimum temperatures – lower 10%" and "extreme warm minimum temperatures – upper 10%". The latter is an important component of climate change. Also not the table below each chart. The columns are sortable and the tables can be downloaded in 'csv' format for quick import into Excel or other analysis software.



NCEI is the nation's scorekeeper in terms of addressing severe weather and putting climate events in historical perspective. As part of its responsibility of monitoring and assessing the climate, NCEI tracks and evaluates climate events in the U.S. and globally that have large economic and societal impacts. NCEI is frequently called upon to provide summaries of global and U.S. temperature and precipitation trends, extremes, and comparisons in their historical perspective. Found here are the weather and climate events that have had the greatest economic impact from 1980 to 2014. The U.S. has sustained 178 weather and climate disasters since 1980 where overall damages/costs reached or exceeded $1 billion (including CPI adjustment to 2015). The total cost of these 178 events exceeds $1 trillion.

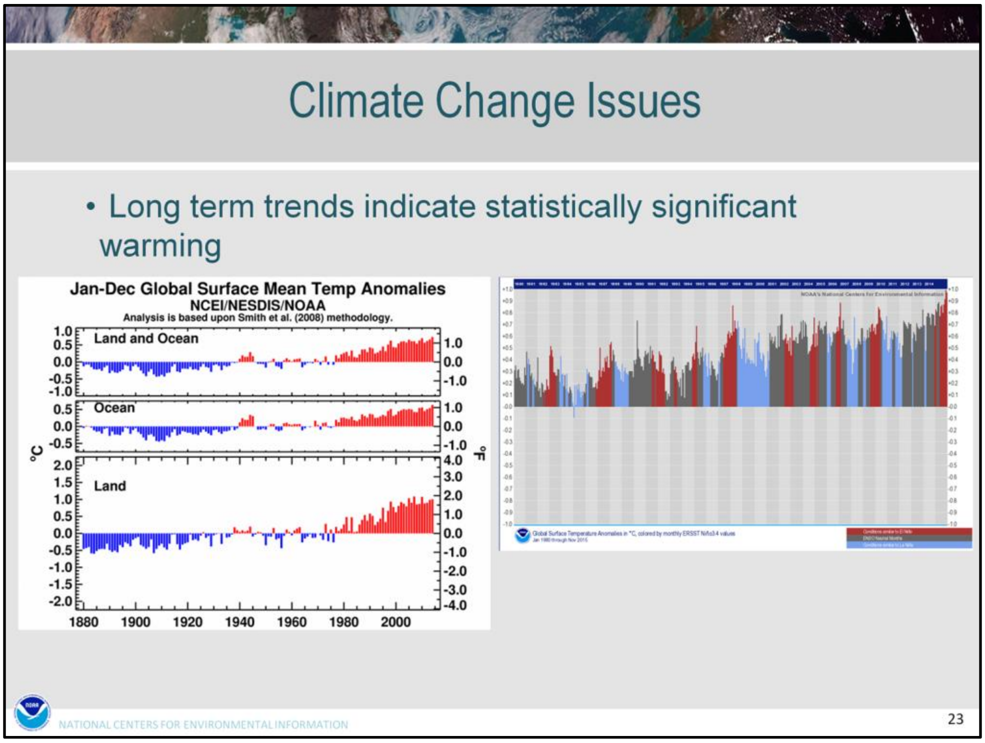


The Regional Snowfall Index (RSI) is based on the spatial extent of snowfall accumulation, the amount of snowfall, and the juxtaposition of these elements with population. Including population information provides a measure of the societal susceptibility for each region. The RSI places each storm analyzed in categories between 0 and 5, with category 5 storms having the most impact and only comprising a few percent of all storms. The RSI is an evolution of the Northeast Snowfall Impact Scale (NESIS) which NOAA’s NCEI began producing operationally in 2006. The RSI is a regional index calibrated to specific regions using only the snow that falls within that region. RSI uses region-specific parameters and thresholds and is calculated for six climate regions in the eastern two-thirds of the nation. The RSI has been calculated for over 600 snowstorms that occurred between 1900 and 2015 providing a century-scale historical perspective for these snowstorms. The RSI is computed for Category 1 or greater storms in near real-time; usually a day after the storm has ended.



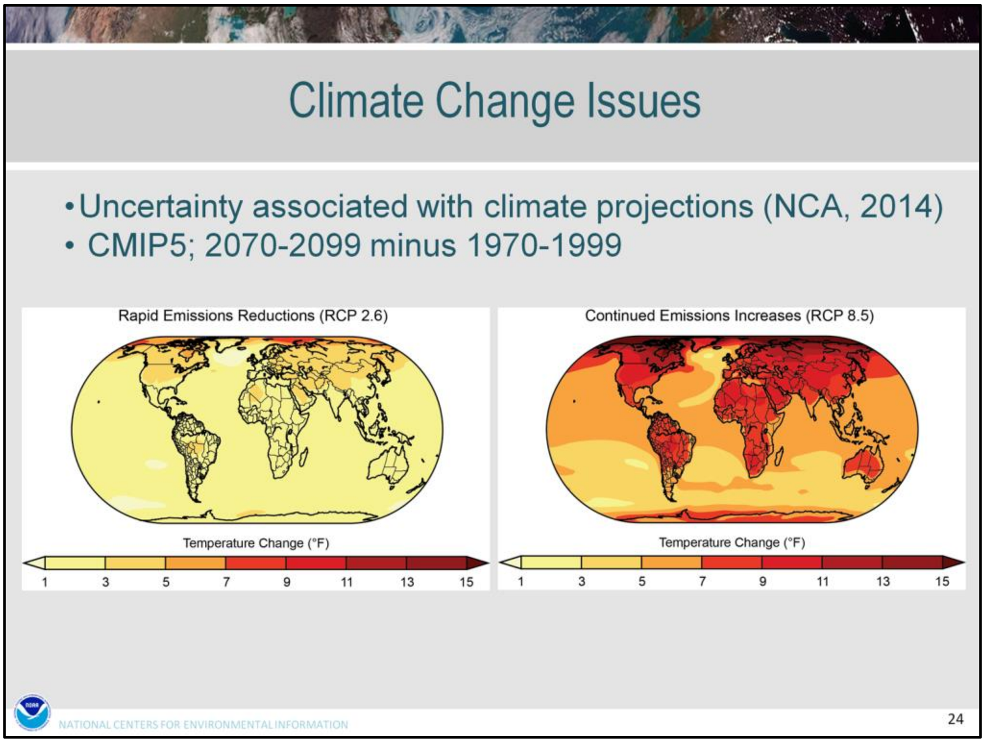
The mission of NOAA's Climate Data Record Program is to develop and implement a robust, sustainable, and scientifically defensible approach to producing and preserving climate records from satellite data. The National Research Council (NRC) defines a CDR as a time series of measurements of sufficient length, consistency, and continuity to determine climate variability and change. For the first time, NOAA is applying modern data analysis methods, which have advanced significantly in the last decade, to these historical global satellite data. This process will unravel the underlying climate trend and variability information and return new economic and scientific value from the records. In parallel, NCEI will maintain and extend these Climate Data Records by applying the same methods to present-day and future satellite measurements.

The results will provide trustworthy information on how, where and to what extent the land, oceans, atmosphere and ice sheets are changing. In turn, this information will be used by energy, water resources, agriculture, human health, national security, coastal community and other interest groups.
NCEI CDR data will improve the Nation's resilience to climate change and variability, maintain our economic vitality and improve the security and well-being of the public.



The chart on the left shows the annual global surface temperature anomalies from 1880 through 2014. The charts depict the land, ocean, and land + ocean combined. These values include the recent additions of additional data and improved adjustments implemented earlier this year (Karl et al., 2015). This chart does not include 2015 which will be far and away the warmest year on record.

The chart on the right shows the monthly global temperature anomalies from January 1980 through November 2015. The monthly bars have been color codded by El Nino/La Nina condition. El Ninos are colored red, La Ninas blue, and neutral months are colored grey. El Nino/Nina episodes are a part of natural variability and tend to warm and cool, respectively, the global climate. In this figure there is a positive overall trend in the temperature anomalies, but the La Ninas typically produce a short negative trend while the El Nino periods are associated with an increased positive trend. Perhaps these could be termed “shocks” to the long term underlying trend. This chart is particularly interesting because it shows the impacts of both natural variability and climate change on the overall temperature trend.

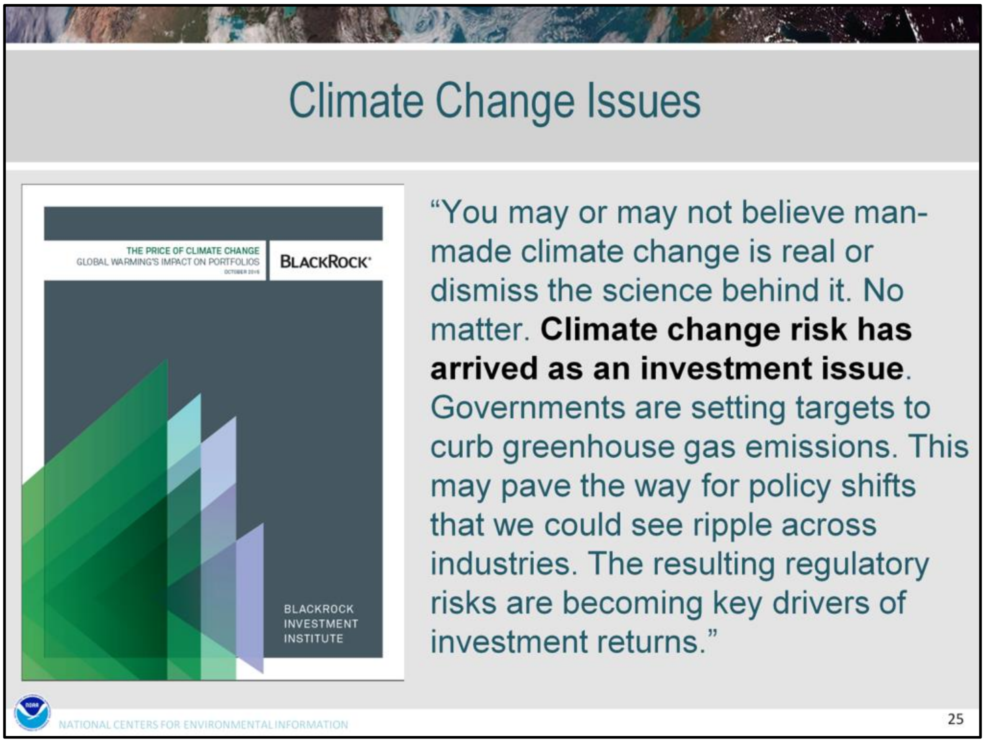


These charts show the projected average temperature change by the end of this century compared to the end of the 1900s. There is some uncertainty with the models themselves but even more so with the rate of emissions over the next 85 years. The rate of emissions is a function of population growth, economic growth, and conservation efforts that will take place in the future.

These projections use data from the Coupled Model Intercomparison Project Phase 5 (CMIP5) which is meant to provide a framework for coordinated climate change experiments and thus includes simulations for assessment in the Assessment Report 5 (AR5) as well as others that extend beyond the AR5. CMIP5 is not, however, meant to be comprehensive; it cannot possibly include all the different model intercomparison activities that might be of value, and it is expected that various groups and interested parties will develop additional experiments that might build on and augment the experiments described here.

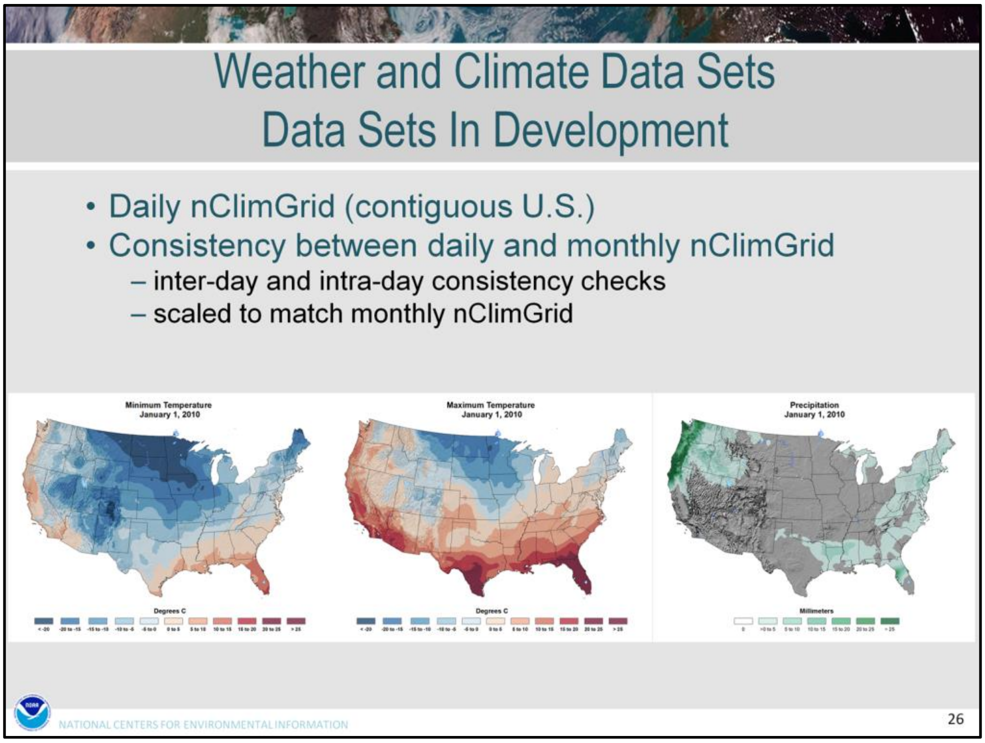
CMIP5 promotes a standard set of model simulations in order to:

1. evaluate how realistic the models are in simulating the recent past,
2. provide projections of future climate change on two time scales, near term (out to about 2035) and long term (out to 2100 and beyond), and
3. understand some of the factors responsible for differences in model projections, including quantifying some key feedbacks such as those involving clouds and the carbon cycle.



According to "THE PRICE OF CLIMATE CHANGE; GLOBAL WARMING'S IMPACT ON PORTFOLIOS", October 2015:

"You may or may not believe man-made climate change is real or dismiss the science behind it. No matter. **Climate change risk has arrived as an investment issue**. Governments are setting targets to curb greenhouse gas emissions. This may pave the way for policy shifts that we could see ripple across industries. The resulting regulatory risks are becoming key drivers of investment returns."



A new product under development at NCEI is a daily version of the monthly nClimGrid. The daily nClimGrid will have data from 1981 through the present and be updated daily when it becomes operational. The new daily grids will have:

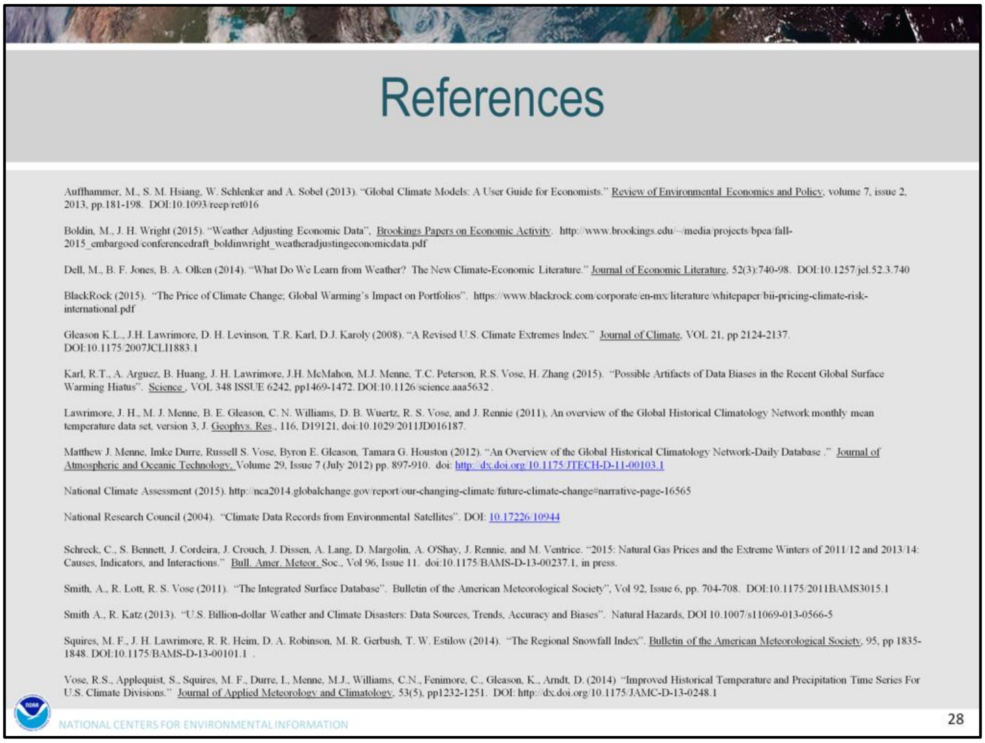
1. inter-day consistency; today's maximum temperature will not be colder than yesterday's minimum temperature and today's minimum temperature will not be warmer than yesterday's maximum temperature
2. Intra-day consistency; the maximum temperature is not less than the minimum temperature
3. The daily values at each grid point aggregate to match the existing monthly grid (mean value for temperature and total for precipitation).

The first two checks are important because the interpolations/predictions for each day for each grid point for minimum and maximum temperature are made independently. So a small percent of the gridpoints on each day will need to be adjusted. These adjustments are typically in mountainous areas where there is a lack of observations.

We are currently evaluating the daily maximum and minimum temperature grids and we will start evaluating the daily precipitation grids early in the new year. An experimental version of this product will be released this summer.

This presentation has shown some examples from the climate-economic literature and presented summaries of some climate data sets available from NCEI. These data sets have been peer reviewed, are available free to the public, and are archived in perpetuity.

The primary goal, however, is to open a dialogue between the NCEI Product Development Branch and the economic community. Our mission is to produce authoritative user-inspired products and data sets that will pass peer review and be easily available to the public. Therefore we need requirements from the economic community. The descriptions of our existing data sets and products should give you some idea of our capabilities and expertise. Of course new development is a balance between validated requirements and available resources. Please feel free to contact Mike Squires (Mike.Squires@noaa.gov) with any questions or comments.



Auffhammer, M., S. M. Hsiang, W. Schlenker and A. Sobel (2013). "Global Climate Models: A User Guide for Economists." Review of Environmental Economics and Policy, volume 7, issue 2, 2013, pp.181-198. DOI:10.1093/reep/ret016

Boldin, M., J. H. Wright (2015). "Weather Adjusting Economic Data", Brookings Papers on Economic Activity. http://www.brookings.edu/~/media/projects/bpea/fall-2015_embargoed/conferencedraft_boldinwright_weatheradjustingeconomicdata.pdf

Dell, M., B. F. Jones, B. A. Olken (2014). "What Do We Learn from Weather? The New Climate-Economic Literature." Journal of Economic Literature, 52(3):740-98. DOI:10.1257/jel.52.3.740

BlackRock (2015). "The Price of Climate Change; Global Warming's Impact on Portfolios". https://www.blackrock.com/corporate/en-mx/literature/whitepaper/bii-pricing-climate-risk-international.pdf

Gleason K.L., J.H. Lawrimore, D. H. Levinson, T.R. Karl, D.J. Karoly (2008). "A Revised U.S. Climate Extremes Index." Journal of Climate, VOL 21, pp 2124-2137. DOI:10.1175/2007JCLI1883.1

Karl, R.T., A. Arguez, B. Huang, J. H. Lawrimore, J.H. McMahon, M.J. Menne, T.C. Peterson, R.S. Vose, H. Zhang (2015). "Possible Artifacts of Data Biases in the Recent Global Surface Warming Hiatus". Science , VOL 348 ISSUE 6242, pp1469-1472. DOI:10.1126/science.aaa5632 .

Lawrimore, J. H., M. J. Menne, B. E. Gleason, C. N. Williams, D. B. Wuertz, R. S. Vose, and J. Rennie (2011), An overview of the Global Historical Climatology Network monthly mean temperature data set, version 3, J. Geophys. Res., 116, D19121, doi:10.1029/2011JD016187.

Matthew J. Menne, Imke Durre, Russell S. Vose, Byron E. Gleason, Tamara G. Houston (2012). "An Overview of the Global Historical Climatology Network-Daily Database ." Journal of Atmospheric and Oceanic Technology, Volume 29, Issue 7 (July 2012) pp. 897-910. doi: http://dx.doi.org/10.1175/JTECH-D-11-00103.1

National Climate Assessment (2015). http://nca2014.globalchange.gov/report/our-changing-climate/future-climate-change#narrative-page-16565

National Research Council (2004). "Climate Data Records from Environmental Satellites". DOI: 10.17226/10944

Schreck, C., S. Bennett, J. Cordeira, J. Crouch, J. Dissen, A. Lang, D. Margolin, A. O'Shay, J. Rennie, and M. Ventrice. "2015: Natural Gas Prices and the Extreme Winters of 2011/12 and 2013/14: Causes, Indicators, and Interactions." Bull. Amer. Meteor. Soc., Vol 96, Issue 11. doi:10.1175/BAMS-D-13-00237.1, in press.

Smith, A., R. Lott, R. S. Vose (2011). "The Integrated Surface Database". Bulletin of the American Meteorological Society", Vol 92, Issue 6, pp. 704-708. DOI:10.1175/2011BAMS3015.1

Smith A., R. Katz (2013). "U.S. Billion-dollar Weather and Climate Disasters: Data Sources, Trends, Accuracy and Biases". Natural Hazards, DOI 10.1007/s11069-013-0566-5

Squires, M. F., J. H. Lawrimore, R. R. Heim, D. A. Robinson, M. R. Gerbush, T. W. Estilow (2014). "The Regional Snowfall Index". Bulletin of the American Meteorological Society, 95, pp 1835-1848. DOI:10.1175/BAMS-D-13-00101.1 .

Vose, R.S., Applequist, S., Squires, M. F., Durre, I., Menne, M.J., Williams, C.N., Fenimore, C., Gleason, K., Arndt, D. (2014) "Improved Historical Temperature and Precipitation Time Series For U.S. Climate Divisions." Journal of Applied Meteorology and Climatology, 53(5), pp1232-1251. DOI: http://dx.doi.org/10.1175/JAMC-D-13-0248.1