Report – Workshop on Quantitative Evaluation of Downscaling (QED) – Evaluation of Downscaled datasets, August 12-16, 2013, NCAR, Boulder, CO

1. Overview, objectives, expected outcomes, and structure of the workshop

Overview:

NCPP organized a weeklong workshop on quantitative evaluation of downscaled climate projections on August 12-16 2013 at NCAR, Boulder, CO. The workshop was focused on evaluation of downscaled datasets and derived climatic indices for the United States and on the applications contexts in which evaluation may be useful. The workshop was a focal point of NCPP's efforts and aligned with NCPP's strategic goals to develop a flexible and extensible evaluation platform that offers important performance metrics on methods, data and tools; to facilitate the development of application-oriented communities; and to become an authoritative source for practices and standards, recommendations and guidance for use of local and regional climate predictions and projections. The workshop was sponsored by the NOAA Climate Program Office and the NSF Cyber-Enabled Discovery and Innovation (CDI) program, which provided funding to the workshop as a testbed for the CoG collaboration environment.

More than 90 participants attended the workshop. Attendees were from non-governmental organizations (NGOs), universities, federal agencies and labs, local decision-making institutions, and private companies. The invited participants represented the communities involved in the development, provision, and the use of downscaled climate data for applications. These include climate scientists, information technology specialists, practitioners and applied scientists in water resources, ecological, agricultural and human health impacts of weather and climate. NCPP engaged social scientist Jillian Wallace to evaluate the workshop through interviews with participants during the workshop, and an online survey after the workshop.

Objectives:

- To introduce elements of a quantitative and objective framework that NCPP is developing for evaluation of the credibility of downscaled projections, and to engage the participants in the advancement of this framework into a community standard.
- To bring the different communities involved in the chain of providers and users of climate information together to advance the development of standards for evaluation.
- To identify the needs for information that aids in translation of numerical climate data into usable knowledge for specific applications sectors.
- To facilitate the productive interactions between these communities and to initiate the development of a collaborative *community of practice* through working groups focused on water resources, ecological, agricultural and human health related impacts of weather and climate.

Expected outcomes: NCPP's core and technical teams in collaboration with the NCPP advisory Climate Science Applications Team (CSAT) identified some expected outcomes and products to be achieved or developed prior, during and after the workshop.

A. Prior to workshop

a) a prototype standardized evaluation framework that uses standard data sets for comparison, common protocols, tests and metrics for evaluation

b) standardized metadata to describe downscaling techniques and tools for collecting, archiving and searching this information
c) IT infrastructure to facilitate community analysis, evaluation and provision of climate information for the purposes of the group activities at the workshop
d) a working bibliography of downscaling techniques related to the available datasets and the selected application problems

B. After the workshop

a) comparative evaluation of downscaled datasets based on the prototype framework
b) sector and problem-specific guidance based on quantitative and comparative evaluation of available downscaled datasets
c) a publically accessible database of metadata that describes downscaling techniques
d) identified gaps and needs for further development of the evaluation framework and metadata
e) identified gaps in the availability and usefulness of downscaled data for the selected applications and sectors.
f) accessible case study descriptions for application problems considered at the workshop

Workshop structure:

The efforts of NCPP are informed and guided by the growing literature on improving the usability of science-based knowledge in planning and management. Successful use of science-based knowledge requires the iterative co-production of application-oriented decision packages and solution strategies. This is a messy process and one of the premises of NCPP's vision is that some level of order can be extracted from representative samples of real-world use cases. The workshop program was designed with this iterative co-development in mind, but also with the realization that there are substantial barriers between knowledge production and knowledge use. How to reduce these barriers is a research question unto itself. The workshop program was conceived around the two ends of the chain that connects knowledge generation and provision and climate-knowledge users. In this case the climate-knowledge providers are identified as the climate scientists, especially from groups who have taken on the provision of downscaled climate projections, and the knowledge users as practitioners. In response to several of the practitioners having the experience of being the "token practitioner" at, fundamentally, a meeting of scientists, many practitioners from four focus areas were invited. Though the first part of the workshop focused on scientists presenting fundamental information, each day also included presentations from the practitioner community as well as panels, which provided broader perspectives of the practitioner's experiences. The meeting also included many of the members of the community of scientists and practitioners who have experience over the past five years of using climate knowledge in planning and management. These people are *de facto* pioneers in translation of climate science for real-world applications. They provided information on practices that have proved effective, and it was the intention that they would help maintain the balance so that the meeting did not fall into the pattern of the scientists and practitioners simply dropping their knowledge and needs over the barriers that limit usability.

The focus of the discussions during the first half (Mon-Wed morning) was on the presentation of tools and protocols whose design was initiated prior to the workshop, the scientific issues related to downscaling and evaluation of credibility of dynamically and statistically downscaled climate projections, practitioner's requirements and exposure of successful and unsuccessful experiences in planning and management. The second half of the workshop was more iterative. Originally conceived around use cases to define the links in the end-to-end chain between climate-knowledge providers and users, much of this iterative discussion evolved to more substantive definition of a common starting point for providers and users. A result of this was more

concrete definition of the needs for guidance and translation related to the applications that were represented. There was substantial improvement of the definition of the link in the end-to-end chain that the practitioner's felt NCPP needed to provide.

The program in more detail: The slides from presentations and in many cases the actual presentation and discussion are available at http://earthsystemcog.org/projects/downscaling-2013/Agenda .

- Keynote presentations on each day were centered on applications and decision perspectives from each of the four featured applications areas
- Day one started with a welcome by Wayne Higgins of NOAA and Tom Bogdan of UCAR. NCPP introduced the proposed evaluation framework including three families of evaluation protocols and three groups of evaluation indices and metrics. NCPP also presented elements of the evaluation infrastructure, which included pre-computed national scale evaluations available at the NOAA ESRL ESGF node, and the development efforts and tools related to access and provision of metadata. The CoG collaboration environment, which hosted the workshop website, was introduced, along with how to access the pre-computed evaluation data and images through the CoG advanced search interface. The first day also included presentations and discussions of the baseline data used in the evaluations – station data and gridded observations.
- Day two was devoted to presenting the comparison of downscaled GCMs to observations and the results of the "perfect model" experiments designed to investigate the impact of non-stationarity of the Earth's climate on use of climate projections. Presentations on process-based evaluation of dynamically downscaled datasets, applications related comparisons using indices of interest to practitioners from a specific sector, and discussions of experiments based on idealized comparisons concluded the demonstration of the specific elements of the proposed evaluation framework. The summary discussion at the end of day two addressed issues related to the evaluation framework, the need for development of standards, and the specific needs for climate information to be used for applications.
- Day three was a transition towards the applications-focused working groups, which included the presentation of the Climate Translator (CT)/Open Climate GIS (OCGIS) capabilities for provision of local and regional information and evaluations. The presentations and panel discussions provided the scientists and practitioners the opportunity to share various experiences and issues that they have encountered when using climate data for adaptation and risk assessments. The first of three applications breakout sessions was dedicated to "problem identification" -- identifying issues related to the evaluation framework, and use of climate information in general, that the working groups would address.
- Days four and five were devoted to discussions and breakout sessions by application in order to identify gaps and needs for the provision of translation and guidance for appropriate use of climate data by practitioners. The last day included the summary presentations of the breakout sessions from the previous day. The final discussion on the last day identified future directions for NCPP efforts and defined a specific "target audience" for the capabilities and tools that NCPP is providing.

2. Workshop outcomes

The outcomes of the workshop are organized in three major categories: data and evaluation, translation, and community

A) Data and Evaluation: Needs and Gaps

Prior to the workshop:

NCPP advanced the development of an evaluation framework for assessment of the credibility of downscaled climate projections. The framework uses standard datasets for comparison, common protocols, tests and metrics for evaluation. Draft Protocols and Metrics documents are available online and always open for comments. A series of releases is envisioned, as experience is gained, the field advances and there is broader engagement. A forum paper that poses the social and the scientific rationale behind the development of this framework was published in EOS-Transactions of the American Geophysical Union, on November 12, 2013 – "The Practitioner's dilemma: How to assess the credibility of downscaled climate projections".

NCPP implemented the initial elements of the evaluation framework – observational and "perfect model" protocols. The observational protocol national-scale index calculations and evaluations were performed on the NCAR high-performance computing machines. The evaluation results are hosted on the NOAA ESRL ESGF node and are available through the CoG advanced data search capability. The "perfect model" experiments, designed to evaluate the stationarity assumption, were performed via collaboration with NOAA GFDL and Texas Tech University. "Perfect model" datasets are available through the ESGF node at GFDL; in addition, NCPP provided access to the standardized evaluation of the experiments results via the NOAA ESGF node in Boulder, CO.

To support the workshop and provide access to these experimental results, the NESII team in collaboration with the NCPP tech team developed an IT infrastructure and tools that facilitated the quantitative evaluation of the downscaled datasets. Novel capabilities to publish unified comparison results (NetCDF), images (png) and metadata (xml) files on an ESGF node were introduced. In addition, the capability of the CoG environment to search and display the evaluation results located on the ESGF node was presented. This search capability relies on a directory structure developed with attention to linking, uniquely, digital data, image and metadata files. The metadata information for the evaluation and comparison results and maps specifically conforms to the set of global attributes for publishing of downscaled datasets in EGSF nodes developed by collaboration between NCPP, GFDL, NASA, and the international community. In addition, the NCPP team made headway in the direction of standardization of metadata that describes the details of downscaling methods and datasets by expanding a model schema and controlled vocabularies developed by the international Earth System Documentation project (ES-DOC, http://earthsystemcog.org/projects/es-doc-models/). This metadata schema, called the Common Information Model (CIM), was initially developed for CMIP5 and has associated tools that include a metadata Questionnaire (developed by NESII, an ES-DOC partner), a Viewer, and a Comparator. NCPP customized the CIM Questionnaire in order to collect information about dynamical downscaling. This questionnaire incorporated the existing CMIP5 GCMs atmosphere and land controlled vocabularies and was intended to allow early-adopter users to develop dynamical downscaling instances during the workshop.

NCPP in collaboration with the NESII team developed additional tools in order to provide access to climate information and evaluation and comparison results on local and regional scales through the Climate Translator (CT) web-based interface to the Open Climate GIS (OCGIS) software package. CT/OCGIS allows the provision of local climate information in a variety of formats (e.g., shapefiles, csv, netCDF) and for user-defined or pre-defined regions (states, counties, ecological regions, several metropolitan areas, HUC-8 regions).

Outcomes and lessons learned:

The discussions during the workshop confirmed that evaluation is relevant and necessary. "How good are the data for my region and application?" and "What are the best choices of downscaled data for my application?" were frequently asked questions. These questions prompted a detailed and engaging discussion on

"If a service similar to NCPP had existed how it could have helped you in your work?" A specific need for the full product value chain (product stream) of climate information – from access to downscaled data, to inter-comparisons of observed data, to historical evaluations of Global Climate Model (GCM) and downscaled climate model data, to provision of ensembles of future downscaled projections and uncertainty descriptions. This full product stream of climate information needs to fulfil the requirements for:

a) relevance to specific applications – many new indices, metrics and variables of interest were proposed such as trends, variability and additional distribution related metrics, spatial autocorrelation metrics, multivariate temporal and spatial coherence metrics, indices of cumulative events (such as growing degree days or forcing and chilling units), indices of events' duration and intensity, snow and drought related indices, evaporation, SPI, additional heat related indices, and additional variables, e.g., wind and solar radiation, humidity.

b) transparency of information, including calculation workflows and metadata describing the datasets. The participants explicitly indicated that metadata of downscaling methods and datasets, as well as calculations workflow information are important and need to be provided to help inform their decisions.

The attendees contributed to the development of standards for evaluation by identifying gaps and needs for the further development of the evaluation framework and environment. For example:

Observational datasets. The discussions concluded that there was no need for the development or use of only one standard observational dataset to serve as a comparison baseline. A thorough inter-comparison of the observational datasets, however, was strongly supported by the majority of the attendees.

Evaluation across scales. The need for evaluation across temporal (seasonal, decadal, century long), and spatial (for user defined areas) scales was indicated.

Process evaluation. The practitioners specifically stated that they were no longer interested in just the projected numbers or in 'black box' answers. Not just digital data but meteorological process-based evaluation (Monsoon precipitation, atmospheric Rivers, etc.) and descriptions of climate change impacts of interest, such as on streamflow and forest fires was indicated as being useful to applications.

Uncertainty. Summarizing projected changes in one number is no longer the goal; uncertainty descriptions are useful and necessary. There was a clearly described need for NCPP to provide ensembles of climate data, not single evaluations.

Web Interfaces. Enhancing the usability and navigability of the web interfaces to access the evaluations through the addition of tutorials and examples was strongly indicated by the participants.

<u>Challenges</u>:

The design and implementation of the first prototype of our evaluation infrastructure identified the range of requirements needed to develop an evaluation service. These included access to observed and downscaled climate data, even when public access entry points were known. Internet bandwidth is not adequate to support transfers and handling of large amounts of fine resolution data or to support services except for small geographical areas. Differing data formats, resolutions, map projections, lack or inadequate

metadata were problems that needed to be addressed and were time-consuming tasks. The introduction of application-specific indices is a large multiplier that increases the computational and storage needs to the institutional scale. The adoption of the CIM metadata standards to downscaling methods proved more difficult than anticipated; however, our efforts exposed significant international interest, which is an effort that has grown since the workshop.

Substantial progress was made with NESII partners in all aspects of the development of an evaluation infrastructure, with unique presentation of metadata, visuals, and digital data of a portfolio of indices calculated from downscaled datasets. The prototypes released at the workshop helped to better define the usability requirements of the evaluation infrastructure. Of particular note is the need to support machine-automated, flexible definition and analysis of large ensembles. There is also the need to determine the balance of pre-computed and as needed computation of indices and metrics.

A future challenge will be in producing the end-to-end user experience that was a strong recommendation of many of the participants. The workshop helped to better define the components of the end-to-end system, and highlighted the technical, scientific and resource challenges. This once again highlights the need to work as a member of a participatory community to develop services that are relevant, extensible and sustainable.

B) Translational information, guidance and metadata

Prior to the workshop:

The area of 'translational information' is a new and relatively uncharted territory. While various forms of guidance have existed in the past in the form of reports and journal articles, it was clear from the discussions that these are not adequately answering the questions that practitioner have. Led by the understanding that there is a need for information that translates digital climate data for use in applications, NCPP made headway in defining 'translational information' and the various elements it encompasses. In preparations for the workshop NCPP identified examples of translational information and guidance (such as guidance on identifying the need for downscaled data, analysis pathways for climate change assessments, and communication ideas such as the "nutrition label" for downscaled datasets). These ideas energized the breakout discussions during the second half of the workshop and allowed identification of the needs of users and climate scientists for translation and interpretation of climate information.

An additional focus of NCPP and NESII efforts related to the provision of translation and interpretation of climate information prior to the workshop was the development of metadata for the description of downscaling methods and downscaled datasets. A standardized description of downscaling methods and datasets allows for traceability, for adequate comparisons and interpretation, and ultimately for better-informed user decisions. We have decided to use the same metadata framework that is used by CMIP5 to describe GCMs, with the idea of being able to link downscaled and GCM-level model descriptions. Working towards the provision of standardized metadata for downscaling methods and datasets, NCPP extended the CIM metadata schema, which was developed initially by the ES-DOC project to accommodate descriptions of CMIP5 GCMs. NCPP and NESII also incorporated the existing atmosphere and land controlled vocabularies for GCMs in a tool, the CIM regional model component Questionnaire, which allows the development of descriptions of Regional Climate Models (RCMs). In collaboration with GFDL, NASA, and the international community, NCPP proposed a set of global attributes to be included when publishing downscaled data on ESGF nodes. These

global attributes were used by GFDL, NCPP and NASA to publish downscaled datasets on their respective ESGF nodes.

To obtain a better perspective on the needs for translation and interpretation, as well as to solicit guidance on evaluation relevance for applications, NCPP established working groups for several sectors – water resources, human health, agriculture and ecological impacts. The second half of the workshop provided opportunities for more in depth discussions on the gaps and needs for climate information for these applications, as well as the needs for interpretation and guidance on the appropriate use of that information. Some of the groups were engaged in discussions prior to the workshop about needs and time scales of climate information, necessary variables and indices, specific questions related to climate change impacts, while other used productively the time during the workshop to specify needs for evaluation and climate information.

Outcomes and lessons learned:

During the workshop NCPP facilitated a productive dialog between climate scientists and practitioners, which lead to community-identified gaps and needs for the development of translational information and metadata. The participants indicated strongly the need for interpretation and translation. Given the wealth of information available through the ESGF stored evaluations the workshop participants requested guidance on how to interpret these evaluations? They specifically indicated that there is a need for structured guidance on the choice of downscaled data with highest credibility for their applications and area – a "Practitioner's guide" as some called it, or a standardized template supporting the selection of a particular suite of downscaled climate projections. Choosing a "stable ensemble" of projections -- giving similar results when two people have similar questions for a similar area -- was indicated as expected and highly desirable in order for the information to be deemed useful.

The participants defined types of translational and guidance information that is needed to inform their decisions such as explaining how the downscaled data are to be used appropriately; interpreting results with uncertainty; succinct summaries of the advantages and disadvantages of downscaled datasets in a format similar to a nutrition label; representative case studies with a common but flexible template for applications; user provided comments on datasets' usefulness and applicability; explanation of climate change and climate processes and their impacts on the applications of interest.

Recognizing the variety of users that need climate information and guidance, the participants indicated the need for provision of tiers of translational information. They also identified a structure for the translational information to be provided by NCPP, for example, providing user guidance to address high-level issues related to the need for downscaled climate data before diving in detail descriptions and guidance. In addition, they indicated that the usefulness of the guidance would increase if it is organized by intended use – education/outreach; risk screening; adaptation options appraisal (scenario ranges).

Finally, the attendees indicated that there is a need for education of practitioners and identified topics of interest such as need for downscaling; types of downscaling methods; advantages and disadvantages of dynamical versus statistical downscaling; considerations for use of downscaled climate information for impact assessments – for example, do you need downscaled information? , how the evaluation information will help with interpreting future projections?

Challenges:

The workshop helped to define usable types of translational information. The workshop also re-affirmed the enormous complexity and challenges to define and organize translational information. There are also gaps in the information technology to manage translational information and link qualitative and quantitative information. The workshop also identified scientific challenges in the development of guidance. In particular, there is a need for more work on how to connect the evaluations to the development of ensembles of climate projections for a specific use.

C) Communities of Practice

Prior to the workshop:

From the beginning NCPP recognized the need for participatory communities to address the scope and complexity of the usability of climate information. The goal is to facilitate the start of communities of practice with the premise that successful practices will emerge and have broader organizing impact. Prior to and during the workshop NCPP made active steps towards building such a community of practice. Notably with our Climate Science Applications Team and Executive Board, we framed the need for evaluation of downscaled data as the "practitioners dilemma" (http://onlinelibrary.wiley.com/doi/10.1002/2013EO460005/pdf). NCPP established and initiated contacts with working groups representing water resources, ecological, agricultural and human health related impacts of weather and climate. These working groups will continue to be engaged to define and inform the future development of translational information and information technology tools that will build capacity to improve the usability of climate knowledge. The invited attendees included many of the most experienced practitioners from the four applications areas water resources, ecological impacts, agricultural impacts, human health related impacts of weather and climate (e.g., the representatives of the North Central Climate Science Center, the Center for Disease Control, USDA); prominent representatives of the dynamical (NARCCAP, University of Maryland) and statistical downscaling community (e.g., "Perfect model" experiments – NOAA GFDL team, ARRM method - Texas Tech University, BCSD and BCCA methods – Santa Clara University, MACA method – University of Idaho); informational technology experts (e.g., from NASA JPL RCMES, NOAA, USGS); decision-makers (e.g., San Francisco Public Utilities, Denver Water, Tampa Bay Water Utilities). In addition to the attendees from the US, foreign participants from Korea, Japan, and France, also took part in the workshop. Christian Page who represented the IS-ENES2, CORDEX-Europe and the COST-VALUE projects from Europe provides valuable and ongoing connections to the emerging European efforts on evaluation of downscaling.

Outcomes and lessons learned:

As was indicated in the exit survey, the workshop was effective at bringing the various communities together to communicate and to collaborate and provided an effective platform for community co-development of standards and sharing of knowledge. NCPP actively promoted the interactions between the GFDL "Perfect Model" experiments team and the community of practitioners and climate scientists, which resulted in identification of several collaborative opportunities. NCPP initiated the development of new working groups related to downscaling and technical infrastructure, which led to identification of opportunities for multi-institutional collaborations. The QED workshop site on CoG and other tools developed by the NESII and NCPP teams provided the technical infrastructure in support of the interactions between the participants throughout the workshop. Some of the attendees became members of the various related projects on CoG – (ES-DOC, Downscaling Metadata, GFDL Perfect Model, for ex.), which will allow them to contribute further to the work of NCPP.

A major outcome from the workshop was the community-driven definition of where the tools, evaluation and translation services envisioned at the start of the NCPP effort most benefitted the supply chain linking climate-knowledge providers and climate-model users. The recommended focus for NCPP is the boundary organization of translators such as RISAs, CSCs, USDA hubs, which have and are building relationships with networks of practitioners. In addition to these boundary organizations there is a set of organizations, which are at the bleeding-edge of trying to incorporate climate science into their plans. These organizations are also critical customers for NCPP. A strong case was made that what is needed is for these boundary organizations to have a place where standard, objective evaluation has been performed, a level of guidance has been provided and practices and procedures have been transparently documented.

With the realization that without regard to limited capital and human resources, that the complexities of the climate change and climate adaptation science cannot be served by a single entity, continued focus on community and partnerships is mandatory.

Challenges:

Although NCPP made significant strides towards initial involvement of the community of practitioners and scientists during the workshop, the strategy for continued engagement needs better definition and even stronger commitment on part of NCPP.

3. Conclusions:

We prepared for the workshop guided by our ideas about community co-development of evaluation and translation standards. The participants reiterated that evaluation of downscaled data and indices is relevant and necessary. As indicated in the exit survey, 72% of the responders agreed that NCPP's approach provides the foundation to address the community's need for evaluation. The workshop attendees helped identify gaps and needs to improve usability and to advance development of the evaluation framework and environment as well as for future development of translational information and metadata. All of the achievements prior and during the workshop were a direct result of the productive collaboration of climate scientists, technical teams and representatives from the various applications.

Bringing groups of developers, providers and users of climate information together and bridging the gap of understanding and communication is not an easy process. However, the week of the workshop was exciting; the participants were very engaged, supportive and enthusiastic about NCPP and the demonstrated capabilities, as well as about the opportunities to contribute to the exchange of knowledge, to the development of standards in evaluation and translation, and to the establishment of lasting collaborations. The legitimacy of this development was enhanced by the openness of the process and by the productive interactions of the different communities. The collaborative engagement of four applications related groups as well as of climate scientists and infrastructure developers increased the salience of the evaluation framework development.

NCPP will use the outcomes from the workshop to inform the development of future directions and implementation plans. NCPP proposes to focus our work in these main areas: data and metadata access, evaluation and screening tools, translation and interpretation tools, and community development of standard and effective practice. We anticipate sustained projects with members of the working groups and other attendees at the workshop.