Environmental Modelling & Software 110 (2018) 38–51

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

Downscaling of climate model output for Alaskan stakeholders

John E. Walsh [a, *], Uma S. Bhatt [b], Jeremy S. Littell [c], Matthew Leonawicz [d], Michael Lindgren [d], Thomas A. Kurkowski [d], Peter A. Bieniek [e], Richard Thoman [f], Stephen Gray [c], T. Scott Rupp [e]

[a] Alaska Center for Climate Assessment and Policy, University of Alaska, Fairbanks, AK, USA
[b] Department of Atmospheric Sciences, Geophysical Institute, University of Alaska, Fairbanks AK, USA
[c] USGS Alaska Climate Adaptation Science Center, Anchorage, AK, USA
[d] Scenarios Network for Alaska and Arctic Planning, University of Alaska, Fairbanks, AK, USA
[e] Alaska Climate Science Center, University of Alaska, Fairbanks, AK, USA
[f] National Weather Service/NOAA, Fairbanks, AK, USA

ARTICLE INFO

Article history:
Received 20 May 2017
Received in revised form
20 February 2018
Accepted 30 March 2018
Available online 30 April 2018

Keywords:
Climate modeling
Downscaling
Visualization
Temperature
Precipitation
Alaska

ABSTRACT

The paper summarizes an end-to-end activity connecting the global climate modeling enterprise with users of climate information in Alaska. The effort included retrieval of the requisite observational datasets and model output, a model evaluation and selection procedure, the actual downscaling by the delta method with its inherent bias-adjustment, and the provision of products to a range of users through visualization software that empowers users to explore the downscaled output and its sensitivities. An additional software tool enables users to examine skill metrics and relative rankings of 21 global models for Alaska and six other domains in the Northern Hemisphere. The downscaled temperatures and precipitation are made available as calendar-month decadal means under three different greenhouse forcing scenarios through 2100 for more than 4000 communities in Alaska and western Canada. The visualization package displays the uncertainties inherent in the multi-model ensemble projections. These uncertainties are often larger than the projected changes.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

The rapid rate of climate warming in Alaska (Thoman and Brettschneider, 2016; Overland et al., 2018) and its consequences (USGCRP, 2014) have created a need for products to help plan for the future. Global climate models run with different greenhouse gas scenarios provide climate scientists with projections of the expected large-scale response to anthropogenic climate change. However, regional changes are not well resolved in these low-resolution models, precluding the detailed landscape level projections often required for understanding impacts on local communities and resources. To date, the downscaled products available for planning and adaptation in Alaska are severely limited. The goal of this paper is to document the development and characteristics of downscaled (finer resolution) products and associated visualization tools recently made available in Alaska.

Development of downscaled climate for Alaska has historically been limited both by a challenging physical geography and by data limitations. Alaska is a particularly difficult region to model, with tall mountains, long complex coastline (Fig. 1), a landscape surrounded by seasonally varying sea ice, and large seasonal swings in temperature, all of which contribute to strong gradients in temperature and precipitation (Fig. 2). Coarse-resolution global climate models (GCMs) do not adequately represent these influences on temperature and precipitation at the landscape level, so downscaling of the GCM information is necessary to provide stakeholders and decision makers with tools to address practical problems such as how climate change will affect local water resources, land use and infrastructure. Downscaling also enables correction of model biases if the downscaling is keyed to historical observational data, although the lack of long-term observations for a variable and location for which downscaling is desired can limit options for downscaling. This limitation is especially problematic for quantities that are not routinely measured, e.g., solar radiation, soil moisture, snow water equivalent.

* Corresponding author. International Arctic Research Center, 2158 Koyukuk Drive, Fairbanks, AK 99775, USA.
E-mail address: jewalsh@alaska.edu (J.E. Walsh).

https://doi.org/10.1016/j.envsoft.2018.03.021
1364-8152/© 2018 Elsevier Ltd. All rights reserved.

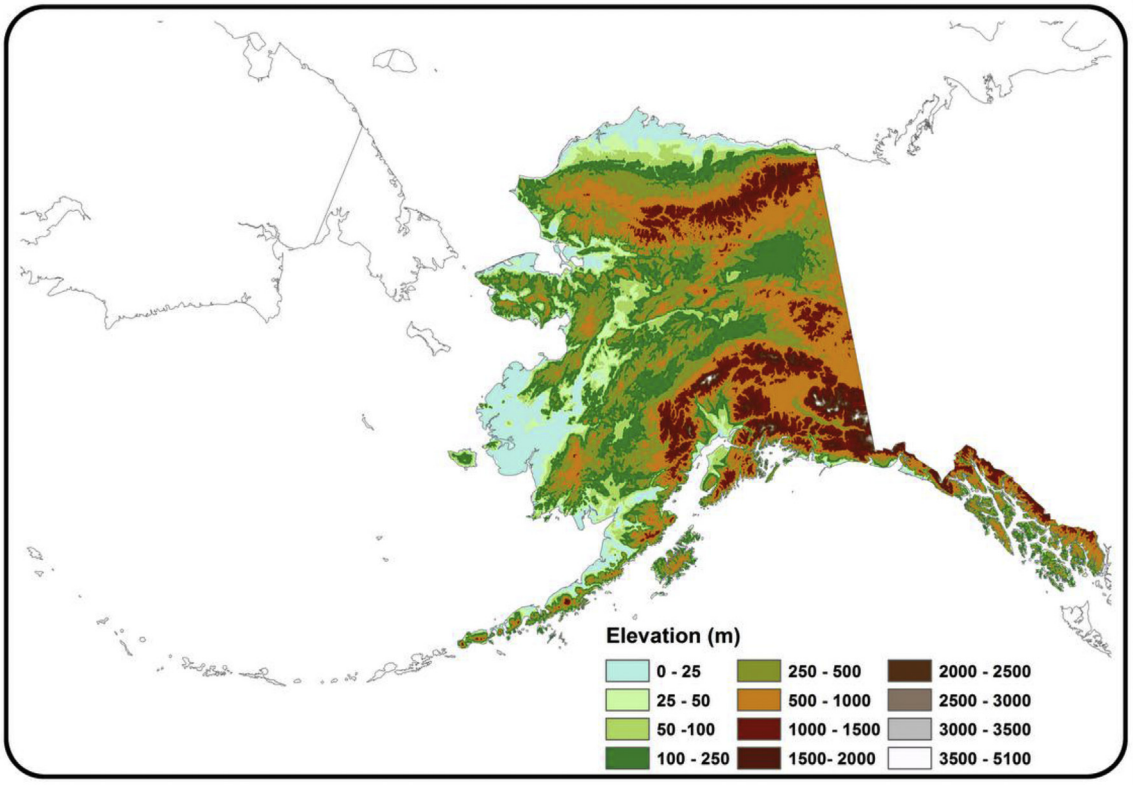

Fig. 1. Topography and coastal configuration of Alaska. Source: Scenarios Network for Alaska and Arctic Planning (SNAP), http://data.snap.uaf.edu/data/IEM/Inputs/ancillary/elevation/iem_prism_dem_1km.tif, modified by J. Littell.

Over the last decade, summaries of global climate model output available for the region have enabled coarse-resolution estimates of regional change. Both temperature and precipitation are expected to increase over Alaska, and as with most high-latitude regions, model agreement on the sign of the precipitation change is favorable. Regional projected changes in temperature and precipitation were calculated for a region (60°-72°N, 103°-170°W) including Alaska based on output from the CMIP3 (Climate Model Intercomparison Project, version 3) generation of models (Christensen et al., 2007). By 2080–2099 (relative to 1980–1999 and for the A1B emissions scenario, across 21 GCMs), they indicated a median annual temperature increase of +4.5°C (range +2.7 °C to +6.4 °C), with greater increases in winter (median +6.0 °C) and autumn (median +4.8 °C) than spring (median +3.7 °C) and summer (median +3.0 °C). The models projected an increase in precipitation in all seasons, with a median annual increase across models of 21% (range +6% to +32%), with more (+28%) in winter and less (+14%) in summer. A global summary of the CMIP5 (Coupled Model Intercomparison Project, version 5) output provided by Collins et al. (2013) for the IPCC (Intergovernmental Panel on Climate Change) Fifth Assessment included sub-regional details. The CMIP5 ensemble average temperature increases for the RCP (Representative Concentration Pathway) 8.5 scenario are generally higher for Alaska compared to global values: they range from +4-5 °C in southeast/Aleutian Alaska to +8°C or more for the North Slope of Alaska. Annual precipitation increases projected for CMIP5 are generally similar throughout Alaska (~+15% for Southeast Alaska and the Aleutians, +20% for the Interior and Yukon-Kuskokwim delta, +30% for the North Slope). While these projections provide useful, coarse-scale information, the model output lacks the details of the topographic and coastal influences that can be important for users.

One of the two primary methods of transforming coarse-resolution climate information to high resolution is statistical downscaling. The goal of statistical downscaling is to reproduce local climate averages over timescales of a decade to several decades. This requires long-term high quality observational data to develop 'training' relationships between coarser-resolution model-derived variables and local conditions. A statistical relationship is established between large-scale climate and observed local variables (temperature, precipitation, winds) over a 'training' period. This method allows downscaling to a local point at whatever time step is most finely resolved by the local observations, typically monthly or daily. The most common statistically downscaled variables are temperature and precipitation, although winds, relative humidity, ocean water temperature, and snow water equivalent have also been downscaled statistically. The procedure implicitly includes a bias-correction of the model output. The so-called "delta" method, used here and described in Section 3.2, obtains a bias correction from model output and corresponding local observations for a historical period; the same correction is then applied to the model's future output for the particular location. Other statistical downscaling methods exist, including several variants of quantile-mapping (e.g., Maurer et al., 2007). Hayhoe (2010) found that the statistical downscaling is most sensitive to the driving GCM, secondly to the statistical method, followed by the evaluation metric. Statistical downscaling is relatively computationally inexpensive, allowing many models/scenarios to be downscaled, and the methods are generally straightforward. The key weakness is that one has to assume the statistical relationship developed on the historical data will not change in the future. This method also requires a robust

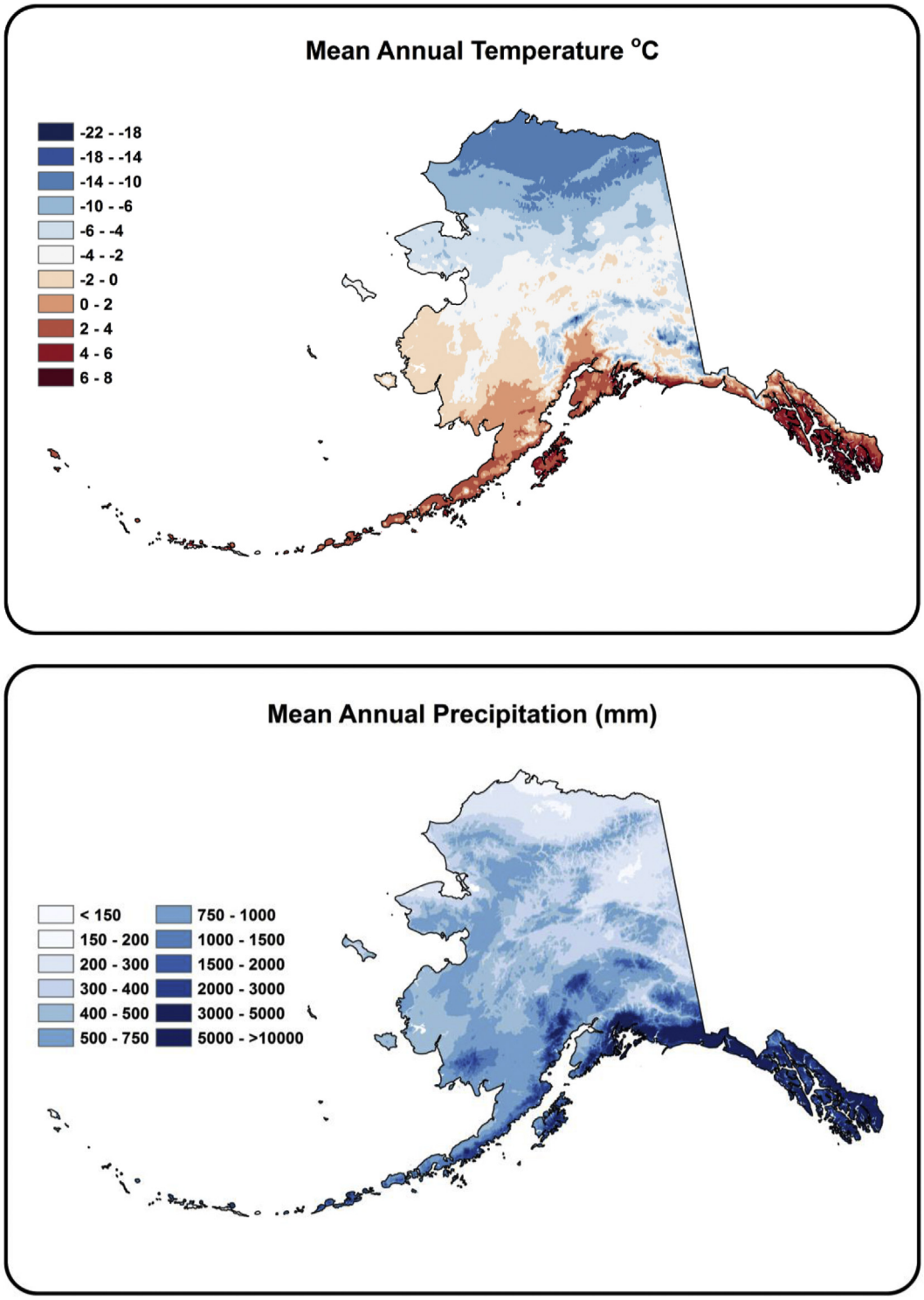

Fig. 2. 1950–1999 mean annual mean temperature (upper) and total precipitation (lower) for Alaska based on downscaled CRU TS 3.1 (Scenarios for Alaska and Arctic Planning, University of Alaska, Fairbanks, after Harris et al., 2014). Alaska station data availability in CRU was most consistent for the 1950–1999 timeframe.

historical climatology based on observations, which are not always readily available.

Dynamical downscaling has the same ultimate goal as statistical downscaling – a finer resolution climate scenario – but employs a regional climate model forced at the boundaries by the large-scale climate model rather than relying on statistical relationships. Dynamical downscaling for Alaska has been conducted using lateral boundary forcing from reanalysis output (e.g., Bieniek et al., 2016; Bhatt et al., 2007) as well as historical and future output from climate models (Zhang et al., 2007a, 2007b; Lader et al., 2017). Dynamical downscaling provides physically consistent projections of many variables, and therefore sufficient data to explore future

climate variability mechanisms. This method is computationally expensive, limiting the number of different models/scenarios that can be downscaled. Dynamical downscaling is also a complex process requiring a relatively high level of modeling expertise to conduct. Biases and other errors in the models are also problematic in dynamical downscaling.

The present paper describes statistical downscaling for Alaska, with an extension of the products into western Canada. It complements and extends previous uses of statistical and dynamical downscaling of global model output for the contiguous United States. For example, the Bureau of Reclamation (2013) supported a downscaling of monthly temperature and precipitation covering the contiguous 48 states at 1/8° resolution for a historical period (1970–1999) and three 30-year future time slices spanning 2010–2099. Similarly, NASA (National Aeronautics and Space Administration) NEX-GDDP (https://cds.nccs.nasa.gov/nex-gddp/) provides daily globalcoverage at 1/4° resolution for an historical period (1950-2005) and continuous projections from 2006-2100 developed using Bias-correction spatial disaggregation (BCSD) methods after Thrasher et al. (2012). To derive higher resolution data for regional climate change assessments, NASA coordinated a statistical downscaling of maximum and minimum air temperature and precipitation from 33 of the CMIP5 climate models to a very fine 800-m grid over the contiguous United States. The product, known as the NEX-DCP30 dataset (https://cds.nccs.nasa.gov/nex/), covers the historical period (1950–2005) and 21st century (2006–2099) under four Representative Concentration Pathways (RCP) emissions scenarios developed for the IPCC's Fifth Assessment Report (AR5). A supporting visualization tool, the National Climate Change Viewer (NCCV), was developed by the USGS (https://www2.usgs.gov/climate_landuse/clu_rd/nccv.asp). The North American Regional Climate Change Assessment Program (NARCCAP) is a dynamical downscaling activity in which regional climate modeling groups performed a coordinated set of high-resolution simulations of North American climate (Mearns et al., 2009). However, the NARCCAP domain boundary passes through the middle of Alaska, placing the state in the buffer zone where the coarse-resolution global model heavily influences the regional model's solution. The Coordinated Regional-climate Downscaling Experiment (CORDEX) has also performed dynamical simulations for an Arctic domain that includes Alaska (Koenigk et al., 2015), although the broader Arctic domain necessitates a resolution of 20–50 km.

The statistical downscaling described here represents a twofold extension of the activities summarized above. First, it extends the downscaling to Alaska, which was not part of the domain of the fine-scale products produced for the rest of the U.S. Second, the downscaling targets communities in Alaska (as well as western Canada) by including a visualization tool for the display of the historical climate and projected changes for more than 4000 specific communities. These communities range from small villages with fewer than 100 people to major population centers such as Fairbanks and Anchorage, where the population exceeds 300,000. The intent of the project was to develop an end-to-end system of climate downscaling, connecting the global modeling enterprise with decision-makers and other users in specific locations.

The downscaling was performed by the Scenarios Network for Alaska and Arctic Planning (SNAP) at the University of Alaska, Fairbanks. It utilized the output of the global models that participated in the Coupled Model Intercomparison Project, version 5 (CMIP5). The downscaling project had three main components: (1) selection of a subset of the CMIP5 models to be downscaled for Alaska, (2) statistical downscaling of the coarse-resolution global model output to a fine-scale grid with 2 km resolution, and (3) the development of the visualization tool that displays output for the 2 km × 2 km pixel corresponding to the particular community selected by a user. In the following sections, we describe these three components.

2. Data and models

Several historical databases were used in the model evaluation and in the downscaling. The European Center for Medium-Range Weather Forecasting's ERA-40 reanalysis provided the observationally-based fields for the model evaluation. The ERA-40 reanalysis spans 45 years (1958–2002) and was available on a horizontal grid with 2.5° resolution in latitude and longitude. The ERA fields used in the model evaluation were surface air temperature, precipitation and sea level pressure.

For the downscaling of the global climate models, two station-based datasets of temperature and precipitation provided the historical climatologies, giving users of the downscaled products the option to choose the database on which the downscaling was based (Section 4). The first database is the PRISM climatology for Alaska (Daly et al., 2008 and subsequent updates). PRISM consists of calendar-month climatologies (1961–1990) of temperature and precipitation with a spatial resolution of 2 km over Alaska and western Canada. PRISM grids represent spatial interpolations of station data, taking into account elevation changes and lapse rates. Finer spatial scale PRISM products exist for Alaska (771 m, 1971–2000), but for consistency of the results for western Canada and Alaska, we used the 2 km PRISM grids in this project. The second database is the University of East Anglia Climate Research Unit's CRU TS 3.2, in which monthly station observations of temperature and precipitation have been binned into grid cells at a resolution of 0.5° latitude x 0.5° longitude (https://crudata.uea.ac.uk/cru/data/hrg). The historical climatological of the two databases differ slightly because their construction (and our interpolation of CRU TS) differed in the two cases. For example, temperature differences of a degree (°C) or so were not uncommon. For this reason, users of our downscaling tool (Section 4) can choose either option for the baseline climatology and can compare the two sets of results if they so desire.

The global climate model output is from the CMIP5 archive, which is the archive utilized in the Fifth Assessment Report (AR5) of the Intergovernmental Panel on Climate Change (IPCC, 2013). As with the observational data, the model output used here consisted of monthly surface air temperature, precipitation, and sea level pressure fields. The fields were from (1) the models' historical simulations (late 20th century, corresponding to the observational fields) and (2) the models' future simulations forced by the RCP 4.5 (low-emission), RCP 6.0 (mid-range) and RCP 8.5 (high-emission) scenarios.

Because the models in the CMIP5 archive were run at different resolutions, all fields were interpolated to a common 2.5° × 2.5° grid of the ERA-40 reanalysis. The 2.5° resolution was used for the global model evaluation and selection. The downscaled products described in Section 4 were based on an interpolation of the global model output from the 2.5° × 2.5° grids to the finer 2 km resolution of the PRISM climatology.

3. Methods

3.1. Model selection

While the CMIP5 archive includes output from more than three dozen models, several considerations led to the choice of a subset of the models for the present downscaling activity. First, the use of the

Table 1
The 21 models with archived output suitable for downscaling.

Model	Country	Model	Coun
CanESM2	Canada	GISS-E2-H	United States
CNRM-CM5	France	GISS-E2-R	United States
CSIRO-Mk3-6-0	Australia	HadCM3	U.K.
GFDL-CM3	United States	HadGEM2-CC	U.K.
GFDL-ESM2G	United States	Had GEM2-ES	U.K.
GFDL-ESM2M	United States	IPSL-CM5A-LR	France
IPSL-CM5A-MR	France	MPI-ESM-LR	Germany
MIROC4h	Japan	MRI-CGCM3	Japan
MIROC5	Japan	NCAR-CCSM4	United States
MIROC-ESM	Japan	NorESM1-M	Norway
MIROC-ESM-CHEM	Japan		

full set of 30–40 models is computationally unwieldy and tends to preclude examinations of differences among models. Second, not all models have archived the simulations (three RCP scenarios in addition to historical runs) and variables at the temporal resolution required for some downscaling applications. Of the approximately three dozen models in the CMIP5 archive, only 21 contained the needed output at the time our model evaluation was performed. These 21 models are listed in Table 1. Third, there are – at best – diminishing returns from the inclusion of models beyond a total of 10–20, in part because models rely on similar physics and are thus not entirely independent (Sanderson et al. 2015). Finally, there are indications (although not conclusive evidence) that retaining a subset of the models deemed to be "best" for a particular application can enhance the utility of the results. The latter consideration has some precedents in the literature, including some for Arctic research, but calls for caveats that we discuss below.

In previous applications to the Arctic, Wang and Overland (2009) chose a subset of CMIP3 models on the basis of their ability to capture the seasonal cycle and mean September extent of Arctic sea ice in order to optimize projections of future sea ice changes. Rogers et al. (2015) used a two-step model selection algorithm to show that the timing of an ice-free Arctic in September advances from 2055 to 2034 when the number of CMIP5 models is filtered from a full set to the subset of five models that best capture recent sea ice trends and other hindcast metrics. In an attribution study of recent Arctic temperature variations, Fyfe et al. (2013) chose a subset of five CMIP5 models on the basis of their simulations of Arctic temperature trends over three historical timeslices. The number of models retained in these studies is consistent with Walsh et al. (2008) finding, based on multimodel composites of historical Arctic simulations, that the mean absolute errors decrease as the number of best-performing models in a composite increases to 5–8, but increases as additional (poorer-performing) models are included in the composites. Nevertheless, model selection is fraught with risks because the best-performing models vary with the choice of the criterion for validation. Moreover, different models perform best for different variables, regions, and other choices in validation methodology. A case may be made that there is still merit in Knutti et al. (2010) assessment that "… there is little agreement on metrics to separate "good" from "bad" models". Given this lack of agreement, our decision to utilize only a subset of the CMIP5 models was based on the more practical considerations listed in the preceding paragraph: computational efficiency and availability of output. Our strategy was to choose the model subset on the basis of the models' ability to reproduce the seasonal cycle of the recent (historical) climate of Alaska and the surrounding area.

In evaluating the models' historical performance for the Alaskan region, the core statistic of the validation was a root-mean-square error (RMSE) of the differences between time-averaged model output for each grid point and calendar month, and the observationally-constrained ERA-40 reanalysis. ERA-40 directly assimilates observed air temperature and sea level pressure observations into a product spanning 1958 through 2002. Precipitation is computed by the model used in the data assimilation. Data from 1958 to 2000 were used here for the comparative evaluation of the global climate models (GCMs). For each of the 21 CMIP5 models, we calculated the monthly root-mean-square-error (RMSE) for each of three variables: surface air temperature, precipitation and sea level pressure. We tested the sensitivity of the model ranking to the choice of the error metric by repeating the calculations using bias-corrected RMSE, mean absolute error (MAE), and bias-corrected MAE. The bias correction removed the domain-average error from the error at each grid cell. The model selection procedure used here has been made available through a web-based application at https://uasnap.shinyapps.io/ar5eval. This application incorporates various degrees of freedom (choice of variable, domain, evaluation metric) described below. Users can select any of the four error metrics through the model evaluation web application.

Fig. 3. The seven domains for which the model evaluation was performed. Note that there are two circumpolar domains in the lower left panel: 60°-90°N and 20°-90°N.

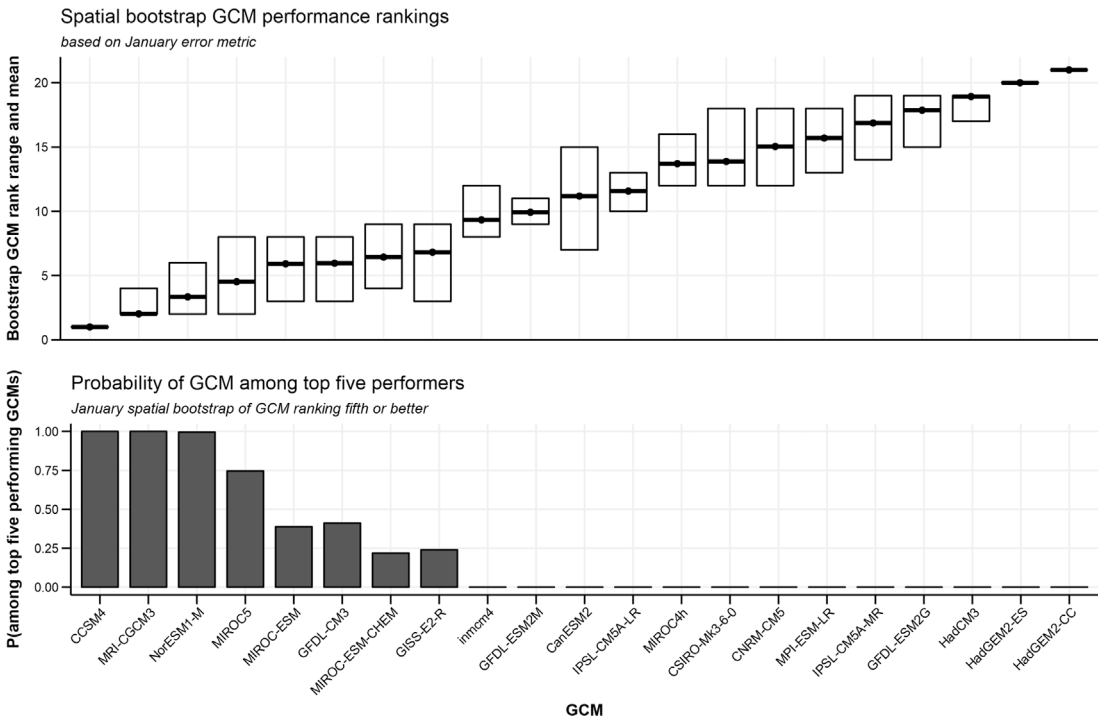

Fig. 4. Upper panel: Mean value and ranges of ranks of RMSE of January temperature based on 1000 random resamplings (with replacement) of Alaska-domain grid cells. Highest-ranking (smallest RMSE) model is on left, lowest-ranking (largest RMSE) model is on right. Lower panel: Probability (based on 1000-member resampling) that a model ranks in the top five, based on the RMSE of January temperature. Source: https://uasnap.shinyapps.io/ar5eval.

The Alaskan domain for the model evaluation covers the area 52–72°N, 130–180°W. For comparison, the same error statistics were also evaluated for the following six other domains (Fig. 3): Canada (49–72°N, 52–141°W), combined Alaska-Canada (49–72°N, 52–172°W), the 48 contiguous United States (25–49°N, 66–125°W), the Pacific Islands (17°S-25°N, 152–228°W), and two circumpolar domains: 60–90°N and 20–90°N. For each domain, the output from each model was interpolated to the 2.5° × 2.5° latitude x longitude grid of the ERA reanalysis.

The skill of the models was evaluated over all the domains of Fig. 3, and the skill over the different domains is compared in the results below. However, we focus on the Alaska domain in our illustration of the methods used for skill evaluation and model selection, as well as in the examples of the products presented in Section 4. This focus on Alaska stems from the availability of complementary downscaled information for Canada produced by the Pacific Climate Impacts Consortium (PCIC). The PCIC methods and products are accessible at https://www.pacificclimate.org/data/statistically-downscaled-climate-scenarios.

While the model evaluation procedure has some commonalty with that used by Walsh et al. (2008) to select a subset of models from the previous generation (CMIP3) of global climate models, there are several notable extensions of the procedure in the present application. First, the ranking of models was based on the models' simulations of three variables: surface air temperature, precipitation, sea level pressure rather than only the first two. Sea level pressure is a proxy for the atmospheric circulation at the surface. Second, rather than summing ranks over all calendar months and variables as in Walsh et al. (2008), the ranking was performed only after a summation of the standardized RMSEs over all calendar months and variables. Third, the robustness of the RMSEs was tested by a bootstrapping procedure in which repeated (~1000) estimates of a particular RMSE were calculated based on randomly selected grid cells (with replacement) from the domain under consideration. The number of points randomly selected was equal to the total number of grid cells in the domain. Figs. 4 and 5

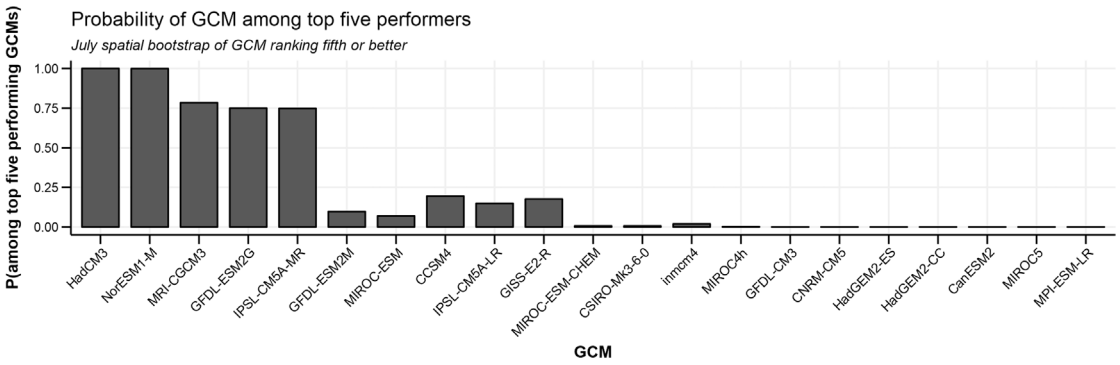

Fig. 5. As in Fig. 4, but for July precipitation over the Alaska domain.

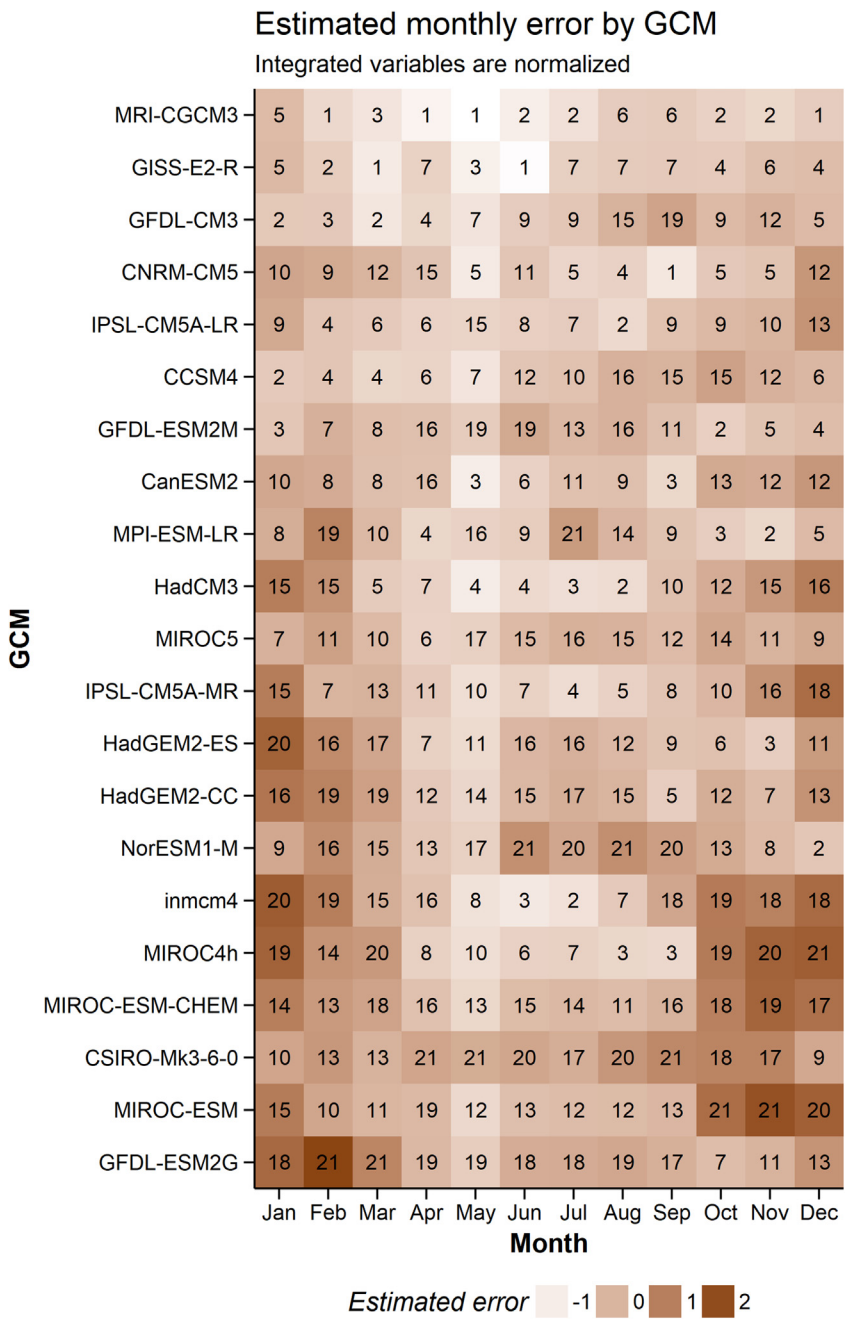

Fig. 6. Error matrix showing the relative magnitudes of the cumulative normalized error (temperature, precipitation, sea level pressure) over the Alaska domain by calendar month (x-axis) and model (y-axis). Darker shading denotes larger RMSE values. Numbers in boxes are the individual models' ranks (1 = smallest error, 21 = largest error) for the calendar month. Source: https://uasnap.shinyapps.io/ar5eval.

provides examples of RMSE values for January temperature and July precipitation from the 21 models over the Alaska domain. While the distributions for the different models overlap, there is clear separation of the models with lower RMSE versus larger RMSE, especially in the case of January temperature. In general, the distributions for the models have the least overlap for sea level pressure and the greatest overlap for precipitation. In all cases, the mean of the distributions is nearly identical to the RMSE obtained from the original (not resampled) grid. However, plots such as Figs. 4 and 5 provide a measure of the robustness of the rankings of the models. The lower panels of these figures show the probability, based on the 1000-member sample of RMSEs for each model, that a particular model will rank in the top five based on the RMSE metric for that particular variable. The probabilities are essentially 100% for the top three models in the case of January temperature and the top two models in the case of July precipitation. Beyond the 8th-ranking model for temperature and the 10th-ranking model for precipitation, the probabilities that resampling would place a model in the top five are essentially zero.

RMSE values for the three variables (temperature, precipitation and sea level pressure) were standardized and summed, and this cumulative sum was the basis for ranking the models from #1 (smallest RMSE) to #21 (largest RMSE). There was reasonably good consistency from one calendar month to the next in the relative

Table 2
Top-ranking (1st through 6th) models for various domains defined in text. Rankings are shown for two metrics: RMSE (Rood Mean Square Error) and MAE (Mean Absolute Error).

	RMSE metric	MAE metric		RMSE metric	MAE metric
Alaska:	MRI-CGCM3	MRI-CGCM3	**Alaska-Canada**:	GISS-E2-R	GISS-E2-R
	GISS-E2-R	GISS-E2-R		MPI-ESM-LR	MPI-ESM-LR
	GFDL-CM3	GFDL-CM3		CNRM-CM5	CNRM-CM5
	CNRM-CM5	CanESM2		NCAR-CCSM4	CanESM2
	IPSL-CM5A-LR	HadCM3		GFDL-CM3	CCSM4
	CCSM4	NCAR-CCSM4		IPSL-CM5A-LR	MRI-CGCM3
Canada:	MPI-ESM-LR	GISS-E2-R	**Lower 48 states**:	MPI-ESM-LR	MPI-ESM-LR
	GISS-E2-R	MPI-ESM-LR		MIROC5	MIROC5
	CNRM-CM5	CNRM-CM5		CNRM-CM5	IPSL-CM5A-MR
	NCAR-CCSM4	CanESM2		CanESM2	CNRM-CM5
	CanESM2	CCSM4		IPSL-CM5A-MR	CanESM2
	IPSL-CM5A-LR	IPSL-CM5A-LR		HadGM2-ES	HadGM2-ES
60–90°N:	MIROC4h	MIROC4h	**20–90°N:**	MPI-ESM-LR	MPI-ESM-LR
	MPI-ESM-LR	MPI-ESM-LR		GFDL-CM3	CanESM2
	GFDL-CM3	GFDL-CM3		CanESM2	MIROC4h
	CanESM2	CanESM2		MIROC4h	GFDL-CM3
	NorESM1-M	NorESM-1		CNRM-CM5	HadGEM2-ES
	GISS-E2-R	GISS-E2-R		HadGEM2-ES	CNRM-CM5

rankings of the models. Fig. 6 shows the relative errors (RMSEs) of all models in all calendar months for the Alaska domain. The individual cells in the error matrix are shaded, with the lightest cells indicating the smallest errors. The numbers in the cells are the model ranks for the calendar month. The models are listed from top to bottom according to their ranks aggregated over the twelve calendar months. These aggregate ranks formed the basis for our model selection.

The evaluation procedure was performed for the eight domains listed above. For each domain, aggregate ranks based on both RMSE and MAE (Mean Absolute Error) were evaluated. As shown in Table 2, the domains with common areas (e.g., Alaska, Alaska + Canada, 60–90°N) generally had several models in common among the best performing models (e.g., with smallest RMSEs and MAEs). There is much less overlap between the lists of best-performing models for the smaller and larger domains, reinforcing the previous caveat that the best-performing models vary by region. The choice of the error metric (RMSE vs. MAE) has only a minor effect on the rankings. Finally, although not shown here, there was also a tendency for the same models to have smaller RMSEs of all three variables in a particular domain, although there were exceptions, especially for precipitation.

As indicated by Figs. 4 and 5, several models had substantially smaller systematic errors than others. The models also vary substantially in their projections of future changes over the Alaska region (Fig. 7). This combination of historical and future spread raises the possibility that the choice of a subset of models might offer a viable approach to narrowing the uncertainty and obtaining more robust estimates of future climate change in regions such as Alaska. Subject to the caveats noted earlier, we further evaluated this strategy by examining the errors generated by compositing subsets of N models selected from the full set of 21. The model selection was done in two ways: (1) all combinations of N models and (2) the best N models based on the RMSE metric. Fig. 8 shows that the average error for a single variable (temperature, precipitation, sea level pressure) generally increases (orange lines) as one

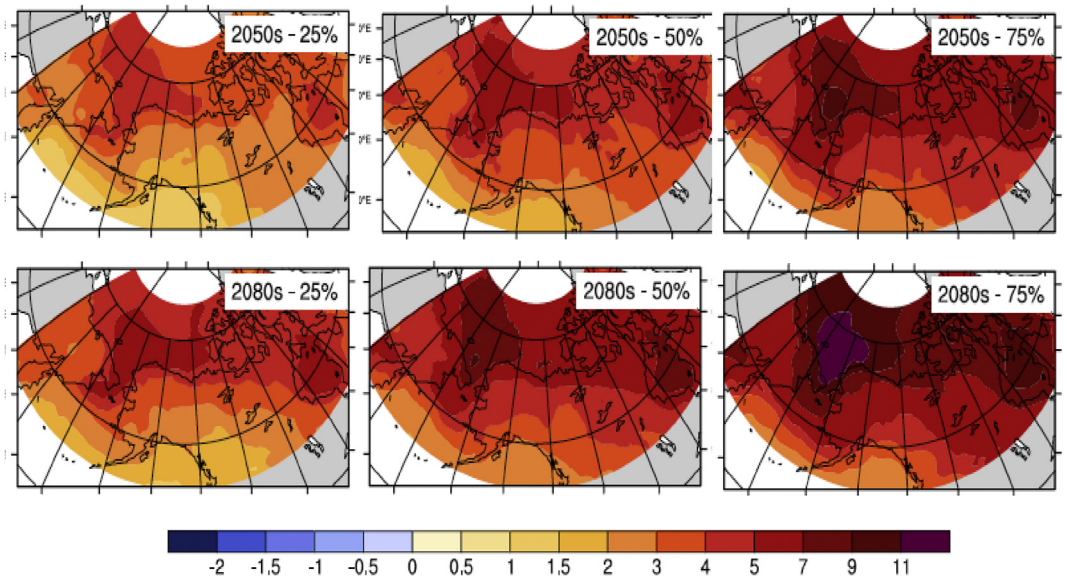

Fig. 7. Projected temperature changes over Alaska and surrounding region in the 2050s (top row) and 2080s (bottom row) under the RCP 4.5 scenario. Changes for each time slice are shown for the 25th-percentile model (left), the 50th-percentile model (center) and 75th-percentile model (right) based on a ranking of the models by the average warming over the domain. Figure provided by Greg Flato and Robin Rong, Environment Canada.

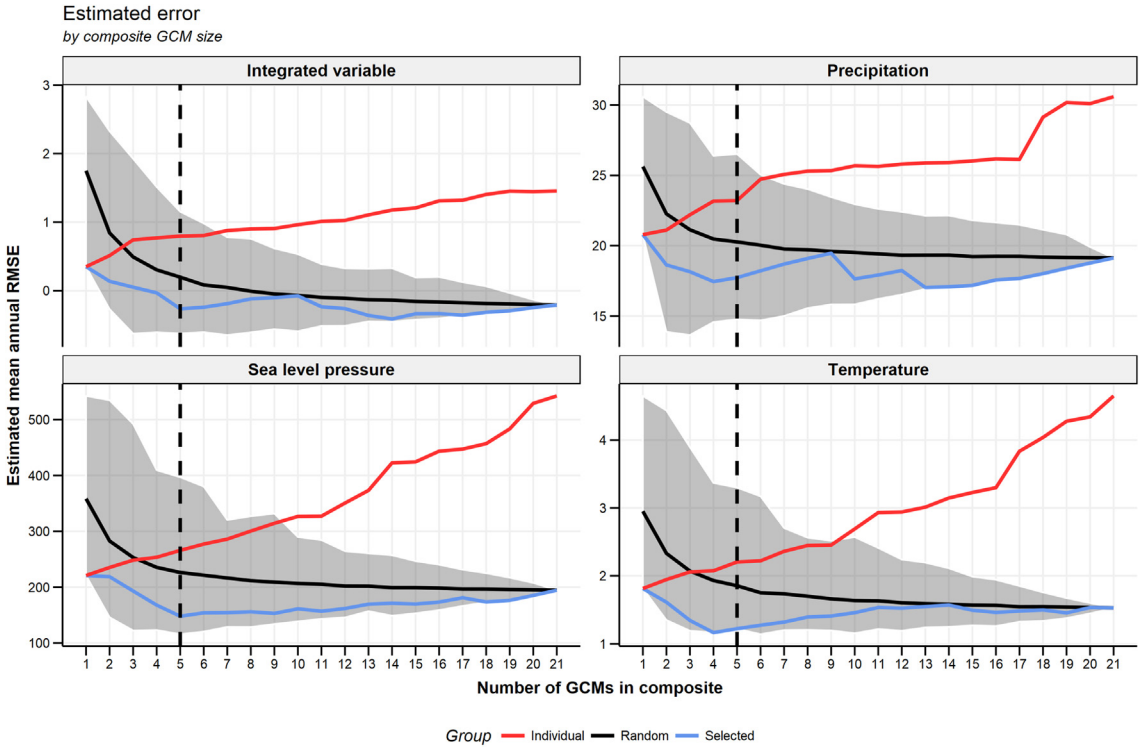

Fig. 8. RMSE for the Alaska domain as a function of the number of models in the composite for temperature (°C), precipitation (mm), sea level pressure (Pa), and the integrated (three-variable, normalized) variable. For each plot (variable), red lines show the RMSE individual 21 GCMs ranked from 1 (smallest RMSE) to 21 (largest RMSE). The narrowing gray band shows the range of RMSE among all possible composites of a given size along the x-axis, and the smooth black line running through it shows the mean RMSE among all these random composite models. The blue lines show the RMSE of the N-model composite made up of the best-performing individuals. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

moves down the list of models that rank successively lower by the aggregate metric. When the composite is based on the N randomly selected models, the error (averaged over all possible combinations of N models) decreases monotonically from N = 1 to N = 21 (black lines, with range indicated by shading). The error of the N-model composite composed of the single set of N best performing models reaches a minimum somewhere between N = 1 and N = 21 (blue lines) for the individual variables (Fig. 8a–c). While the decrease of the error with increasing N is not monotonic, there are indications of a minimum in the range of N = 4 to N = 6. However, the minimum is ill-defined for the integrated three-variable metric (Fig. 8d). The values of N at which the minimum is reached vary with region as well as with the variable. Nevertheless, even in the case of multimodality, where a single choice of optimal composite size may not be clear, there is still a prominent decrease in RMSE during the initial compositing of several models. On the basis of these results and in the interest of computational economy, we chose N = 5 for the Alaskan downscaling application.

3.2. Downscaling by the delta method

The downscaling procedure is an application of the so-called "delta" method, in which a model's future change ("delta") in a variable at a particular location and calendar month is added to the historical mean value of the same variable for the same location and calendar month. The delta is computed as the model's change from the period of the historical climatology (1961–1990 in the case of the PRISM data) to a future time slice (e.g., the 2050s). This delta is added to the higher-resolution observationally-based climatology, thereby effectively bias-correcting the model's output. A key assumption in this procedure is that the model's bias is the same in the future time slice as in the historical reference period. In all likelihood, the bias will undergo some change over time, thereby limiting the validity of the delta method (and other statistical downscaling methods). Nevertheless, the delta method has been widely used in downscaling applications, and its validity has been found to be comparable to that of more sophisticated downscaling methods when applied to monthly fields, although this is not the case for daily fields and their corresponding extremes (Hayhoe, 2010).

4. Downscaled products

4.1. Examples of community charts

Many users requiring climate information for planning or adaptation purposes are located in villages or larger population centers. Given our target of the North American Arctic, we therefore performed the downscaling for the largest available collection of community locations, covering all of Alaska, The Yukon Territory, British Columbia, Alberta, Saskatchewan, and Manitoba. This resulted in downscaled temperature and precipitation data for more than 4000 communities. In order to illustrate the fusion of the model selection and the downscaling, we focus here on Alaska and apply the downscaling methodology to the five models that ranked highest by the RMSE metric in the historical simulations across the Alaska domain (cf. Table 2): MRI-CGCM3, GISS-E2-R, GFDL-CM3, IPSL-CM5A-LR and NCAR-CCSM4. For every year and calendar month, the downscaling consisted of calculating the "delta" value for each GCM grid cell, interpolating to the same spatial resolution as the high resolution baseline climatology, followed by adding these high resolution "deltas" to the same high resolution

climatology. The resulting values for the high-resolution grid cell containing a particular community became the downscaled values for that community. Downscaled monthly values were then averaged across decadal time slices (2010–2019, 2020–2029, …, 2090–2099). The downscaled values were computed separately using the historical baselines from PRISM and CRU TS 3.2 for 1961–1990, and separately for the RCP 4.5, RCP 6.0 and RCP 8.5 forcing scenarios in order to provide an indication of the sensitivities of the downscaled products. The downscaled values for any particular community are the values for the 2 km grid cell containing that community. While there is no consideration of the measurement site's location within the 2 km grid cell (e.g., valley vs. mountain), the use of fine (2 km) grid cells reduces the impact of within-call variability for most communities. However, the same cannot be said for the model grid cells, typically 100–200 km in size, for which the average elevation or land-sea fraction may be a poor representation of the community's location. For this reason, the bias correction inherent in the downscaling is an important attribute of the downscaling procedure.

Fig. 9 is an example of the downscaled temperatures for Kotzebue, a community on the northwest coast of Alaska. In this example, the downscaling is based on the PRISM climatology and the RCP 6.0 (mid-range) emissions scenario. The results are shown for each calendar month (x-axis). The gray bar represents the historical (1961–1990) climatology based on observational data, while the colored bars are means for individual decades of the 21st century based on the addition of the models' deltas for those decades.

Consistent with the forcing, warming is apparent in all calendar months. However, the warming is greater in the cold season (November–March) than in the warm season. Moreover, the inter-model spread is generally larger than the overall change in the 5-model composite mean, pointing to the range in projected changes associated with the combination of internal variability and across-model differences in formulation, resolution, components that are coupled, and other model characteristics. The across-model spread decreases as the averaging is performed over time slices longer than a single decade, e.g., over 30-year period, pointing to the influence of internal variability on decadal averages.

As a second example, Fig. 10 shows the downscaled precipitation for McGrath, a small community in interior Alaska. In this case, the precipitation is for the RCP 8.5 (high-emission) scenario and is based on the CRU 3.2 historical climatology. (We display the results based on CRU 3.2 and not PRISM in order to avoid redundant graphics while illustrating the choices available to users; there is no evidence that either of the precipitation climatologies is better for a particular region). The monthly clusters of bars show that precipitation is projected to increase in all calendar months, with the largest increases in the warm season, corroborating similar results in Walsh et al. (2008) and Stewart et

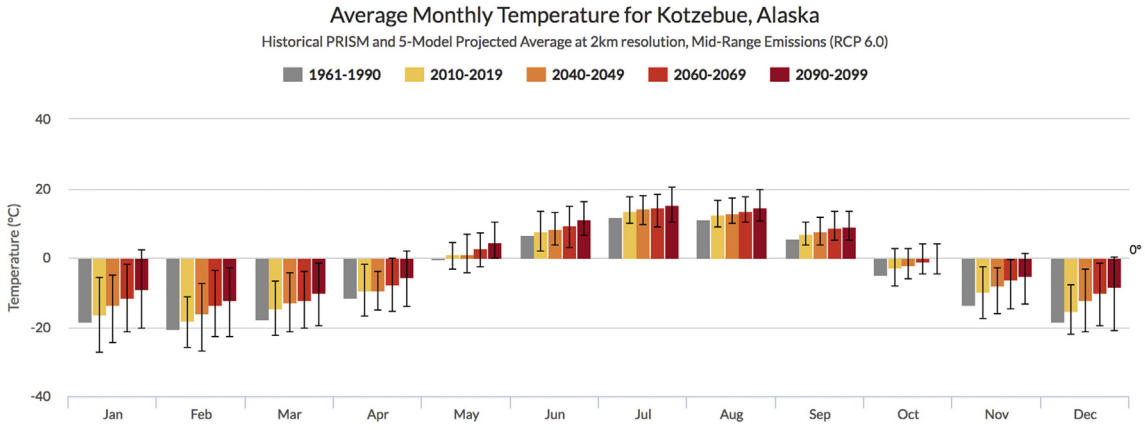

Fig. 9. Decadal means of downscaled temperatures (°C) for Kotzebue, Alaska as a function of calendar month (x-axis). Colored bars are means for individual decades. Thin vertical lines denote range of the temperatures obtained from the five models. Results are for RCP 6.0 (mid-range) emissions scenario.

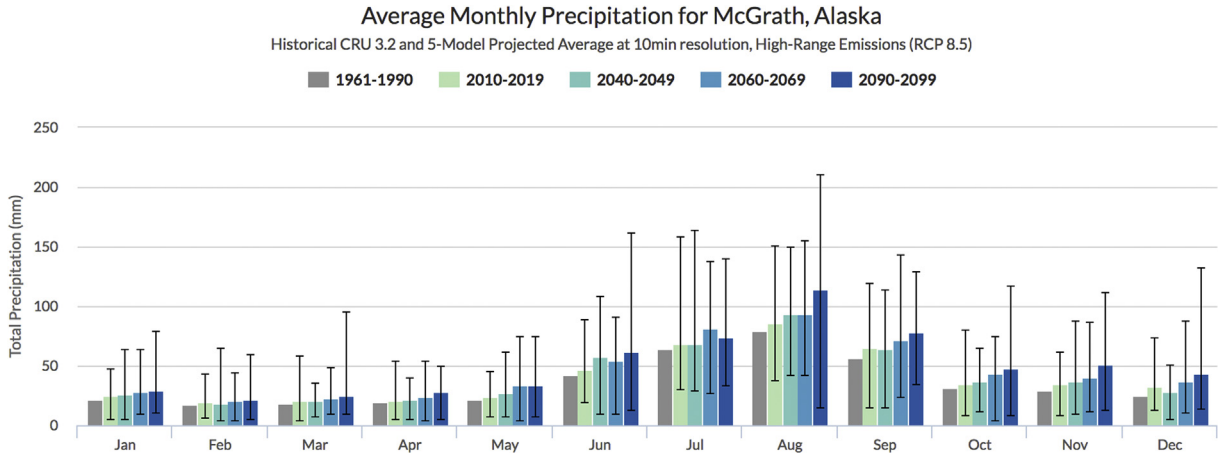

Fig. 10. As in Fig. 8, but for calendar-month precipitation (mm) at McGrath, Alaska under the RCP 8.5 scenario and referenced to the CRU 3.2 historical climatology.

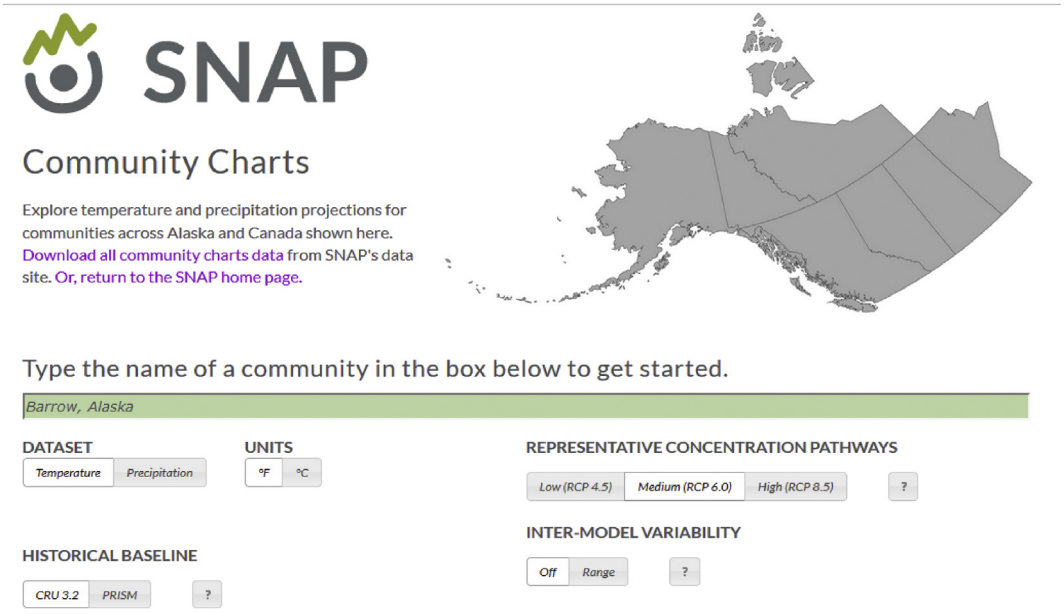

Fig. 11. Screen capture of the home page of the SNAP website for visualization of the community charts. The home page for the visualization tool is accessible at https://www.snap.uaf.edu/sites/all/modules/snap_community_charts/charts.php.

al. (2013). The across-model spread is even larger than in Fig. 9, indicative of a general tendency for greater spread in precipitation projections than in temperature projections. In this case, the across-model uncertainties are far larger than the changes in the composite (five-model mean) values. The spread generally increases with time, indicating greater uncertainty in the late-century projections than in the mid-century projections. The 5-model mean projections in Fig. 10 even show occasional decreases from one time slice to the next (e.g., the blue bars for June and July), pointing to a role of internal variability in the decadal means. Because internal variability is a source of uncertainty in addition to the uncertainty associated with across-model differences in formulation, the future changes have the character of a "bumpy ride" rather than a steady progression, especially in the case of precipitation.

The values downscaled for each model were based on a single simulation (ensemble member). The across-model spread would decrease if the estimates were based on averages of multiple ensemble members from each model rather than a single ensemble member, since internal variations would be reduced by averaging over multiple simulations.

4.2. User interface

A key aim of the SNAP downscaling was the facilitation of use by stakeholders. For this reason, a user interface was developed not only to provide public access to the products, but to encourage users to visualize and experiment with the downscaled projections for their particular locations of interest. User-driven exploration of the sensitivities of the output was one of the priorities in the design and implementation of the user interface. This interface allows users to select different options for various calculation and display parameters: the variable (temperature or precipitation), the units (°F or °C, inches or millimeters), the reference database for the historical period (PRISM, CRU 3.2), the forcing scenario (RCP 4.5, RCP 6.0, RCP 8.5), and the inclusion (or not) of the across-model ranges in the display of the projections. Fig. 11 is a screen capture of the user interface, which also provides the option to download a user-created chart for a particular community.

As an example of the sensitivities that a user can explore, Fig. 12 shows a comparison of the projected changes of temperature (°F) at Point Hope, Alaska under the RCP 4.5 (upper) and RCP 8.5 (lower) scenarios. The warming shows clear signs of leveling off in the RCP 4.5 scenario, but continues to increase in the RCP 8.5 scenario. The difference in warming between the two scenarios is approximately 10 °F in the winter months (December–February). Perhaps more importantly, the monthly mean temperatures in the transition months (May, October) rise above freezing by 2100 under RCP 8.5 while remaining at or below freezing under RCP 4.5. Implications for duration of the ice-free season, which affects over-land travel as well as offshore activities (e.g., whaling, subsistence hunting) are significant in coastal areas where daily activities are closely tied to the state of the land and ocean surfaces. As in the preceding examples, Fig. 12 shows the results based on only one climatology (PRISM) in order to avoid redundant graphics; there is no evidence that either of the temperature climatologies is better for a particular region.

As a second example of exploration of sensitivities, Fig. 13 shows the downscaled precipitation values for Juneau, a relatively wet location in southeast Alaska, based on the PRISM (upper) and CRU (lower) reference climatologies. In both cases, the forcing is the RCP 8.5 scenario. While the projected changes are the same in the two cases, the actual amounts are larger with PRISM, which has a wetter reference climatology for Juneau. During the late summer and autumn months, which are Juneau's wettest, the differences in the two climatologies are as large as 5 cm (2 inches). Such differences are comparable to the projected changes from the late 1900s to the 2090s, pointing to the importance of a robust base climatology in the use of downscaled climate projections.

The SNAP visualization tool for the community charts has been accessed by users within and outside of Alaska. It has provided reference material for the Alaska section of the Third U.S. National Climate Assessment (Stewart et al., 2013), and it has provided input to climate adaptation planning efforts for Alaskan communities (Nome Eskimo Community, 2017). Feedback from users has led to additions to the original capabilities, including the capability to download user-generated charts.

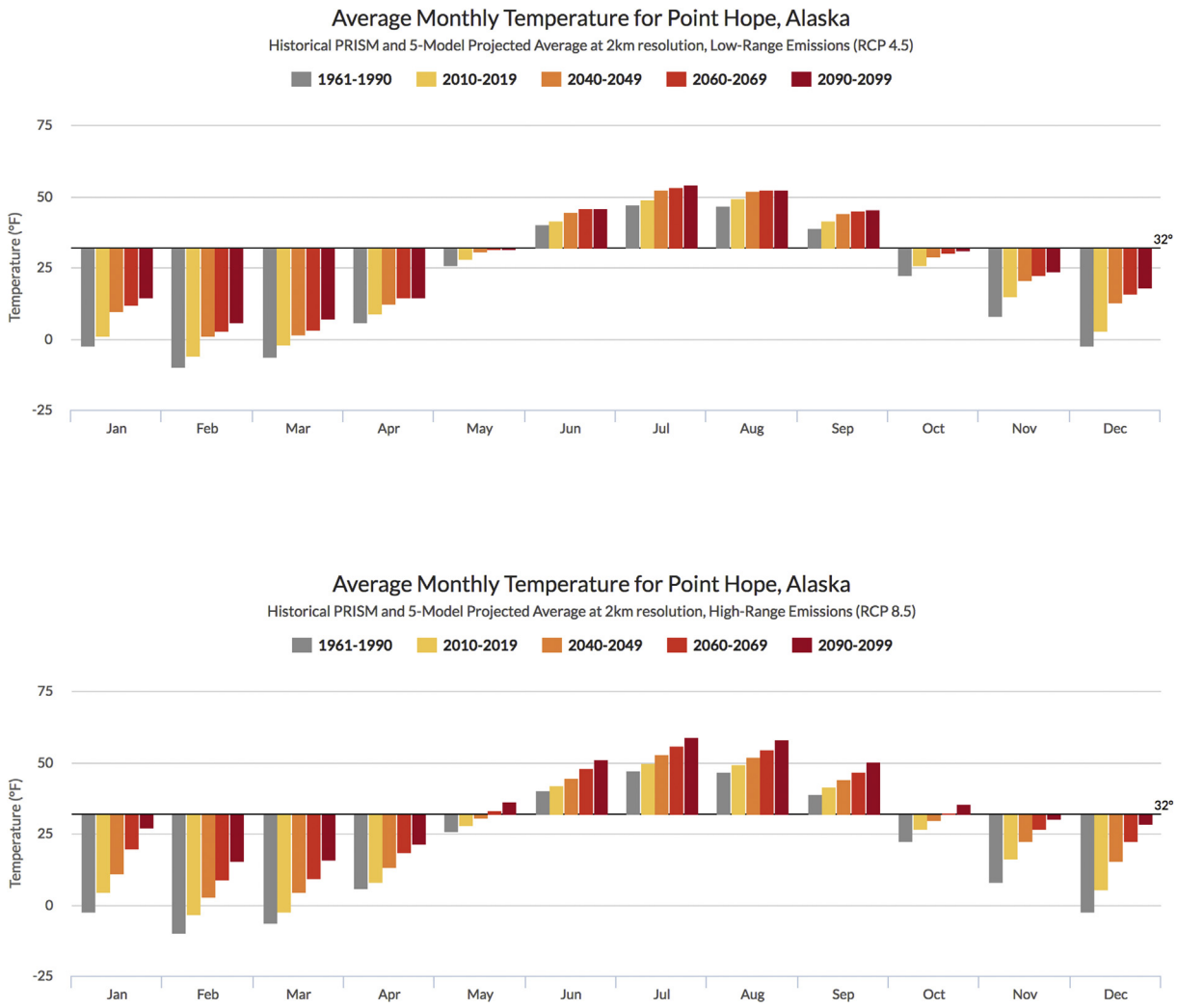

Fig. 12. Decadal means of downscaled temperatures (°F) for Point Hope, Alaska as a function of calendar month (x-axis) under the RCP 4.5 scenario (upper) and RCP 8.5 scenario (lower).

5. Software and data availability

All downscaled climate data and software tools discussed in this paper were produced by the Scenarios Network for Alaska and Arctic Planning (SNAP) and are available under a Creative Commons 4.0 International License (https://creativecommons.org/licenses/by/4.0/), where only attribution to SNAP is needed with no additional restrictions allowed.

The AR5 GCM Evaluation Tool (https://uasnap.shinyapps.io/ar5eval/) was developed in January of 2016 using the R programming language (https://cran.r-project.org/) Shiny web application framework (https://shiny.rstudio.com/). The developer is Matthew Leonawicz, mfleonawicz@alaska.edu The only hardware requirement is a computer with an internet connection. There are no special software requirements, and there is no charge for public users. The raw data utilized by this app were obtained from the Coupled Model Intercomparison Project, version 5 (CMIP5, http://cmip-pcmdi.llnl.gov/cmip5/data_portal.html) and the University of East Anglia's Climate Research Unit (CRU, http://www.cru.uea.ac.uk/data).

The SNAP Community Charts, including the downscaling software and the visualization tool, (https://www.snap.uaf.edu/sites/all/modules/snap_community_charts/charts.php) were developed in 2009 and updated in 2015 to display the latest CMIP5 climate data. The Community Charts utilize jQuery, jQueryUI, HighCharts, MySQL, and PHP programming languages. SNAP downscaled monthly climate data are available in geotiff format for download from https://www.snap.uaf.edu/tools/data-downloads.

6. Conclusion

The project described here represents an end-to-end activity connecting the global climate modeling enterprise with planners, decision-makers and other users in Alaska. The effort has included retrieval of the requisite observational datasets and model output, a model evaluation and selection procedure targeted at the Alaska region, the actual downscaling by the delta method with its inherent bias-adjustment, and the provision of the data to a range of users through a visualization tool that empowers users to explore the downscaled output and its sensitivities. The website's documentation of the visualization tool provides users with a summary of the main components of the downscaling, but there have also been frequent requests for a reference that can be cited. The present paper responds to those requests.

Because the downscaled products have been accessible to users for several years, “lessons learned” have begun to accumulate. One

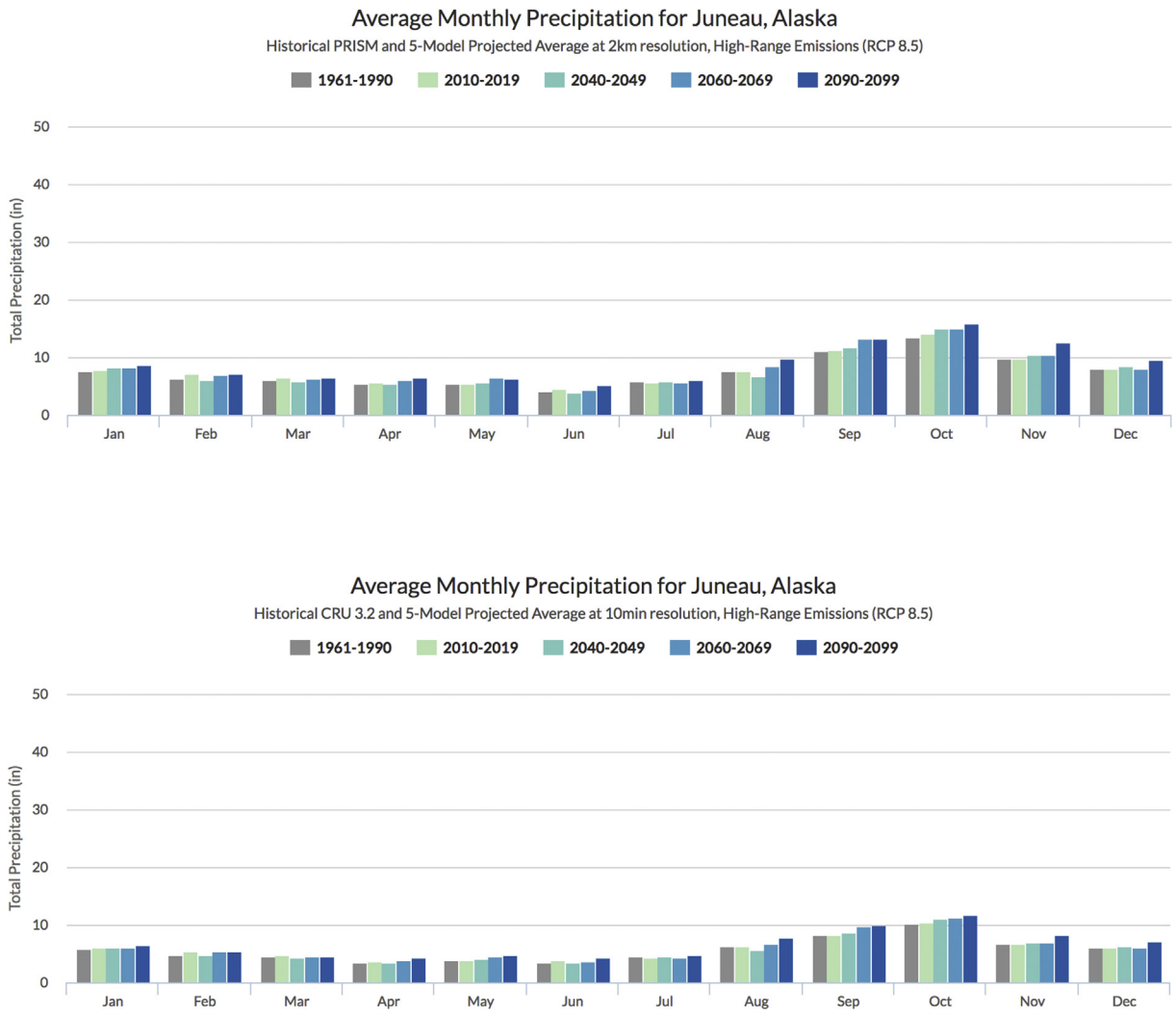

Fig. 13. Decadal means of monthly precipitation (inches) for Juneau, Alaska as a function of calendar month (x-axis) based on the historical climatologies of PRISM (upper) and CRU 3.2 (lower).

lesson is that users desire the actual plots or digital data for presentation purposes or for supporting statements about potential climate change in their area. Second, there is need for caution with regard to the internal variability that can affect decadal means but can also be obscured by compositing of projections from several models (five, in this case). In recognition of this need for caution, the decision was made to include the across-model range indicators as a user option. However, our experience has been that many users do not realize that these range indicators include uncertainties due to both internal variability and differences in model formulations. It has been necessary to make this point in a more complete framework of uncertainties in future projections (e.g., Hodson et al., 2013).

Finally, the downscaled products and visualization tool have proven to be useful for messaging about the role of human activities, especially alternative futures as they may result from different emissions scenarios (RCP 4.5 vs. RCP 8.5, the contrasting options of the user interface). We did not include the RCP 2.6 scenario because the emissions reductions (with negative emissions by 2100) are so extreme that this scenario is rapidly becoming impossible to achieve. Using the other three primary RCP scenarios, the community charts take this scenario dependence down to the local scale that is of greatest interest and concern to a user. The messages conveyed by the charts are consistent with broader depictions of Arctic change and Overland et al. (2014) "adaptation" and "mitigation" timeframes: (1) climate change (especially warming) is already built into the system over the next few decades, even under emissions reduction scenarios, so adaptation will be necessary; and (2) the choice of the emissions scenarios substantially alters the trajectory of local climate in the second half of the century, so mitigation will ultimately make a difference in a community's future climate.

Acknowledgments

This work was supported by the Alaska Climate Science Center through a Cooperative Agreement G10AC00588 from the USGS and by NOAA's Climate Program Office through Grants NA15OAR4310169 and NA16OAR4310162.

We acknowledge the World Climate Research Programme's Working Group on Coupled Modeling, which is responsible for CMIP, and we thank all the climate modeling groups for producing and making available their model output. For CMIP the US Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and leads development of software infrastructure in partnership with the Global

Organization for Earth System Science Portals. Finally, we thank an anonymous reviewer for insightful and constructive comments on the original submission. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

Bhatt, U.S., Zhang, J., Tangborn, W.V., Lingle, C.S., Phillips, L., 2007. Examining glacier mass balances with a hierarchical modeling approach. Comput. Sci. Eng. 9 (2), 60–67. https://doi.org/10.1109/MCSE.2007.29.

Bieniek, P.A., Bhatt, U.S., Walsh, J.E., Scott Rupp, R., Zhang, J., Kreiger, J.R., Lader, R., 2016. Dynamical downscaling of ERA-Interim temperature and precipitation for Alaska. J. Appl. Meteorol. Climatol. 55, 635–654.

Bureau of Reclamation, 2013. Downscaled CMIP3 and CMIP5 Climate Projections Release of Downscaled CMIP5 Climate Projections, Comparison with Preceding Information, and Summary of User Needs. U.S. Department of the Interior, Bureau of Reclamation, p. 104 available at: http://gdodcp.ucllnl.org/downscaled_cmip_projections/techmemo/downscaled_climate.pdf.

Christensen, J.H., Hewitson, B., Busuioc, A., Chen, A., Gao, X., Held, I., Jones, R., Kolli, R.K., Kwon, W.-T., Laprise, R., Magaña Rueda, V., Mearns, L., Menéndez, C.G., Räisänen, J., Rinke, A., Sarr, A., Whetton, P., 2007. Regional climate projections. In: climate change 2007: the physical science basis. In: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M., Miller, H.L. (Eds.), Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W.J., Johns, T., Krinner, G., Shongwe, M., Tebaldi, C., Weaver, A.J., Wehner, M., 2013. Long-term climate change: projections, commitments and irreversibility. In: climate change 2013: the physical science basis. In: Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M. (Eds.), Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Daly, C., Halbleib, M., Smith, J.I., Gibson, W.P., Doggett, M.K., Taylor, G.H., Curtis, J., Pasteris, P.P., 2008. Physiographically-sensitive mapping of temperature and precipitation across the conterminous United States. Int. J. Climatol. 28, 2031–2064.

Fyfe, J.C., von Salzen, K., Gillett, N.P., Arora, V.K., Flato, G., McConnell, J.R., 2013. One hundred years of Arctic surface temperature variation due to anthropogenic influence. Nat. Sci. Rep. 3, 2645. https://doi.org/10.1038/srep02645.

Harris, I., Jones, P.D., Osborn, T.J., Lister, D.H., 2014. Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. Int. J. Climatol. 34, 623–642. https://doi.org/10.1002/joc.3711.

Hayhoe, K., 2010. A Standardized Framework for Evaluating the Skill of Regional Climate Downscaling Techniques. Ph. D. Thesis. Dept. of Atmospheric Sciences, University of Illinois at Urbana-Champaign, p. 153.

Hodson, D.L.R., Keeley, S.P.E., West, A., Ridley, J., Hawkins, E., Hewitt, H.T., 2013. Identifying uncertainties in Arctic climate projections. Clim. Dynam. 40, 2849–2865.

IPCC, 2013. Annex I: atlas of global and regional climate projections. In: van Oldenborgh, G.J., Collins, M., Arblaster, J., Christensen, J.H., Marotzke, J., Power, S.B., Rummukainen, M., Zhou, T. (Eds.), Climate Change 2013: the Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Knutti, R., Furrer, R., Tebaldi, C., Cermak, J., Meehl, G.A., 2010. Challenges in combining projections from multiple climate models. J. Clim. 23, 2739–2758.

Koenigk, T., Berg, P., Doescher, R., 2015. Arctic climate change in an ensemble of regional CORDEX simulations. Polar Res. 34, 24603. https://doi.org/10.3402/polar.v34.24603.

Lader, R., Walsh, J.E., Bhatt, U.S., Bieniek, P.A., 2017. Projections of 21st-century climate extremes for Alaska via dynamical downscaling. J. Appl. Meteorol. Climatol. 56, 2393–2409. https://doi.org/10.1175/JAMC-D-16-0415.1.

Maurer, E.P., Brekke, L., Pruitt, T., Duffy, P.B., 2007. Fine-resolution climate change projections enhance regional climate change impact studies. Eos Trans. Am. Geophys. Union 88, 504. https://doi.org/10.1029/2007EO470006.

Mearns, L., Gutowski, W., Jones, R., Leung, L., McGinnis, S., Nunes, A., Qian, Y., 2009. A regional climate change assessment program for North America. Eos Trans. Am. Geophys. Union 90, 311–312. https://doi.org/10.1029/2009EO360002.

Nome Eskimo Community, 2017. Nome Tribal Climate Adaptation Plan. Nome Eskimo Community and Alaska Center for Climate Assessment and Policy (submitted for publication).

Overland, J.E., Wang, M., Ballinger, T.J., 2018. Recent increased warming of the Alaska marine Arctic due to mid-latitude linkages. Adv. Atmos. Sci. 35, 75–84. https://doi.org/10.1007/s00376-017-7026-1.

Overland, J.E., Wang, M., Walsh, J.E., Stroeve, J.C., 2014. Future Arctic climate changes: adaptation and mitigation timescales. Earth's Future 2, 68074. https://doi.org/10.1002/2013EF000162.

Rogers, T.S., Walsh, J.E., Leonawicz, M., Lindgren, M., 2015. Arctic sea ice: use of observational data and model hindcasts to refine future projections of ice extent. Polar Geogr. 38, 22–41.

Sanderson, B.M., Knutti, R., Caldwell, P., 2015. Addressing Interdependency in a Multimodel Ensemble by Interpolation of Model Properties. J. Clim. 28, 5150–5170. https://doi.org/10.1175/JCLI-D-14-00361.1.

Stewart, B.C., Kunkel, K.E., Stevens, L.E., Sun, L., Walsh, J.E., 2013. Regional Climate Trends and Scenarios for the U.S. National Climate Assessment: Part 7. Climate of Alaska. NOAA Technical Report NESDIS 142-7, p. 60. http://www.nesdis.noaa.gov/technical_reports/NOAA_NESDIS_Tech_Report_142-7-Climate_of_Alaska.pdf.

Thoman, R., Brettschneider, B., 2016. Hot Alaska: as the climate warms, Alaska experiences record high temperatures. Weatherwise 69 (6), 12–20.

Thrasher, B., Maurer, E.P., McKellar, C., Duffy, P.B., 2012. Technical Note: Bias correcting climate model simulated daily temperature extremes with quantile mapping. Hydrol. Earth Syst. Sci. 16 (9), 3309–3314. https://doi.org/10.5194/hess-21-589-2017.

USGCRP, 2014. In: Melillo, J.M., Richmond, T.C., Yohe, G.W. (Eds.), Highlights of Climate Change Impacts in the United States: the Third National Climate Assessment. U.S. Global Change Research Program, p. 148.

Walsh, J.E., Chapman, W.L., Romanovsky, V.E., Christensen, J.H., Stendel, M., 2008. Global climate model performance over Alaska and Greenland. J. Clim. 21, 6156–6174.

Wang, M., Overland, J.E., 2009. A sea ice free summer Arctic within 30 years? Geophys. Res. Lett. 36 https://doi.org/10.1029/2009GL037820. L07502.

Zhang, J., Bhatt, U.S., Tangborn, W.V., Lingle, C.S., 2007a. Response of glaciers in northwestern North America to future climate change: an atmosphere/glacier hierarchical modeling approach. Ann. Glaciol. 46, 283–290. https://doi.org/10.3189/172756407782871378.

Zhang, J., Bhatt, U.S., Tangborn, W.V., Lingle, C.S., 2007b. Climate downscaling for estimating glacier mass balances in northwestern North America: validation with a USGS benchmark glacier. Geophys. Res. Lett. 34 https://doi.org/10.1029/2007GL031139. L21505.