Metrics: bridging between science and applications

- (mainly) Role of metrics in longer-term predictions
- Science-centric vs. applications-centric?
- Weather and climate phenomena – a common language
- Barriers to use

Joseph Barsugli
CIRES, University of Colorado at Boulder
NCPP, WWA, NC-CSC

CLIVAR SUMMIT
DENVER CO 8 JULY 2014

A couple of anecdotes based on recent CMIP5 evaluations

* ENSO Teleconnections

* North American Monsoon

*ThePractitioner's Dilemma (EOS)

- Proliferation of [downscaled climate projection] datasets
- Choice of data/method/ensemble often needs justification
- Not enough information to help with this choice
- Comparison and Evaluation (metrics are part of this) is essential

http://onlinelibrary.wiley.com/doi/10.1002/2013EO460005/pdf

What use are metrics for applications?

Credibility of Information

- "refers to whether an actor perceives information as meeting standards of scientific plausibility and technical adequacy. Sources of knowledge must be deemed trustworthy and/or believable, along with facts, theories, and causal explanations invoked by these sources" (Cash et al, 2002)

- "best available science"; justification of data choices

- Is this data/method suitable for my purposes? (e.g. does it do well with heat extremes?)

Legitimacy (open, reproducible, transparent evaluation of methods and data)

Uncertainty/Risk Characterization

- "Can you narrow the uncertainty?" "What is the best (sic) model?"
- Choice of methods, ensembles, datasets to adequately characterize possible outcomes
- Most decision-makers are used to dealing with uncertainty in some form, and climatic uncertainty just adds to the mix. Knowing limits to predictability (practical and theoretical), are of great use.
- Communication of uncertainty/confidence to stakeholders

What are Application-centric [climate] metrics?

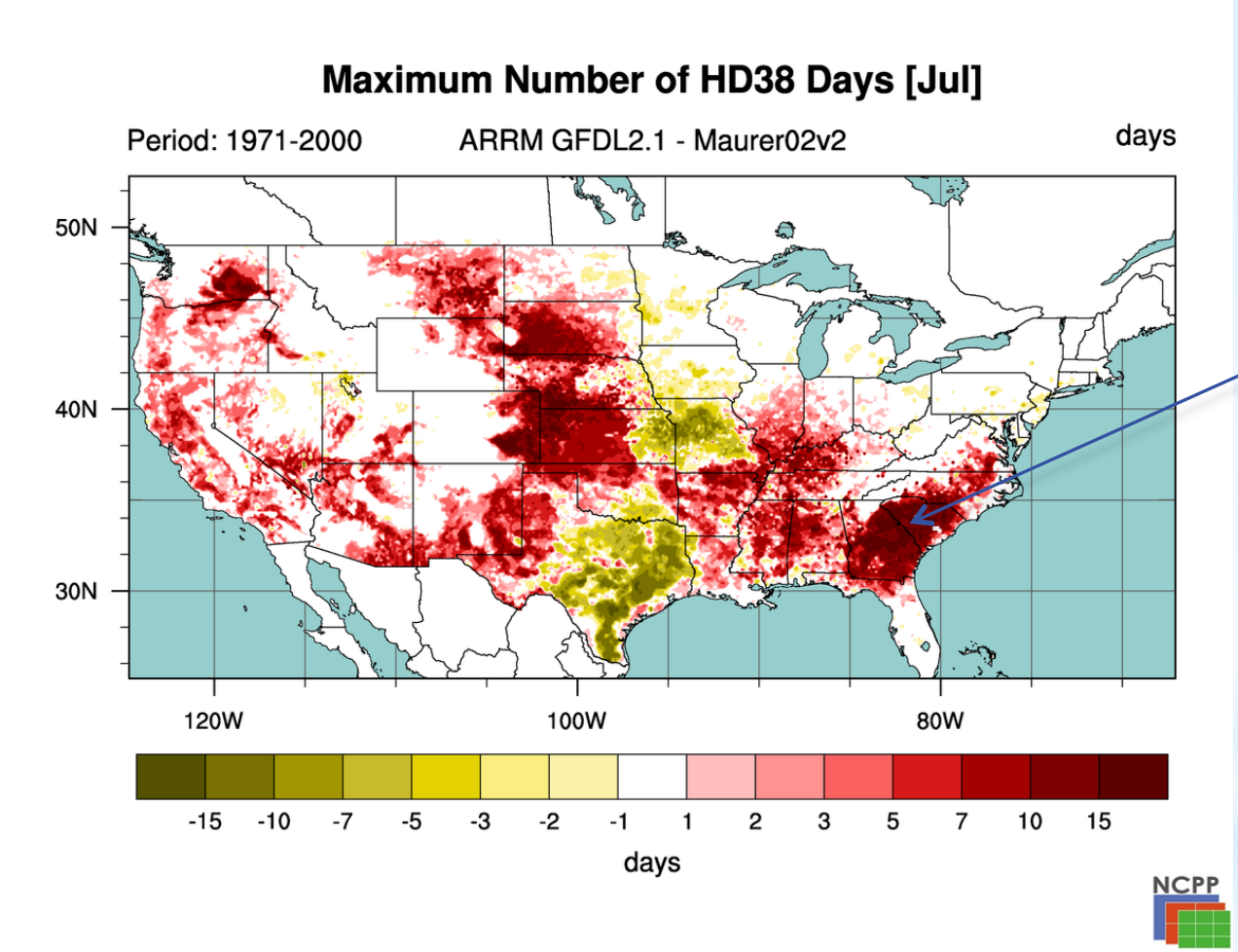
One type consists of.. metrics based on climatic indices that are commonly used in a sector/region

- Example: Health/Heat stress commonly use threshold-based temperature indices, and the threshold may vary depending on the region.
- Application-centric indices are more commonly based around extremes or cumulative indices.
- These may have different predictability/simulation skill than central tendencies.
- There are often common families of indices used in a sector for typical impacts modeling.

Other examples: spatial scales of precipitation (for hydrologic impacts); recurrence intervals for precipitation; heat/humidity indices; Bioclim indices; PET; …

Evaluating the quantity of interest!

Days > 38 C in July --- Maximum for the 30 year historical period.



15 days bias out of 31!

The downscaling method removes most GCM biases in the central tendencies, but it doesn't with this extreme. Large biases remain.

Scientist-centric metrics should help answer a scientific question that is of relevance and interest to a stakeholder..

I. **Processes** that are important regionally [globally] – particularly skill-process relationships

Q. Is a model in the correct hydrologic regime for this region?
A. Budyko Curve metric (water limited vs. energy limited)

II. **Emergent phenomena**

Q. Are ENSO effects well-reproduced in my region?
A. Power spectrum of ENSO, EOFs, teleconnections, etc…

III. Method testing through **idealized settings**

Q. How well does a statistical (downscaling) model handle non-stationarity?
A. Perfect-model (high-res/low-res GCM) evaluation

Climate and weather events as a common language

Metrics based on phenomena that are regionally and locally important and already "on the radar" of local decisionmakers hold the most promise for building a **common language to evaluate predictability**

PDO/decadal variability
Great Plains Organized Convective Systems
Landfalling Tropical Storms
Atmospheric Rivers/West Coast Extreme Precip.
Upslope Springtime Storms
Lake Effect Snows
…..

Good News: We have already embarked on this journey…
Bad News: What we have so far has mainly (with some exceptions) been developed by scientists, for scientists, little in the way of practical recommendations.

Barrier #1a Access (Data and Knowledge)

"Scientists generate a ton of data, but it goes into journals and then nobody uses it,"

Henry Markram, neuroscientist and director of the Human Brain Project
"We are building the technology to bring all of that together." NYT, 7/9/2014

(Straw) Proposals:

Support (require?) the publication of the data along with the articles, including metrics and indices for individual models.

Make salient figures available with captions that are understandable by stakeholders/applied scientists.

Improve open-access for summary/assessment articles

Support the publication of sectoral and regional assessments for widely used datasets.

Barrier #1b Access (IT backplane)

Support the use of existing standards and data archives for the dissemination of metrics

Support the development of metadata standards for evaluation/metrics

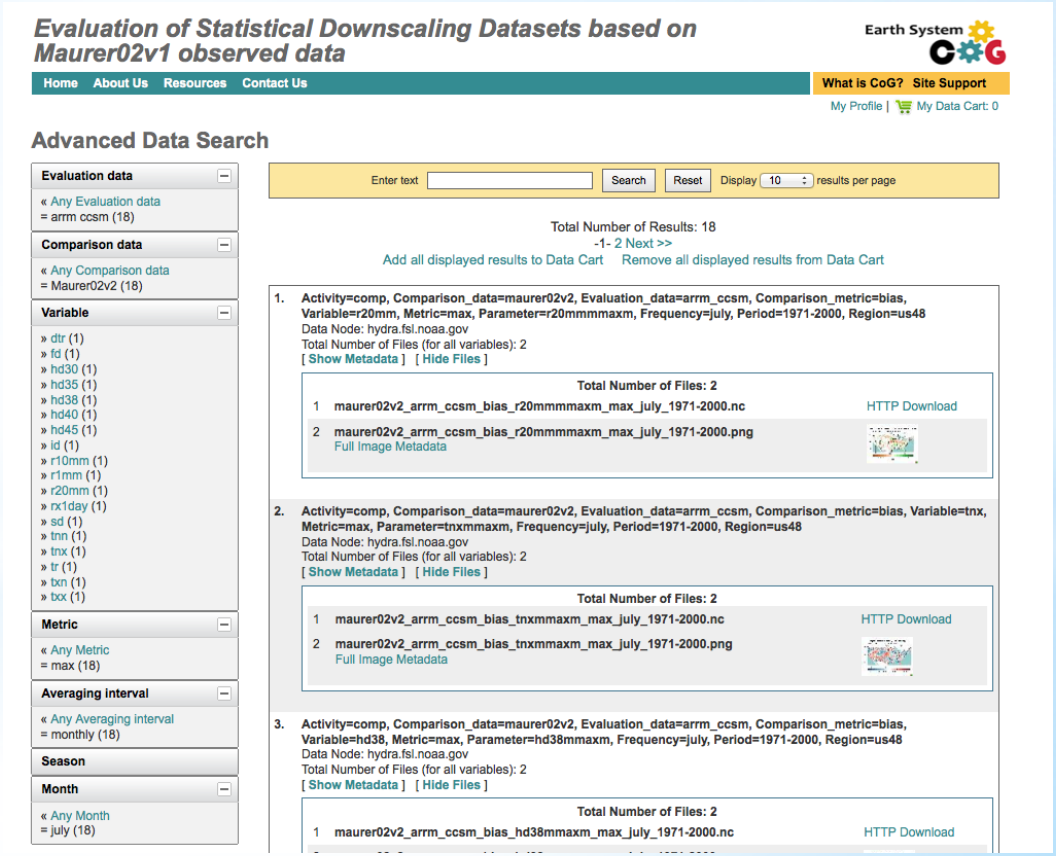
Support the development of new Information Technology to store, organize, locate, annotate, disseminate, regionalize, and integrate metrics into subsequent analyses.

NCPP downscaled data evaluation search

Evaluation of downscaled data
CLIMDEX (ETCCD:I) indices + heat/health sector indices
Stored on Earth System Grid
Faceted (“shopping”) search, + “Data Cart”

This is (or soon will be)
Based on open standards

And can be extended into
a community archive



www.earthsystemcog.org/search/ncpp

Barrier #2 Lack of standardization makes comparison difficult

[Further] support the development of standards and protocols for evaluation and comparison -- not just the production of standard sets of model runs.

Support international collaboration to this end

Examples:
Standard lists of indices (CLIMDEX/ETCCDI)
Evaluation Protocols (e.g. CLIVAR-decadal WG; NCPP; EU-COST/ VALUE…)

Towards standardized metadata for metrics…

Dataset Metadata
Activity = eval
Calendar = 365_day
Code Instance = URL: https://svn-jnt-ncpp.cgd.ucar.edu/branches/metrics/source; Revision: 525
Contact = ncpp_core@list.woc.noaa.gov
Creation Date = 2013-08-15T04:41:10Z
Data Node = hydra.fsl.noaa.gov
Data Type = downscaled
Downscaling Model Type = dynamical
Driving Data Type = gcm
Driving Model = ncep
Evaluation Data = narccap_regrid_50km_WRFG.ncep
Evaluation Project Code = QED2013
Experiment Id = 20C3M
Frequency = period
Index Node = hydra.fsl.noaa.gov
Institution Id = NOAA-NCAR-NCPP
Metadata Format = THREDDS
Metric Group = SectorEco
Model = narccap_regrid_50km_WRFG
Model Name = Weather Research and Forecasting Model - Grell convective parameterization
Number Of Files = 2
Parameter = bioclim1
Period = 1979-2000
Plot Title = BioClim1: Annual Mean Temperature
Product Filename = narccap_regrid_50km_wrfg.ncep_none_bioclim1_period_1979-2000
Project = NCPP
Project Id = NCPP
Protocol = observational
Region = US48
Relative Path = ncpp/eval/narccap_regrid_50km_wrfg.ncep/bioclim1/none/bioclim1/period/1979-2000/us48/
Resolution = 50km
Source Variable = bioclim1
Units = degC
Variable = bioclim1

Inherits metadata from the data being evaluated

Application-oriented metric groups

Evaluation protocol

Barrier #3 Evaluation is often not meaningful to applications practitioners -- obscure variables, wrong spatial/temporal scales, difficult to go “the last mile” to likely impacts

- Promote the development of application-relevant metrics.
- Promote a view on both sides of the “bridge” that focuses on metrics that characterize weather and climate events
- Engage sophisticated climate data users and though leaders in various sectors to blaze the trail.
- Promote “boundary organizations” as places where this interchange happens
- Support the interaction with impacts intercomparison projects such as AgMIP

Barrier #4 But still, what do the numbers mean?

Support interpretation and translation through regional and sectorally focused activities.

Provide recommendations for using the metrics/evaluations, even if these are “we can’t see daylight between the models”.

Concluding remarks

Scientist-centric metrics are necessary, but not sufficient for most applications

Applications-centric metrics are usually needed when

- Basic metrics are obscure to the non-climate/weather/ocean scientist
- Climate indices or processes have a strongly nonlinear impact (extremes, thresholds, for example) OR correlations among variables are important (heat indices, snowpack)

Metrics have the following uses in applications:

- As PART of the assessment of credibility of model(s) used and suitability for purpose
- As an aid in choosing models/methods/ensembles for uncertainty analysis

Metrics of climate and weather phenomena/events can provide a common language.

- Regional
- Sector/application
- Boundary organizations should be involved in developing this common language (RISAS, Climate Hubs, Climate Science Centers).

Coda: Comments on limits to predictability

Most decision-makers are used to dealing with uncertainty in some form, and climatic uncertainty just adds to the mix. Uncertainty does not have to be reduced to zero to be useful. Knowing limits to predictability (practical and theoretical, are of great use.

But… I see problems in communication around uncertainty, and expectations about climate projections (in particular).

The “Uncertainty Prayer”

Grant us…

The ability to reduce the uncertainties we can;

The willingness to work with the uncertainties we cannot;

And the scientific knowledge to know the difference.