Incorporating Climate Change Projections into Risk Measures of Index-Based Insurance

Zhuoli Jin and Robert J. Erhardt

Department of Mathematics and Statistics, Wake Forest University

jinz16@wfu.edu, erhardrj@wfu.edu

November 5, 2017

Abstract

We present a complete working example of using regional climate model projections to estimate the changing risks of temperature index-based insurance products defined for a series of locations in California. This region is a major agricultural producer. The climate model projections are an ensemble of three regional climate models obtained from the North American Regional Climate Change Assessment Program. Hindcasts for the period of 1971-2000 are compared to historical observed temperature data for bias and variance corrections. Adjusted future model projections are used to estimate distributions of cooling degree days for 2041-2070, which are then used to estimate risk measures for index-based insurance products defined for cooling degree day indices.

Keywords: climate risk, NARCCAP, regional climate model, weather derivative

DRAFT

Contents

DRAFT

1 Introduction

Index-based insurance specifies payments to the policyholder based on the settlement value of some external index. The contract specifies the index, the mathematical function which describes payments for all possible settlement values of the index, and the time period over which the insurance is in force. Whereas traditional insurance compensates policyholders based on the actual realized loss from a covered peril, index-based insurance compensates policyholders based on an external trigger, regardless of the actual realized loss. The advantages of this approach include reduced administrative costs for the insurer – moral and morale hazards are virtually eliminated as the index is objectively verified and beyond anyone's control or manipulation, and there is no need to confirm the actual policyholder losses. Payments can be issued remotely and quickly. Index-based insurance has been widely touted in developing regions underserved by traditional insurance markets (see, for example the review papers by Carter et al. (2014) and Collier et al. (2009), examples from Barnett and Mahul (2007) and Alderman and Haque (2007), or the Global Index Insurance Facility at `https://www.indexinsuranceforum.org/`).

As an example, consider the Livelihood Protection Plan from the Munich Climate Insurance Initiative ((MCII), 2013). This index-based insurance specified payments to policyholders in the Caribbean triggered when the maximum daily wind speed exceeded a high threshold, or when the total rainfall exceeded some high threshold. The motivation was to issue payments for severe weather events, likely to cause physical damage and economic distruption. Rainfall was monitored by the Danish Hydrological Institute, and wind speeds were monitored by the Caribbean Catastrophe Risk Insurance. Policyholders were sent SMS text messages to their mobile phones warning them of upcoming storms which may trigger the insurance, and after a triggering event they were sent notice and payment without any loss adjustment or claims verifcation.

Other index-based insurance products have been defined with triggers based on precipitation extremes (Fischer et al., 2012), the *normalized vegetation depth index* (NDVI, a satellite-based measurement of vegetation which correlates with crop risk (Turvey and Mclaurin, 2012)), temperature (Turvey, 2005), a temperature-humidity index designed to correlate with dairy production risk (Deng et al., 2007), and more complex indices meant to capture livestock risk (Chantarat et al., 2013).

To price index-based insurance accurately, one needs an accurate estimate of the distribution function of payments, but these are mathematically determined for all given index-values, and so the distribution of the index variable fully specifies the distribution of the payments. The focus of this paper is on weather index-based insurance products, which includes all of the examples mentioned above, and so a starting point for actuarial work is the estimation of the index distribution. This paper will explore risk measures for

DRAFT

a temperature index-based insurance product in Fresno, CA. Historical data on daily temperature can be obtained from publicly available sources (see the National Centers for Environmental Information (NCEI) for data in the US, `https://www.ncei.noaa.gov/`). When using a few decades worth of historical temperature data to estimate the distribution of temperature in the near future, one must ask (Q1) if the distribution of temperatures has been *stationary* over this time period – that is, can we regard all historical data as being draws from the *same* distribution of daily temperatures, or is it possible that the distribution has been changing over time? Furthermore, if the period of interest for the index-based insurance product is far enough in the future, one must also ask (Q2) if the distribution in the future will match the distribution being estimated in the present, or if instead further adjustments will be needed.

Historical data show that for many weather variables, distributions have not been stationary over the past few decades. Consider Figure 1 from the National Centers for Environmental Information. This image shows global mean temperature anomolies temperature from 1880 - present (a temperature anomaly is a year's mean temperature minus the average taken over 1900 - 2000). As is strikingly evident, average temperatures have been trending up over the past few decades, and therefore the mean of the distribution of temperature has been trending up. For an excellent visualization of shifting temperature distributions, see also Hansen et al. (2010) and Hansen et al. (2012). The Actuaries' Climate Index (ACI, `http://actuariesclimateindex.org/home/`) similarly looks at historical data to identify shifts in climate. Revisiting (Q1) from the previous paragraph, we conclude that the distribution of global mean temperatures has not been stationary, and therefore it makes sense to consider adjustments to past data over several decades.

Investigation of (Q2) is the primary focus of this paper. To ask question (Q2) is to ask what scientists project forward for climate trends, and how those projections could be utilized by actuaries. These projections come from climate models, which are high quality computer simulations of the Earth's climate which can be run forward in time. Full discussion of global and regional climate models is an extensive topic beyond this paper, but for more on the basics of the underlying physical science, climate model construction and primary results, see the Society of Actuaries white paper Erhardt and Von Burg (2018), or the Intergovernmental Panel on Climate Change (IPCC, `http://www.ipcc.ch/`) reports (Pachauri et al., 2014).

In this paper, we use an ensemble of three regional climate models from the North American Climate Change Assessment Program (NARCCAP, `http://www.narccap.ucar.edu/`, Mearns et al. (2007)) to estimate changing risks of temperature index-based insurance in central California. We first describe how to obtain, visualize, and manipulate NARCCAP regional climate model output. Comparing hindcasts of climate model output to historical data, we demonstrate the need for bias and variance correction, and next show how these corrections are applied to future climate model projections to estimate future distributions

DRAFT

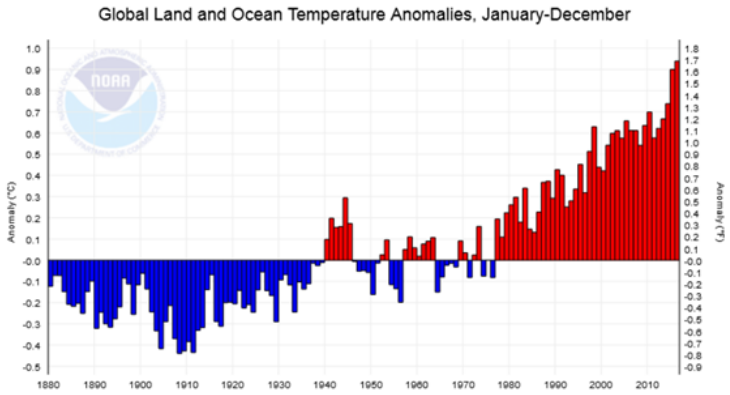

Figure 1: Global average temperature anomalies (land and ocean) from 1880-2016, obtained from the National Centers for Environmental Information (https://www.ncdc.noaa.gov/cag/). Anomalies are defined as a year's mean surface temperature minus the average temperature taken from 1900 - 2000. Years shown in red are warmer than the 20th average, whereas years shown in blue are cooler.

of temperatures. From these, we compute the temperature index for a hypothetical index-based insurance product following Fleege et al. (2004). This product will make payments based on an aggregate temperature index computed from May 1 through July 31, in Fresno CA. We end by demonstrating how risk measures of this increase as a result of a warmer projected climate.

2 Exploratory Data Analysis

2.1 Historical Data

We obtained historical daily temperature for 24 locations in California for the period Jan 1, 1971 through Dec 31, 2000. This time period was selected to match the available hindcast climate model data. The loctions are shown in Figure 2, and were chosen to span the Central Valley region with the requirement that each have less than 10% missing values. There are also several cities located on the west coast and in the inland mountains, for comparison. These data were obtained from the National Centers for Environmental Information (NCEI) Climate Data Online (CDO, `https://www.ncdc.noaa.gov/cdo-web/`). This portal provides free access to the National Climate Data Center (NCDC) archive of global historical weather and climate data.

DRAFT

Summary statistics of the twenty-four cities are provided in Table 1. Figure 3 shows kernel density estimates of the 24 cities by season. In this paper we will refer to winter as "December/January/February" (DJF), each subsequent season as exactly three months (MAM, JJA, SON). The three cities shown in heavy black lines are Fresno, Huntington Lake and Los Angeles. Histograms and qqplots of historical temperature in May, June and July (MJJ) are shown in figure 4.

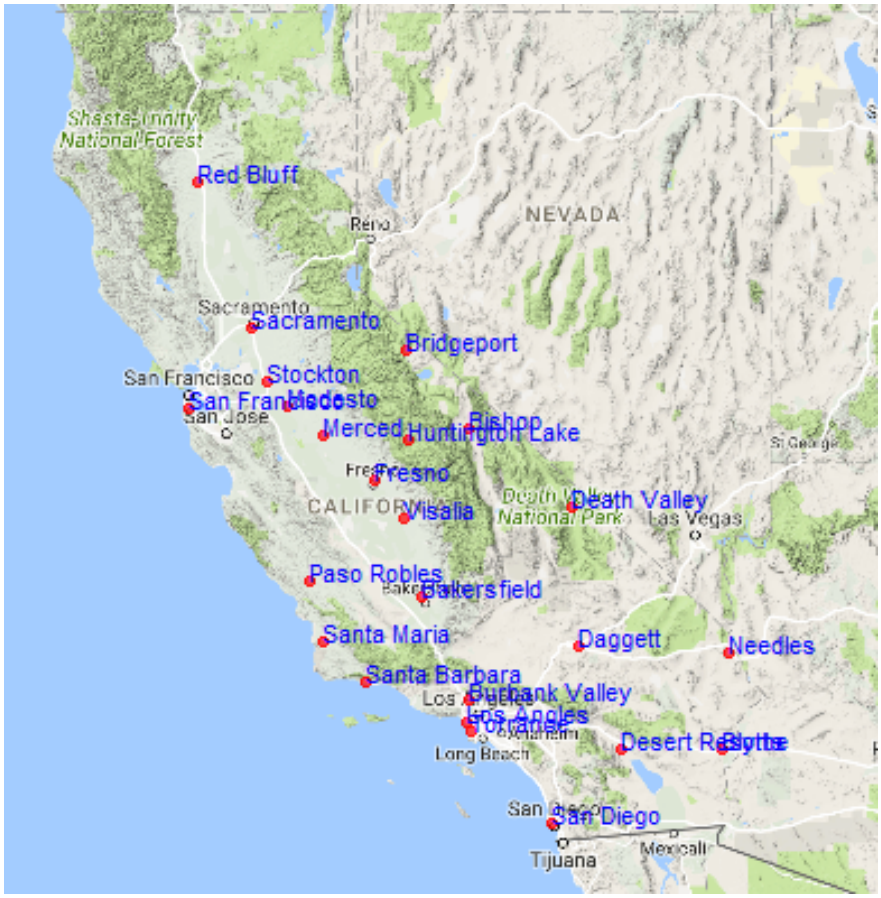

Figure 2: Study region of the 24 cities in California. Each city has less than 10% missing values for the study period January 1, 1971 through December 31, 2000. Cities cover the major regions of the state, including the coast, central valley, desert south, and high mountain regions.

DRAFT

City	% missing days	Mean	Std Dev	Skewness	Kurtosis
Bakersfield	0.00	65.5	14.0	0.05	1.99
Bishop	0.01	56.0	14.9	-0.04	1.96
Blythe	6.00	72.2	15.2	0.01	1.79
Bridgeport	5.70	43.5	14.3	-0.30	2.51
Burbank Valley	1.51	64.7	9.4	0.08	2.34
Daggett	0.28	68.0	15.6	0.02	1.87
Death Valley	0.43	76.4	18.2	-0.02	1.79
Desert Resorts	1.17	72.3	14.2	-0.04	1.88
Fresno	0.00	63.8	13.9	0.03	1.93
Huntington Lake	3.59	45.9	11.6	-0.01	2.27
Los Angeles	0.00	63.3	6.4	0.09	2.82
Merced	4.01	62.0	12.8	0.01	2
Modesto	1.01	62.4	12.2	-0.03	2.11
Needles	1.94	73.7	16.3	0.03	1.79
Paso Robles	0.60	59.4	10.4	0.08	2.25
Red Bluff	6.91	63.2	14.1	0.12	1.97
Sacramento	0.18	61.1	11.9	-0.02	2.09
San Diego	0.00	64.4	6.4	0.13	2.7
San Francisco	0.00	57.6	6.6	-0.02	2.98
Santa Barbara	8.51	59.5	6.5	0.00	2.81
Santa Maria	0.02	57.7	6.5	-0.11	2.91
Stockton	0.01	62.0	12.6	-0.01	2.08
Torrance	0.68	63.1	6.8	0.08	2.77
Visalia	1.16	63.2	13.2	-0.06	1.95

Table 1: Summary statistics for the 24 locations in the study. Missing data ranges from 0.00% to 8.51% in Santa Barbara.

DRAFT

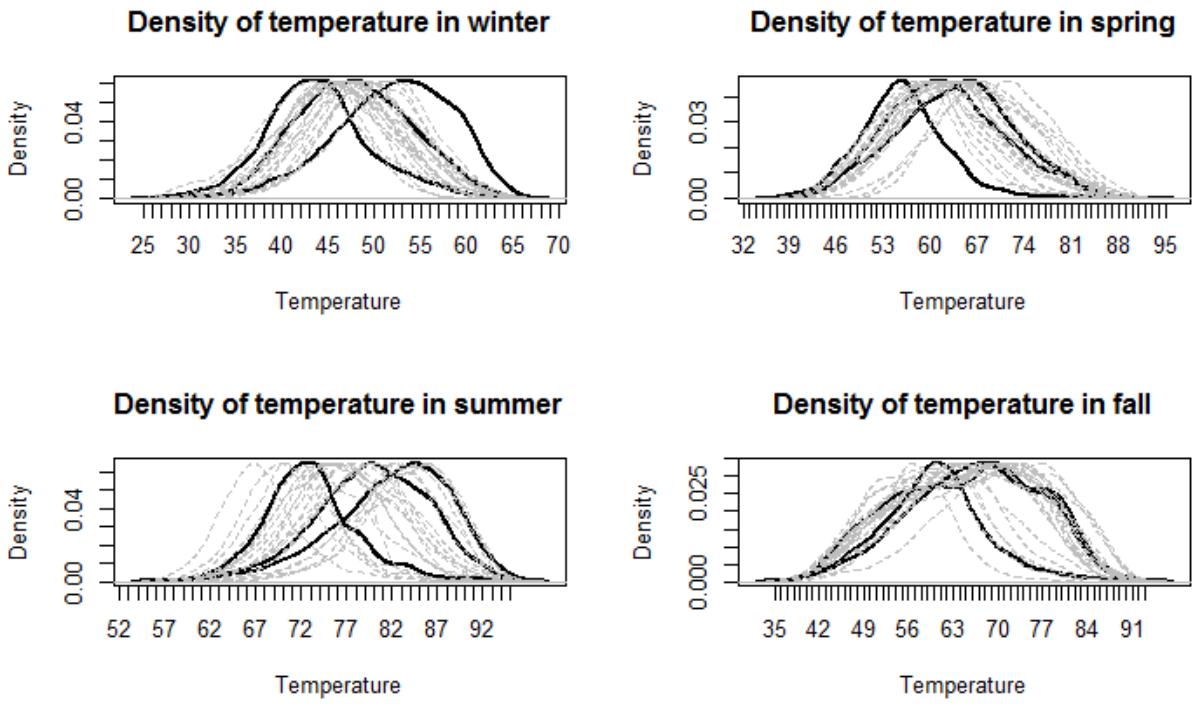

Figure 3: Kernel density plots for each city in each season in a "winter - spring - summer - fall" order. Three black lines in each season plot show the three selected cities, and other gray lines indicates other cities.

DRAFT

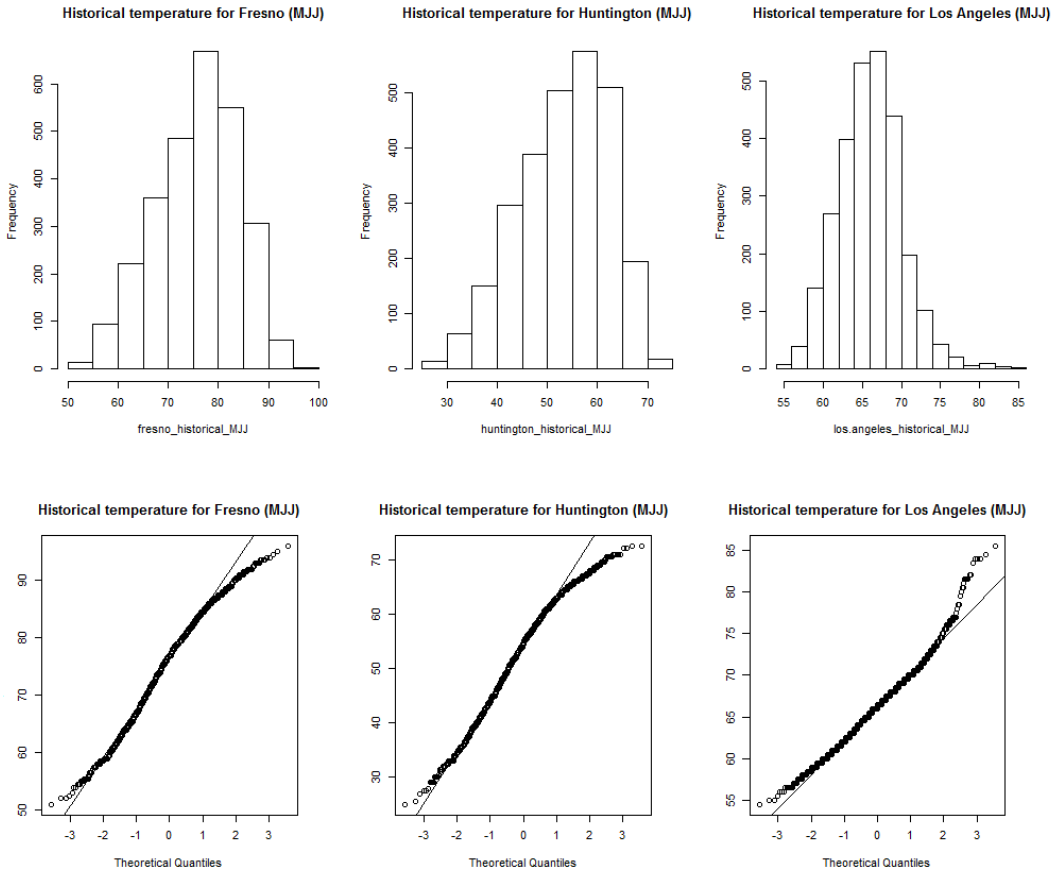

Figure 4: Top: histograms of historical temperatures of Fresno, Huntington Lake and Los Angeles in MJJ, demonstrating a mild skew. Bottom: qqplots of historical temperatures of Fresno, Huntington Lake and Los Angeles in MJJ.

DRAFT

2.2 Regional Climate Model Data

We obtained regional climate model (RCM) output from the North American Regional Climate Change Assessment Program. An introduction and overview of this program can be found in Mearns et al. (2005) and Mearns et al. (2007). NARCCAP is an international program to produce climate projection outputs for North America for the years 2041 to 2070 from multiple regional climate models (RCMs) nested within multiple general circulation models (GCMs). Background about RCMs and GCMs can be found in Music et al. (2015) and Eagleson (2011), respectively, and for an actuarial perspective see Erhardt and Von Burg (2018). What is relevant for this paper is that RCMs are dynamically downscaled climate models designed to allow for the study of more localized, regional consequences of projected climate change. Climate model output from NARCCAP is available for two time periods, a hindcast period from 1971 - 2000 and a future period from 2041-2070. RCM output can be thought of as a cube, with latitude and longitude forming 2 dimentions, and time the third. Temperature is obtained every 3 hours, so for each year we extracted a cube of data with dimensions $16 \times 33 \times 2920$. The latitude and longitude coordinates were chosen to include those forming a rectangle covering California, and can best be seen in Figure 5.

Due to differences in climate models, we chose an ensemble of three different GCM/RCM combinations to explore a range of models. This ensemble contains (1) the Canadian Regional Climate Model nested within the Community Climate System Model (CRCM-CCSM), the (2) Weather Forecasting Research Group model nested within the Third Generation Coupled Global Climate Model (WRFG-CGCM3); and (3) the Hadley Regional Model 3 nested within the Hadley Centre Coupled Model, version 3 (HRM3-HadCM3). All three GCMs were run under the SRES A2 emissions scenario (see `http://www.narccap.ucar.edu/about/emissions.html` or Nakicenovic et al. (2000) for more detail). We will highlight results from the CRCM-CCSM (Caya and Laprise, 1999), (Gent et al., 2011), but the others are qualitatively similar.

2.3 Cooling Degree Days

The index-based insurance product we will consider is triggered by an index measured in *cumulative cooling degree days*, so here we define this index. A *degree day* is a measure of the demand for heating or cooling. Cooling degree days (CDDs), which we discuss in this paper, measure the excess temperature above a certain threshold. Here we use the common 65 degrees Fahrenheit, in which case for a single day d, a CDD is defined as

$$CDD_d = max(0, \bar{T}_d - 65),$$

where $\bar{T}_d$ is the average of daily maximum temperature and minimum temperature measured in degrees

DRAFT

Fahrenheit. We can sum CDDs over some time period, denoted as $\mathcal{D}$, and call this sum the cumulative cooling degree days (cCDDs) for a place/location in time $\mathcal{D}$, defined as

$$cCDD = \sum_{d \in D} CDD_d = \sum_{d \in D} max(0, \bar{T}_d - 65). \tag{1}$$

cCDDs are a measure of overall temperature excess above 65 degrees Fahrenheit, without regard to precisely how or when the temperature exceeded the threshold. They serve as a convenient proxy for the overall warmth of a season, and are widely used by utility companies to measure demand for heating and cooling energy, and by agricultural producers to measure the warmth of a particular season.

Erhardt (2015) quantified the projected positive trends in North American cCDDs in the mid-twenty-first-century using NARCCAP data. Here we will explore changes in the overall distribution of cCDDs comparing the *future* period 2041-2070 to the *current* period 1971-2000 (these names follow NARCCAP convention). Following the example in Fleege et al. (2004), we define $\mathcal{D}$ as the period from May 1 to July 31, for both current and future years. We computed the annual cCDDs for each grid cell by summing the daily CDDs over the period of May 1 through July 31. Figure 5 shows the average of these cCDDs taken over the 30 year current period (left panel) and future period (center panel). Owing to warmer projected temperatures in the future, the future period shows higher average cCDDs, best seen by the third panel which shows the difference between the other two panels. For every grid cell in the study region, the cCDDs are projected to increase in the CRCM-CCSM model. This projected increase is largest in the southeast and central valley of CA, with only mild increases projected for the Pacific coast.

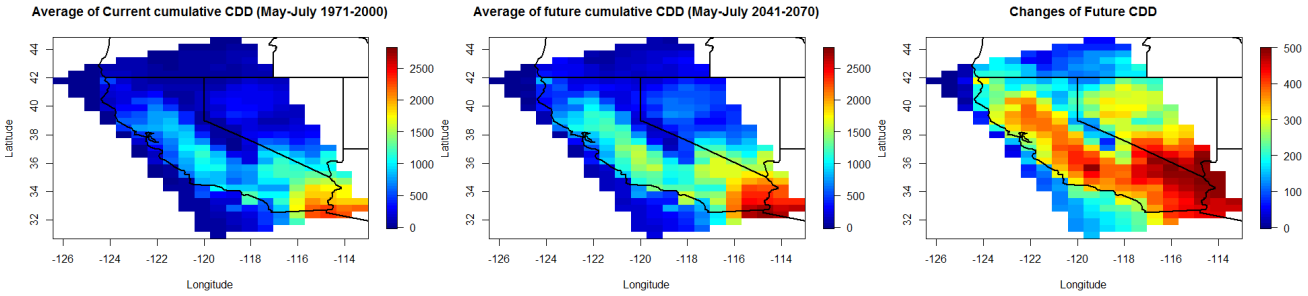

Figure 5: The left and center graphs are cCDDs for current and future time periods for California, respectively. The right graph shows the difference for cCDDs (future cCDDs minus current cCDDs).

Figure 6 shows a pointwise 95% confidence surface for the projected increase in cCDDs, with the lower bound

DRAFT

(left panel) and upper bound (right panel) computed simply as the difference ± 2 standard errors of the difference. Even after incorporating uncertainty in the two estimates, the lower bound is everywhere greater than zero.

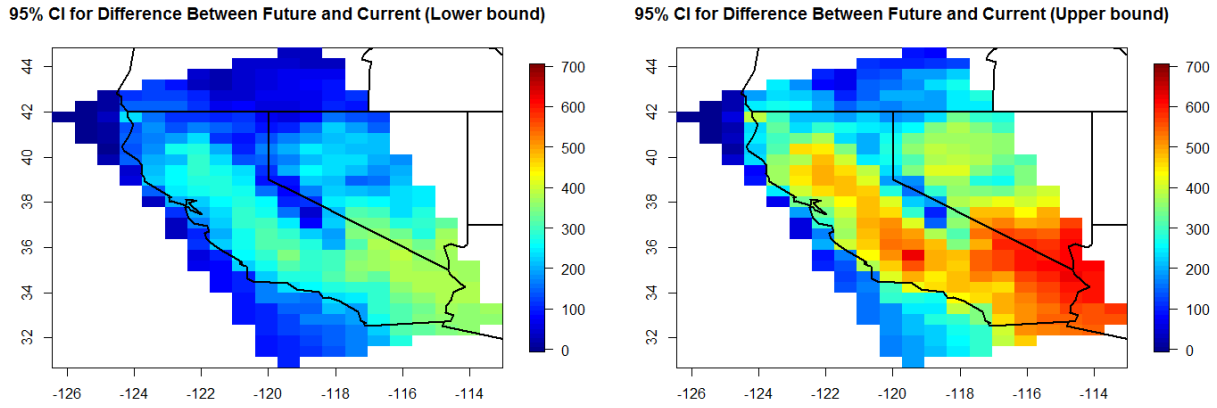

Figure 6: These two panels show the pointwise 95% confidence surface for the projected change in cCDDs from the CCSM-CRCM model over California. The lower bound (left panel) is computed as the mean minus 2 standard errors, and the upper bound (right panel) is computed as the mean plus two standard errors.

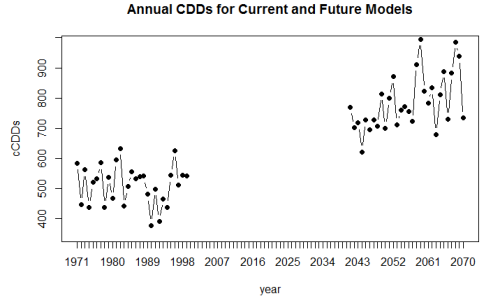

Figure 7: A line chart of current and future average annual cCDDs for all grid cells. The left part is showing average of annual cCDDs from year 1971 to year 2000, and the right part is showing average of annual cCDDs from year 2041 to year 2070.

To further visualize the shift in cCDDS, Figure 7 shows a line chart of the average annual cCDDs for all grid cells, for both current and future time periods. In this chart, it is easy to tell that annual cCDDs are higher

DRAFT

in the future, and therefore any payments in an index-based insurance product based on cCDDs would be similarly impacted.

3 Downscaling

3.1 Introduction to Statistical Downscaling

Consider the three cities of Fresno, Huntingdon Lake, and Los Angeles, and consider only the 3 grid cells which contain each. Figure 8 shows distributions for observed historical daily tmperatures (solid line), 1971-2000 hindcast RCM output for daily temperatures (dashed line), and 2041-2070 projections of daily temperatures (dot-dash line). It is clear that the means and variances of the hindcast (dashed) and observed data (solid) do not match. It is also clear that distributions of temperature for the future period (dot-dash) have noticeably shifted to the right of the hindcast distributions (dashed).

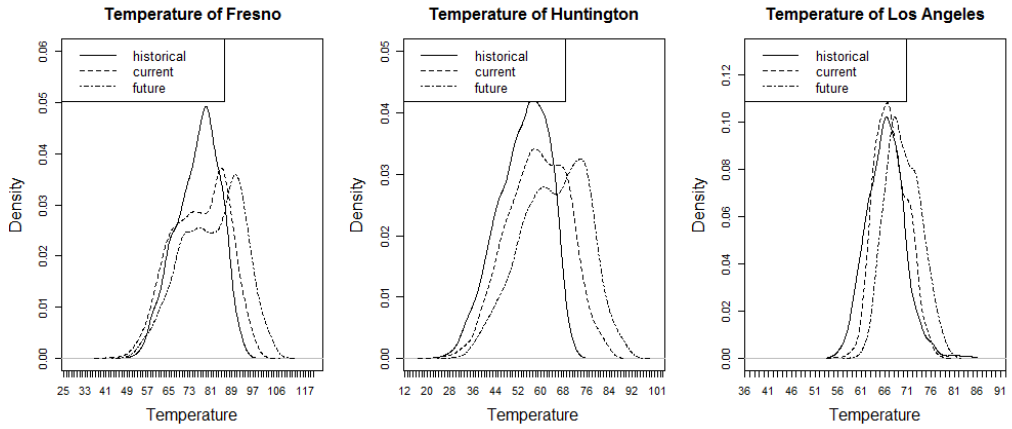

Figure 8: Density plots for historical (solid line), current (dashed line), and future (dot-dashed line) temperature data before downscaling.

Figure 8 demonstrates that the RCM output show climate change trends as the distributions have shifted to the right, but it also shows the need for a bias and variance correction as hindcasted distributions do not match observed climate data. To help remove some of this bias, we match the means and variances of these two distributions. We then apply those same bias and variance corrections to output from the RCM in the future period. The result is what we will refer to as downscaled future temperatures.

DRAFT

For each of the 24 cities and for each regional climate model, we compute a bias correction ψ which measures the mismatch in mean temperatures for hindcast and historical observations, along with a variance correction ϕ which measures the relationship between the variability of daily temperatures for the two periods. That is, we compute the necessary additive and multiplicative constants needed to match the means and variances of the observed and hindcast data. Following Algorithm 1, we apply those same bias and variance corrections to raw NARCCAP future projections to obtain downscaled future projected temperatures $\mathbb{T}_{future,t}$. More detail is provided in the appendices. This correction would eliminate biases in the mean and variance that are preserved throughout time by regional climate models.

Algorithm 1 Simple downscaling

See Appendix for all variable names.

1: Compute $\psi = \bar{T}_{current} - \bar{T}_{historical}$
$\phi = \sigma_{historical}/\sigma_{current}$

2: $\mathbb{T}_{current,t} = (T_{current,t} - \bar{T}_{current}) \times \phi + \bar{T}_{current} - \psi$,
$\mathbb{T}_{future,t} = (T_{future,t} - \bar{T}_{future}) \times \phi + \bar{T}_{future} - \psi$.

3.2 Downscaling Results

After applying the statistical downscaling algorithm 1, Figure 9 (left column) confirms that the bias and variance corrections have matched up downscaled hindcasts and observed distributions. The right column further shows that corrected future distributions are still noticeably shifted to the right as a result of a warmer climate. Figure 10 highlights the shifts from current to future distributions. Tables 2 and 3 summarize some of the statistics from this downscaling procedure. Finally, Figure 11 helps visualize the shifts in mean for each of the three cities, and shows a statistically significant increase after taking year to year variability into consideration. Taken together, these Figures and Tables demonstrate the motivation for and success of the downscaling adjustment.

DRAFT

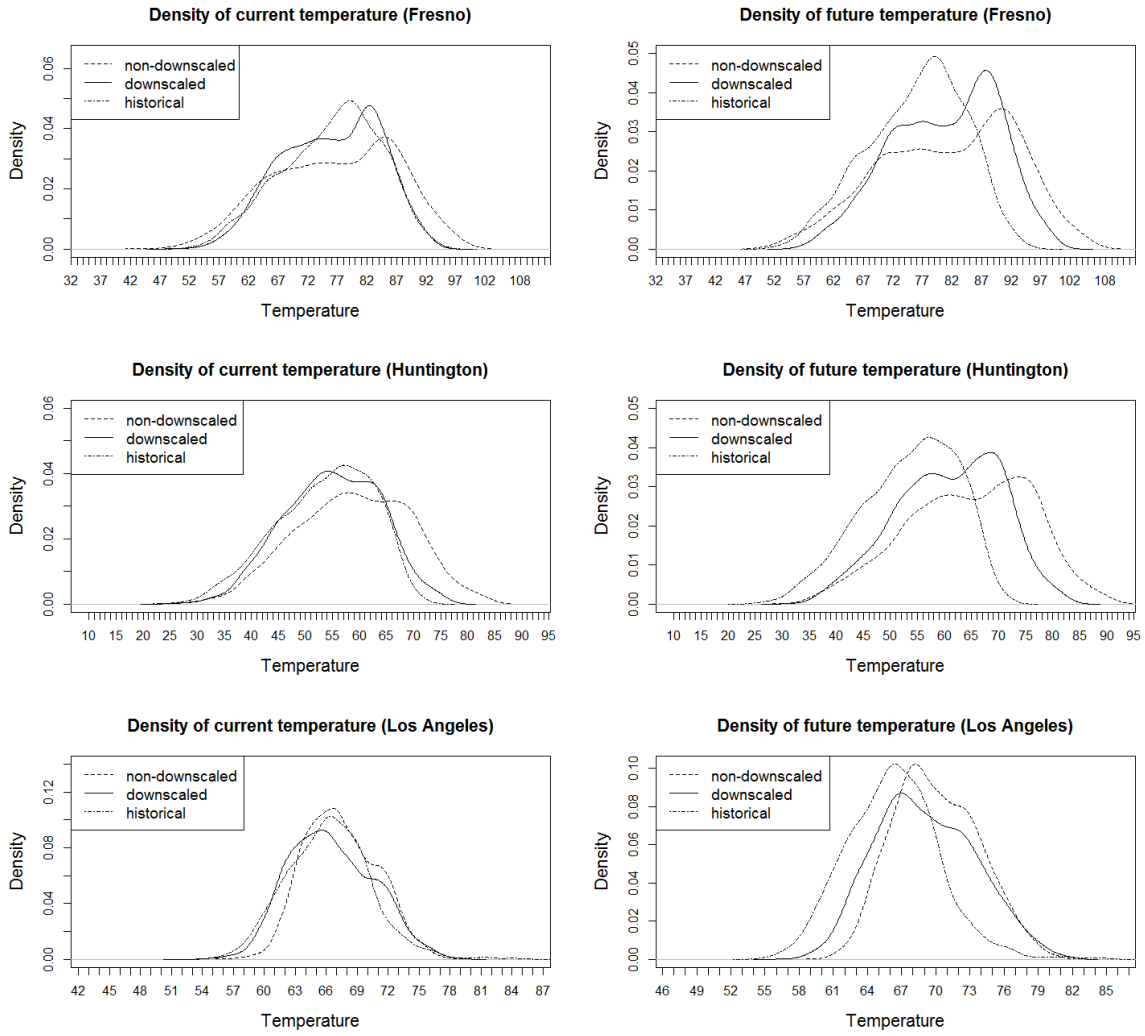

Figure 9: Comparison plots for non-downscaled (dashed line), downscaled (soild line) and historical (dot-dashed line) temperatures in each city. Left panels show the current time period, and right panels show the future time period. The downscaler matches up distributions for the current, and after applying the downscaling on the future, it's clear that the downscaled future distributions are still shifted to the right compared to historical, which is a result of projected warming due to climate change.

DRAFT

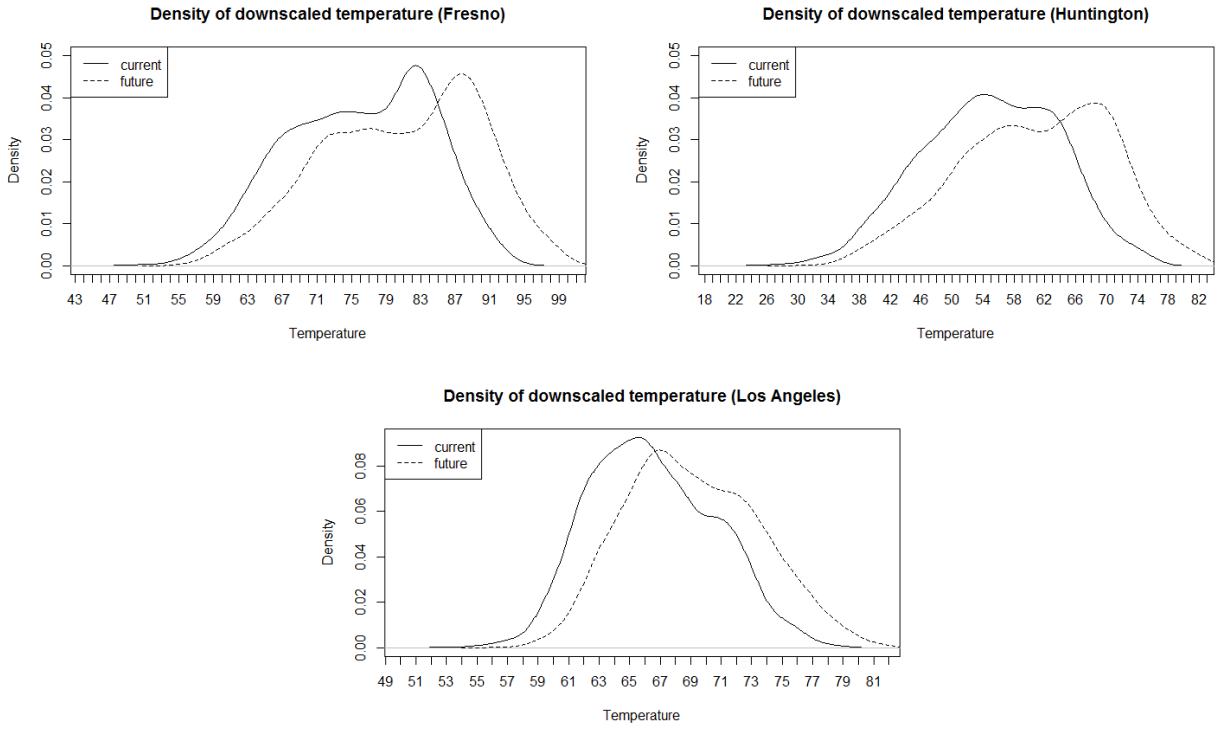

Figure 10: Comparison plots for downscaled current and future temperature for each city. Downscaled temperature in future tends to be higher for each selected city.

	Mean			Std Dev			10^{th} Quantile			90^{th} Quantile		
	F	HL	LA	F	HL	LA	F	HL	LA	F	HL	LA
Current, Non-downscaled	72.4	76.5	66.8	9.9	10.6	3.1	58.9	84.9	62.1	89.7	63.0	71.1
Current, Downscaled	76.1	54.5	66.4	8.2	8.8	4.1	64.9	86.5	42.6	65.5	61.4	72.1
Future, Non-downscaled	77.0	81.2	69.4	10.5	11.5	3.4	62.8	90.1	65.5	95.1	65.3	73.9
Future, Downscaled	80.8	59.3	69.1	8.7	9.6	4.4	68.9	91.6	46.2	70.9	63.5	75.0

Table 2: This table provides mean, standard deviation, 10^{th} and 90^{th} quantile for both non-downscaled and downscaled temperature for Fresno (F), Huntingdon Lake (HL) and Los Angeles (LA). This table shows huge difference in distribution of non-downscaled and downscaled temperature.

DRAFT

Variable	Fresno	Huntington Lake	Los Angeles
ψ	0.73	3.86	0.91
ϕ	0.78	0.84	1.17

Table 3: This table provides bias correction and variance correction for Fresno, Huntington Lake and Los Angeles used in downscaling progress. The variance correction shrinks variance for Fresno and Huntington, but enlarges it for LA.

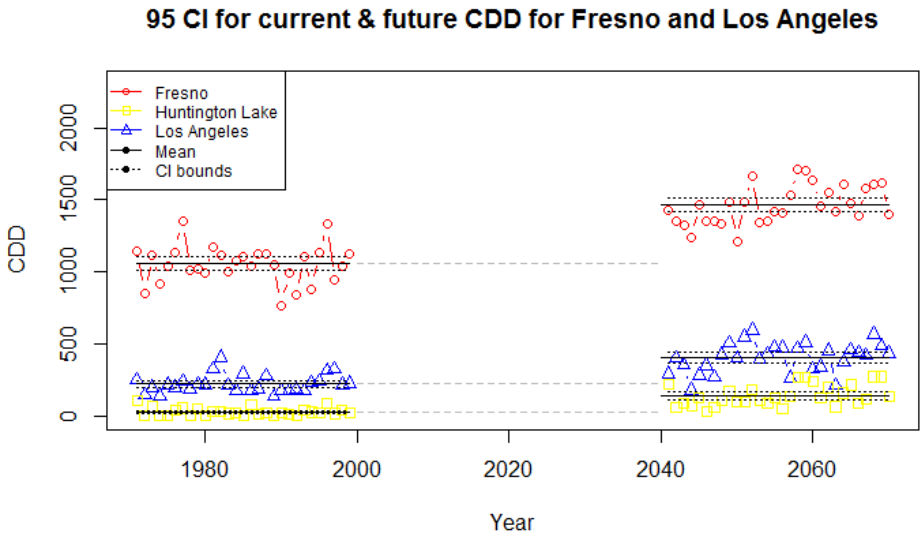

Figure 11: The circles, squares and triangles represent Fresno, Huntington Lake and Los Angeles, respectively. Current downscaled temperature stay on the left, and future downscaled temperature stay on the right. For each city and each time period, its mean downscaled temperature are drew along with a 95% confidence interval of this mean. From this figure, one can tell future annual cCDD in each city tends to be higher than current.

DRAFT

4 Application

Here we describe a hypothetical index-based insurance product that makes payments whenever the cCDD index from May 1 - July 31 exceeds some high threshold. This case is inspired by Fleege et al. (2004), and is meant to mimic an insurance product which provides financial relief to unusually hot weather during the critical growing in central California. Other papers describing the construction of similar products include Zeng (2000), Alaton et al. (2002), and Deng et al. (2007). Specifically, the payment we will consider is

$$Payment = \begin{cases} 0 & cCDD < cCDD_{threshold}, \\ k(cCDD - cCDD_{threshold}) & cCDD \geq cCDD_{threshold} \end{cases} \tag{2}$$

where $cCDD$ is the index settlement value, and $cCDD_{threshold}$ is the value which first triggers payment. For each degree that $cCDD$ exceeds the threshold $cCDD_{threshold}$, a payment of k is issued.

We will consider two simple risk measures for payments. First, the Value at risk (VaR) at probability level p is defined as the p^{th} percentile of the payment. The usual levels of VaR are $p = 0.95$ and $p = 0.99$. Second, we will consider the Conditional tail expectation (CTE) as a risk measure that quantifies the expected value of the loss, given that an event outside a given probability level has occurred. Mathematically, the CTE is defined as

$$CTE_p = E(P|P > P_p), \tag{3}$$

where $1 - p$ is the risk level, P is annual payment, P_p is the p^{th} percentile of P.

We show risk measures for two different settings. In the first, we use only historical temperature data from January 1, 1971 through April 30, 2000 to build a statistical model which is then used to predict forward 92 days. These 92 days of predictions are used to compute the index value, which is in turn used to compute the payment. Using bootstrapping with our fitted model, we obtain 10,000 projected paths of temperature and 10,000 associated payments.

In the second setting, our goal is to produce 10,000 payments for the index-based insurance product for the period of May 1 through July 31, not under the climate of the year 2000 but rather under the climate described by one of the regional climate models from 2041-2070. Comparing this second setting to the first setting allows a direct comparison of changes in payments or risk measures which are due to the changed climate only. To obtain the predictive distribution of payments under this changed climate, we first apply inverse bias and variance corrections map the temperatures to the "current" distributions of each RCM (Algorithm 1, with all parameters shown in Table 5), and then apply the differences in means from "current" 1971-2000

DRAFT

to "future" 2041-2070 for each RCM (demonstrated in Figure 10, with all parameters shown in Table 6). The end result of these two mappings is to produce a distribution of 10,000 payments under the hypothetical future climate described by each of the three RCMs.

Here we describe the construction of the model used in the first setting to form a predictive distribution for daily temperatures for the 92 day period from May 1, 2000 through July 31, 2000. We obtained historical daily temperatures in Fresno, CA from Jan 1, 1971 to Apr 30, 2000. After exploratory data anlysis which demonstrates the periodic structure, linear trend, and time-dependent residual structure, we fit a linear model with Fourier terms and ARMA(3,3) residuals as:

$$T_d = c_0 + c_1 d + \alpha_1 sin\left(\frac{2\pi d}{365}\right) + \beta_1 cos\left(\frac{2\pi d}{365}\right) + \alpha_2 sin\left(\frac{4\pi d}{365}\right) + \beta_2 cos\left(\frac{4\pi d}{365}\right) + \alpha_3 sin\left(\frac{6\pi d}{365}\right) + \beta_3 cos\left(\frac{6\pi d}{365}\right) + w_d,$$

where T_d is temperature for day d, $c_0, c_1, \alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3$ are coefficients, and error term $w_d \sim ARMA(3,3)$. The number of periodic terms in the model was determined using BIC. All parameters were estimated using maximum likelihood estimation/ordinary least squares.

Call the full paramater vector from the fitted model $\widehat{\theta}$ (not shown). This model with $\widehat{\theta}$ was used to to predict 92 days ahead, i.e. predicted daily temperatures in May 1 to July 31, and from these the cCDD index was computed. To obtain a distribution of predicted index values and payments from our model, one which incorporates parameter uncertainty of our model, we employed bootstrapping. The residuals from our fitted model were sampled with replacement to obtain $\widehat{\epsilon}^*$, added to the fitted values to obtain a bootstrapped dataset $y^* = \widehat{y} + \hat{\epsilon}^*$, and to this new sample a new model was fit which yielded parameter estimate $\widehat{\theta}^*$. We then used this new fitted model to project forward 92 days and obtain another predicted index value. This procedure was repeated 10,000 times, and the result was a predictive distribution of 10,000 cCDD settled index values. From these, we computed the 75^{th} and 95^{th} percentiles as 1155 and 1170, and used them as thresholds along with $k = 1$ in the pricing formula in Equation 2 for the index-based insurance product. Therefore, our two payments we consider are

$$Payment1 = max(0, CDD_d - 1155) \quad (4)$$

$$Payment2 = max(0, CDD_d - 1170). \quad (5)$$

We computed the hypothetical 10,000 payments under each threshold, and then computed the VaR and CTEs at the $p = 0.95$ and $p = 0.99$ levels from these distributiuons. These risk measures are shown under

DRAFT

the columns labelled Current in Table 4.

Next, we re-computed all quantities of the the predictive distribution of payments under the hypothetical future climate described by each regional climate model. This required two mappings, from "observed" to "current", and then from "current" to "future". The mapping from observed to current is simply the inverse of the bias and variance correction described in Algorithm 1. The specific parameters used are shown in Table 5. The mapping from "current" to "future" is simply the adjustment to the mean and variance as shown in Figure 10 for CRCM-CCSM. For the other RCMs, the adjustment parameters are shown in Table 6.

After making these adjustments to the 10,000 bootstrapped temperature realizations, we computed payments using equations 4 and **??** and then computed the risk measures VaR and CTE for these distributions. Results are shown in Table 4, and distributions of payments can be seen in Figure 12. As expected, risk measures are sharply higher, since higher average temperatures have pushed the right tail of cCDD distribution far to the right, piercing into higher payments.

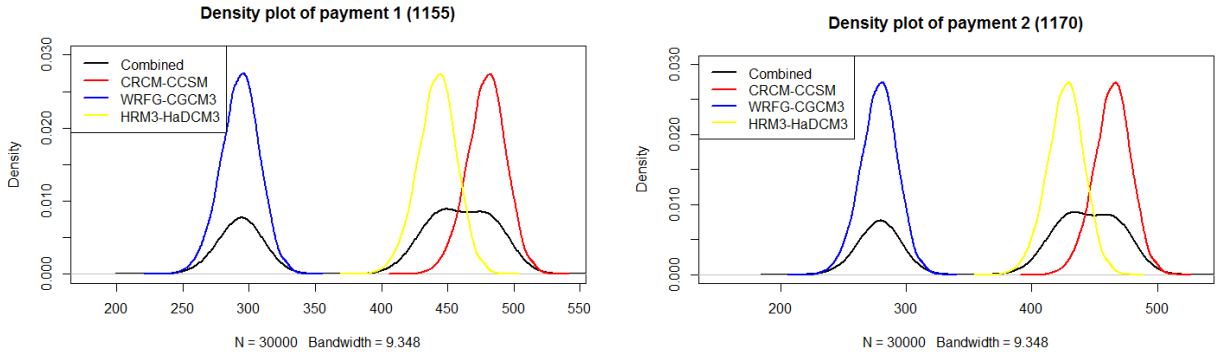

Figure 12: These are density of payments under different thresholds. The left panel shows payments when threshold is 1155, and the right panel shows payments when threshold is 1170. The red line shows density of payments under CRCM-CCSM model projection; the blue line shows density of payments under WRFG-CGCM3 model projection; the yellow line shows density of payments under HRM3-HadCM3 model projection; the black line shows mixed density of payments which puts 1/3 weights on each of these three models.

DRAFT

	Payment 1 (1155)				Payment 2 (1170)			
	Current	CRCM-CCSM	WRFG-CGCM3	HRM3-HaDCM3	Current	CRCM-CCSM	WRFG-CGCM3	HRM3-HaDCM3
95% VaR	14.22	502.81	316.81	465.36	0	487.81	301.81	450.36
99% VaR	24.28	512.93	326.94	475.48	9.28	497.93	311.94	460.48
95% CTE	20.17	508.75	322.75	471.3	20.77	493.75	307.75	456.3
99% CTE	28.46	517.11	331.12	479.66	28.46	502.11	316.12	464.66

Table 4: This table compares future payments under three different RCM-GCM models. The first payment has threshold 1155, and the second one has threshold 1170. First two rows show VaR, and last two rows show CTE. The "current" columns show risk measures for predictions of May.1st to July.31st, 2000, and other columns show risk measures for future payments calculated from corresponding models.

		CRCM-CCSM	WRFG-CGCM3	HRM3-HaDCM3
Bias Correction	Spring	-9.73	-4.35	-3.92
	Summer	0.46	-8.86	9.2
	Autumn	-6.76	-1.87	2.46
	Winter	-8.16	-1.99	-1.05
Variance Correction	Spring	0.84	1.03	0.88
	Summer	0.8	0.92	0.8
	Autumn	0.78	1.07	0.81
	Winter	0.83	0.81	0.97

Table 5: This table shows downscaling parameters used to map historical cCDDs to current model cCDDs for different models and different seasons. The first four rows are bias correction parameters for three models, and last four rows are variance correction parameters for the models. From bias correction parameters, we can tell all three models project 3 to 10 degrees different than true temperature. In spring and winter, the models project higher temperatures, and in summer, the models project lower temperatures. In autumn, CRCM-CCSM and WRFG-CGCM3 model project higher temperatures, and HRM3-HaDCM3 projects lower temperatures.

DRAFT

		CRCM-CCSM	WRFG-CGCM3	HRM3-HaDCM3
Bias Correction	Spring	-3.62	-2.69	-3.89
	Summer	-6.17	-3.59	-5.42
	Autumn	-4.3	-3.74	-4.96
	Winter	-3.63	-3.41	-3.74
Variance Correction	Spring	1.02	1.02	1.07
	Summer	1.01	0.96	1.05
	Autumn	0.99	1.03	1.05
	Winter	0.95	1.02	0.94

Table 6: This table shows downscaling parameters used to map downscaled current temperature and downscaled future temperature. After getting bias correction, we can adjust daily temperatures, and regard them as future temperatures in 2041-2070. Then we can use the adjusted temperatures to compute cCDDs and corresponding payments. From the table, we can also notice downscaled future temperature is about 2 to 4 degrees higher than current, which results in increase of payments.

DRAFT

5 Discussion

In this paper, we carefully documented how to obtain, visualize, and utilize both historical data as well as regional climate projections for the purposes of exploring shifts in distributions over time. We documented the known issue of bias and variance in regional climate models by comparing RCM hindcasts to observed temperatures for the same period, and showed how one can define bias and variance corrections to produce statistically downscaled future climate model projections. We defined an ensemble of multiple regional climate models, and showed how actuaries can explore projected consequences of changes across an ensemble. The primary purpose of this paper is to give actuaries a working example of all of these issues. All code for this paper was written in `R`, and is freely obtained at `http://users.wfu.edu/erhardrj/`.

Working with climate model output and environmental data may well require that actuaries look to resources beyond those commonly encountered to build a better foundation in climate and environmental risks. Erhardt (2017) contains a summary of a wide range of datasets, background climate literacy books, and examples of measuring and managing climate risks within the insurance industry. IPCC summaries are also excellent sources to build background climate literacy.

Although this paper focused on temperature variables in climate models, other variables in RCMs which may be of interest to actuaries include wind speed and precipitation. Precipitation is zero-inflated, and wind speeds are strictly positive and more heavily right-skewed than temperature data. These may alter the specific choice for a bias and variance correcting algorithm, however the overall procedure described in this paper would still apply.

The precise changes in risk measures that were documented in the hypothetical index-based insurance are not the primary conclusions of the research. Indeed, the regional climate models considered here were all run under the RCP 8.5 carbon emissions scenario. This is the most pessimistic scenario considered by the IPCC, and there are other more modest scenarios for anthropogenic carbon emissions (namely, RCM 2.6, 4.5, and 6). Unfortunately, NARCCAP is only run under RCP 8.5, and this therefore is a limitation. However, as the number and quality of regional climate models grows, actuaries should increasingly be able to identify regional climate models which best suit their needs. The techniques described in this paper should help them develop algorithms to utilize those projections to continue the precise measure and management of risks.

DRAFT

Acknowledgements

The authors wish to thank the Society of Actuaries Research Expanding Boundaries (REX) Funding Pool for their generous financial support of this project. The authors further wish to thank Scott Lennox and Erika Schulty at the SOA for their leadership and guidance, and the volunteers who formed the Project Oversight Group. All provided valuable perspectives and feedback which strengthened this paper.

References

Alaton, P., Djehiche, B., and Stillberger, D. (2002). "On modelling and pricing weather derivatives." *Applied mathematical finance*, 9, 1, 1–20.

Alderman, H. and Haque, T. (2007). *Insurance against covariate shocks: The role of index-based insurance in social protection in low-income countries of Africa*. No. 95. World Bank Publications.

Barnett, B. J. and Mahul, O. (2007). "Weather index insurance for agriculture and rural areas in lower-income countries." *American Journal of Agricultural Economics*, 89, 5, 1241–1247.

Carter, M., de Janvry, A., Sadoulet, E., Sarris, A., et al. (2014). "Index-based weather insurance for developing countries: A review of evidence and a set of propositions for up-scaling." *Development Policies Working Paper*, 111.

Caya, D. and Laprise, R. (1999). "A semi-implicit semi-Lagrangian regional climate model: The Canadian RCM." *Monthly Weather Review*, 127, 3, 341–362.

Chantarat, S., Mude, A. G., Barrett, C. B., and Carter, M. R. (2013). "Designing index-based livestock insurance for managing asset risk in northern Kenya." *Journal of Risk and Insurance*, 80, 1, 205–237.

Collier, B., Skees, J., and Barnett, B. (2009). "Weather index insurance and climate change: opportunities and challenges in lower income countries." *The Geneva Papers on Risk and Insurance-Issues and Practice*, 34, 3, 401–424.

Deng, X., Barnett, B. J., Vedenov, D. V., and West, J. W. (2007). "Hedging dairy production losses using weather-based index insurance." *Agricultural Economics*, 36, 2, 271–280.

DRAFT

Eagleson, P. S. (2011). *Land surface processes in atmospheric general circulation models*. Cambridge University Press.

Erhardt, R. (2015). "Mid-twenty-first-century projected trends in North American heating and cooling degree days." *Environmetrics*, 26, 2, 133–144.

— (2017). *Climate, Weather and Environmental Sources for Actuaries*. Society of Actuaries.

Erhardt, R. and Von Burg, R. (2018). *How do they know and What could we do? The science of 21st century climate projections and opportunities for actuaries*. Society of Actuaries.

Fischer, T., Su, B., Luo, Y., and Scholten, T. (2012). "Probability Distribution of Precipitation Extremes for Weather Index–Based Insurance in the Zhujiang River Basin, South China." *Journal of Hydrometeorology*, 13, 3, 1023–1037.

Fleege, T. A., Richards, T. J., Manfredo, M. R., Sanders, D. R., et al. (2004). "The performance of weather derivatives in managing risks of specialty crops." In *NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management, St. Louis, Missouri*, 19–20.

Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne, S. R., Lawrence, D. M., Neale, R. B., Rasch, P. J., Vertenstein, M., et al. (2011). "The community climate system model version 4." *Journal of Climate*, 24, 19, 4973–4991.

Hansen, J., Ruedy, R., Sato, M., and Lo, K. (2010). "Global surface temperature change." *Reviews of Geophysics*, 48, 4.

Hansen, J., Sato, M., and Ruedy, R. (2012). "Perception of climate change." *Proceedings of the National Academy of Sciences*, 109, 37, E2415–E2423.

(MCII), M. C. I. I. (2013). *Climate risk adaptation and insurance. Reducing vulnerability and sustaining the livelihoods of low-income communities. Report No. 13. Bonn: United Nations University Institute for Environment and Human Security*. (UNU-EHS).

Mearns, L., McGinnis, S., Arritt, R., Biner, S., Duffy, P., Gutowski, W., Held, I., Jones, R., Leung, R., Nunes, A., et al. (2007). "The North American regional climate change assessment program dataset." *National Center for Atmospheric Research Earth System Grid data portal, Boulder, CO*.

DRAFT

Mearns, L. O., Arritt, R., Boer, G., Caya, D., Duffy, P., Giorgi, F., Gutowski, W., Held, I., Jones, R., Laprise, R., et al. (2005). "NARCCAP, North American Regional Climate Change Assessment Program." In *16th Conference on Climate Variability and Change*.

Music, B., Frigon, A., Lofgren, B., Turcotte, R., and Cyr, J.-F. (2015). "Present and future Laurentian Great Lakes hydroclimatic conditions as simulated by regional climate models with an emphasis on Lake Michigan-Huron." *Climatic Change*, 130, 4, 603–618.

Nakicenovic, N., Alcamo, J., Grubler, A., Riahi, K., Roehrl, R., Rogner, H.-H., and Victor, N. (2000). *Special Report on Emissions Scenarios (SRES), A Special Report of Working Group III of the Intergovernmental Panel on Climate Change*. Cambridge University Press.

Pachauri, R. K., Allen, M. R., Barros, V. R., Broome, J., Cramer, W., Christ, R., Church, J. A., Clarke, L., Dahe, Q., Dasgupta, P., et al. (2014). *Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change*. IPCC.

Turvey, C. G. (2005). "The pricing of degree-day weather options." *Agricultural Finance Review*, 65, 1, 59–85.

Turvey, C. G. and Mclaurin, M. K. (2012). "Applicability of the Normalized Difference Vegetation Index (NDVI) in index-based crop insurance design." *Weather, Climate, and Society*, 4, 4, 271–284.

Zeng, L. (2000). "Weather derivatives and weather insurance: concept, application, and analysis." *Bulletin of the American Meteorological Society*, 81, 9, 2075–2082.

DRAFT

A Appendices

Table A shows the names of the variables used in downscaling algorithm.

Variable	Explanation
$T_{historical,t}$	Re-stated historical temperature in year t
$T_{current,t}$	Re-stated current temperature in year t
$T_{future,t}$	Re-stated future temperature in year t
$\bar{T}_{historical}$	Average of re-stated historical temperature
$\bar{T}_{current}$	Average of re-stated current temperature
$\bar{T}_{future}$	Average of re-stated future temperature
$\mathbb{T}_{current,t}$	Downscaled current temperature before re-trending
$\mathbb{T}_{future,t}$	Downscaled future temperature before re-trending
$\sigma_{historical}$	Standard deviation of re-stated historical temperature
$\sigma_{current}$	Standard deviation of re-stated current temperature
ψ	Bias correction
ϕ	Variance correction

A.1 Re-stating and Re-trending Algorithms

A.1.1 Re-stating

Algorithm 2 Re-stating

Variables:
T: temperature before re-stating
Y_t: year t
Y: base year
$\tilde{T}$: re-stated temperature
$\hat{\beta}$: linear coefficient
$\hat{\epsilon}$: fitting error

1: Fit linear models:
$$T = \hat{\beta}Y + \hat{\epsilon}$$
2: $\tilde{T} = T + \hat{\beta}(Y - Y_t)$

DRAFT

A.1.2 Re-trending

Algorithm 3 Re-trending

Variables:
T: downscaled temperature
Y_t: year t
Y: base year
$\tilde{T}$: re-trended downscaled temperature
$\hat{\beta}$: linear coefficient

1: $\tilde{T} = T - \hat{\beta}(Y - Y_t)$

DRAFT