Weather and Climate Extremes 10 (2015) 10–28

Contents lists available at ScienceDirect

Weather and Climate Extremes

journal homepage: www.elsevier.com/locate/wace

# Interpreting climate model projections of extreme weather events

Stephen J. Vavrus *, Michael Notaro, David J. Lorenz

*Nelson Institute Center for Climatic Research, University of Wisconsin-Madison, 1225 W. Dayton Street, Madison, WI 53706, USA*

ARTICLE INFO

*Article history:*
Received 31 December 2014
Received in revised form
4 September 2015
Accepted 24 October 2015
Available online 28 October 2015

*Keywords:*
Climate Model
Uncertainty
CMIP
Downscaled
Extremes

ABSTRACT

The availability of output from climate model ensembles, such as phases 3 and 5 of the Coupled Model Intercomparison Project (CMIP3 and CMIP5), has greatly expanded information about future projections, but there is no accepted blueprint for how this data should be utilized. The multi-model average is the most commonly cited single estimate of future conditions, but higher-order moments representing the variance and skewness of the distribution of projections provide important information about uncertainty. We have analyzed a set of statistically downscaled climate model projections from the CMIP3 archive to assess extreme weather events at a level aimed to be appropriate for decision makers. Our analysis uses the distribution of 13 global climate model projections to derive the inter-model standard deviation, skewness, and percentile ranges for simulated changes in extreme heat, cold, and precipitation by the mid-21st century, based on the A1B emissions scenario. These metrics provide information on overall confidence across the entire range of projections (via the inter-model standard deviation), relative confidence in upper-end versus lower-end changes (via skewness), and quantitative uncertainty bounds (derived from bootstrapping).

Over our analysis domain, which covers the northeastern United States and southeastern Canada, some primary findings include: (1) greater confidence in projections of less extreme cold than more extreme heat and intense precipitation, (2) greater confidence in relatively conservative projections of extreme heat, and (3) higher spatial variability in the confidence of projected increases in heavy precipitation. In addition, we describe how a simplified bootstrapping approach can assist decision makers by estimating the probability of changes in extreme weather events based on user-defined percentile thresholds.



## 1. Introduction

Utilizing climate model projections can be challenging for both climatologists and decision makers. Projections from a set of models often exhibit considerable scatter and may even differ on the sign of a future climate change, such as whether a location will become wetter or drier. For many scenarios, there may be consensus on the direction of change (e.g., a warming climate due to greenhouse forcing), but the models may differ greatly on the projected magnitude of the change. Improvements in the quality of climate models have not rectified these discrepancies, as evidenced by the similar spread among projections between the newer Coupled Model Intercomparison Project phase 5 (CMIP5) and the older CMIP3 collection (Knutti and Sedláček, 2013). Differences among the projected changes in extreme weather are of particular relevance for decision makers, because of the disproportionate socioeconomic impact these events exert.

The need to usefully interpret inconsistent model simulations has spurred numerous assessment efforts aimed at quantifying uncertainty in projections and increasing the reliability of projections. The simplest and most widely reported metric is the arithmetic average or multi-model mean among a set of model simulations, which provides the most relevant single piece of guidance on expected change. Information on uncertainty has often been expressed in terms of basic metrics, such as the range of projections among all models (Scherrer and Baettig, 2008) or the inter-model standard deviation of a projection (Maloney et al., 2013). Common alternative approaches are to identify where a vast majority of models agree on the sign of a change (Meehl et al., 2007) and somewhat more refined versions that depict a combination of high inter-model agreement within regions of significant changes (Kirtman et al., 2013; Knutti and Sedláček, 2013). At the other end of the complexity spectrum are highly advanced statistical approaches, such as Bayesian methods (Tebaldi et al., 2005), hierarchical statistical models (Cressie and Wikle, 2011), and Reliability Ensemble Averaging (Giorgi and Mearns, 2002). These more sophisticated strategies are useful in their own right,

* Corresponding author.
*E-mail addresses:* sjvavrus@wisc.edu (S.J. Vavrus), mnotaro@wisc.edu (M. Notaro), dlorenz@wisc.edu (D.J. Lorenz).

http://dx.doi.org/10.1016/j.wace.2015.10.005
2212-0947/

but they are unlikely to be adopted by managers seeking practical guidance on how to digest model information for the purpose of decision making. Regardless of the method, uncertainty assessments are useful for indicating how much confidence should be placed in the inter-model average, but directly translating such information for a particular application is not straightforward.

An alternative approach is to winnow a set of projections by giving more credence to models considered to be the most accurate and downgrading the others. This intuitively satisfying strategy has been explored in a number of studies (e.g., Georgi and Mearns, 2002; Murphy et al., 2004; Annan and Hargreaves, 2010) and has been applied widely in an attempt to optimize various projections (Schmittner et al., 2005; Chapman and Walsh, 2007; Wang and Overland, 2012). Unfortunately, these efforts have been hindered by the lack of a theoretical justification for weighting model projections and by the practical difficulties in doing so (Knutti, 2010), in part because there is no clear relationship between a model's skill in simulating past climate and the magnitude of its projected changes (Knutti et al., 2010). Furthermore, deriving uncertainty information from a weighted model-mean poses additional statistical challenges, and basing a weighting scheme on smaller sample sizes constituting extreme events creates further difficulties.

In this study, we aim to strike a practical balance between simple and complex methods of quantifying uncertainty in climate projections for use by decision makers. Our focus is on extreme weather events, because of their severe societal impacts and overall positive trends (Gleason et al., 2008; Karl and Katz, 2012; Walsh et al., 2014). We concentrate on short-term (daily) extremes of heat, cold, and precipitation over the northeastern United States and southeastern Canada, based on statistically downscaled projections. Although the methodology used here can be applied generally, we focus on simulated climate change by the mid-21st century, a time period of increasing relevance for practical decision-making. Our three primary statistical measures to characterize uncertainty in a set of model projections are considered to be basic to intermediate in complexity: standard deviation (converted to the coefficient of variation), skewness, and percentile ranges derived from bootstrapping. The primary goal of this study is to provide an assessment of projected changes in extreme weather that is relevant for decision makers. We focus on *statistical* uncertainty among model projections, each of which is deemed equally plausible, rather than addressing the underlying causes of model sensitivity that might explain why the projections differ from each other.

## 2. Data and methods

We use a high-resolution (0.1°) data set of daily maximum temperature, minimum temperature, and precipitation that was statistically downscaled from 13 global climate model (GCM) simulations included in CMIP3 (Table 1). We focus on the late-20th century (1961–2000) and mid-21st century (2046–2065), based on the "middle-of-the-road" A1B emissions scenario, along with a supplemental analysis using the lower-emission B1 scenario. Our downscaled projections are an improved and spatially expanded version of a data set originally covering the state of Wisconsin, which has been widely used for a variety of climate change studies and assessments (WICCI, 2011; Notaro et al., 2011, 2012; Veloz et al., 2012; Vavrus and Behnke, 2013). Details of the downscaling procedure are given in Notaro et al. (2014), but an important feature is that a particular large-scale atmospheric pattern does not yield a unique temperature or precipitation value at the surface. Instead, the downscaling is probabilistic by virtue of parameter values translated into a probability density function that varies in time and space according to large-scale atmospheric fields. In addition, the late-20th century model output was debiased, following the cumulative distribution function algorithm of Wood et al. (2004). The same debiasing method was applied to the original Wisconsin-based downscaled data, which was found to produce an excellent match with observations of extreme daily weather events (WICCI, 2011), unlike some downscaling methods that rely on linear regression and analogs (Gutmann et al., 2014). The statistical downscaling was trained on observations from the National Weather Service's Cooperative Observer Program and Environment Canada's Canadian Daily Climate Data.

**Table 1**
List of CMIP3 GCMs used in this study, along with their original resolution before downscaling was applied. Horizontal resolution is listed in degrees of latitude/longitude, translated into approximate values for models that use spectral truncation. Vertical resolution is expressed by the number of levels (L) used by the model.

| # | Model Name | Country | Horizontal resolution (deg) | Vertical resolution |
|---|---|---|---|---|
| 1 | CGCM3.1 (T47) | Canada | 3.75 | L31 |
| 2 | CGCM3.1 (T63) | Canada | 2.8 | L31 |
| 3 | CNRM_CM3 | France | 2.8 | L45 |
| 4 | CSIRO_mk3.0 | Australia | 2.8 | L18 |
| 5 | CSIRO_mk3.5 | Australia | 2.8 | L18 |
| 6 | GFDL_CM2.0 | United States | 2.5 | L24 |
| 7 | GISS_AOM | United States | 3.5 | L12 |
| 8 | IAP_FGOALS | China | 2.8 | L26 |
| 9 | MIROC3.2 (medres) | Japan | 3.5 | L20 |
| 10 | MIROC3.2 (hires) | Japan | 1.2 | L56 |
| 11 | ECHO_G | Germany/ Korea | 3.75 | L19 |
| 12 | MPI_ECHAM5 | Germany | 2.8 | L31 |
| 13 | MRI_CGCM2.3.2a | Japan | 3.5 | L30 |

In this study, we focus on a domain approximating that of the Northeast Climate Science Center (NECSC) and associated Landscape Conservation Cooperatives (LCCs) that are part of the United States Department of the Interior (Fig. 1). The NECSC is one member of a federal network of eight regional centers created to provide scientific information, tools, and techniques that managers can use to anticipate, monitor, and adapt to climate change. The use of this relatively small area in our study allows more in-depth analysis of spatial variations and makes absolute temperature and precipitation thresholds of extremes more meaningful. This domain encompasses several of the LCCs, the primary stakeholders and partners of the Climate Science Centers, whose mission is to connect scientific information with on-the-ground conservation

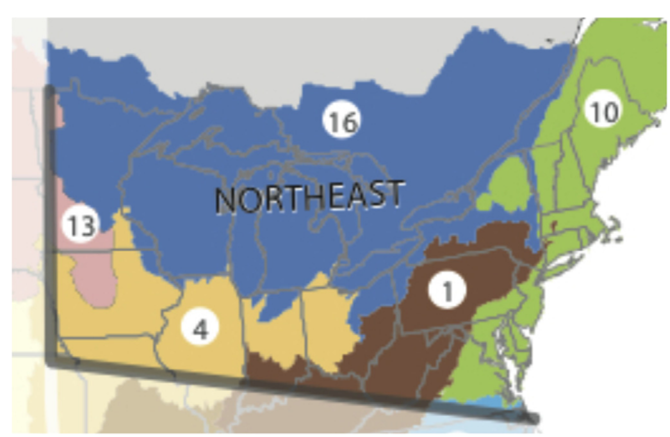


**Fig. 1.** Domain used in this study, encompassing the Northeast Climate Science Center area (thick gray lines) and several Landscape Conservation Cooperatives (numbered): (1) Appalachian, (4) Eastern Tallgrass Prairie and Big Rivers, (10) North Atlantic, (13) Plains and Prairie Potholes, and (16) Upper Midwest and Great Lakes.