

Incorporating Spatial Dependence and Climate Change Trends for Measuring Long-Term Temperature Derivative Risk

by Robert Erhardt

ABSTRACT

In this paper we explore a method to model the financial risks of holding portfolios of long-term temperature derivatives for any subset of the 30 North American cities whose derivatives are actively traded on the Chicago Mercantile Exchange (CME). Long-term derivatives are those whose period of accrual for degree days is substantially longer than the temporal autocorrelation of daily temperature data, and therefore accruals can be modeled with a multivariate normal distribution. One commonly traded temperature derivative on the CME has a 6-month index period, which satisfies this long-term condition. The method presented incorporates spatial dependence among the cities, and allows for possible trends in degree days due to climate change. Though limited to long-term contracts, the method is mathematically and computationally quite simple and applicable to some of the most commonly traded temperature derivatives. Possible implications for the insurance industry are discussed.

KEYWORDS

Chicago Mercantile Exchange, cooling degree day, heating degree day, weather derivative

1. Introduction

Weather risk is any financial impact that a business or institution may face as a result of a weather-related cause. A commonly discussed example of weather risk is when a utility company in a particular region faces an unusually hot summer or unusually cold winter; unable to meet the increased energy demand for heating or cooling, this company may be forced to import electricity from farther away (Alexandridis and Zapanis 2013). Other weather risks include snowfall total amounts (skiing industry), dates of a first frost (agriculture), hurricane activity (tourism), and so forth. In each case a business faces a direct increase in expenses or decrease in revenue as a result of a weather event. The Chicago Mercantile Exchange (CME) has a weather products division (<http://www.cmegroup.com/trading/weather>) where tailored financial contracts designed to address weather risk are traded. A particular class of financial products known as weather derivatives were first listed on the CME in 1999 (Kunreuther and Michel-Kerjan 2009), the same year that the Weather Risk Management Association (WRMA, <http://www.wrma.org/>) was formally chartered. In 2011 the size of the weather risk market had grown to an estimated 11.8 billion U.S. dollars (Benth and Benth 2013).

Weather derivatives link a set of payments to a set of predetermined weather outcomes. Common references include Richards, Manfredo, and Sanders (2004), Dischel (2002), and Geman (1999). The most commonly traded weather derivative is the temperature derivative, which defines payments solely based on temperature outcomes. The building block for temperature derivatives are *heating and cooling degree days*, defined for a single day at a single location as

$$HDD = \max(65 - T, 0)$$

$$CDD = \max(T - 65, 0),$$

where T is the average daily temperature measured in degrees Fahrenheit. The use of 65 degrees Fahrenheit

in both definitions is arbitrary. A convenient consequence of using a common value in both definitions is that one can always recover the daily observed temperature as

$$T = 65 + CDD - HDD.$$

Heating and cooling degree days are meant to measure the overall demand for heating and cooling as a function of a day's temperature. Consider an arbitrary location where on day one the average temperature T is 70 degrees F. This one day generates 5 CDDs and 0 HDDs. If on the following day the average temperature were again $T = 70$ degrees F, this second day would generate 5 CDDs and 0 HDDs. One could then begin running sums of these two quantities—called *cumulative cooling degree days* (cCDDs) and *cumulative heating degree days* (cHDDs)—which would be 10 and 0, respectively. These indices are used to capture the overall demand for heating or cooling within some time frame.

A special class of weather derivatives known as *temperature derivatives* can be built from these indices. Here, the term *derivative* is used to indicate that the product derives its value from something else—in this case, the actual temperature. To illustrate by example, consider a large utility company that provides electricity to a city in the month of June. Suppose that the average daily temperature in June in this city is around 73 degrees F, and thus cumulative cooling degree days in the month of June often settle around 240 (8 CDDs times 30 days). To protect against a surge in demand, the company may wish to buy a weather derivative which pays in proportion to excess heat. One method would be to purchase a temperature derivative which pays \$10,000 for each CDD in excess of 360, and \$0 otherwise; mathematically, the payoff is

$$P = 10,000 \cdot \max(cCDD - 360, 0). \quad (1)$$

It should be clear that this sort of contract offers proportional protection during exceptionally hot months of June, and no financial protection during more typical months of June. It should also be clear that one

actuarial method of determining the premium would consider the *cCDD* index as a random variable. Deriving or estimating this distribution would allow the actuary to infer the resulting distribution of payments P , which of course then serves as the starting point for an actuarial method of pricing or ratemaking.

The temperature derivative described thus far applies to a single location. For multiple locations, one should realize that the financial performance of those locations are dependent, as the underlying weather variable could be dependent. In this paper we explore how the spatial dependence of temperature derivatives impacts financial risks. Currently, temperature derivatives are actively traded on the CME for 30 North American cities, whose locations are shown in Figure 1. Since the temperature at two nearby locations can be closely related, so too can the financial outcomes of temperature derivatives for those cities. Spatial dependence in weather derivatives has been explored in limited cases by Saltyte-Benth, Benth, and Jalinskas (2007), Saltyte-Benth and Saltyte (2011), and Erhardt and Smith (2014), but each paper considers limited cases. Here we aim to extend the reach of such cases by covering a wider class of temperature derivatives.

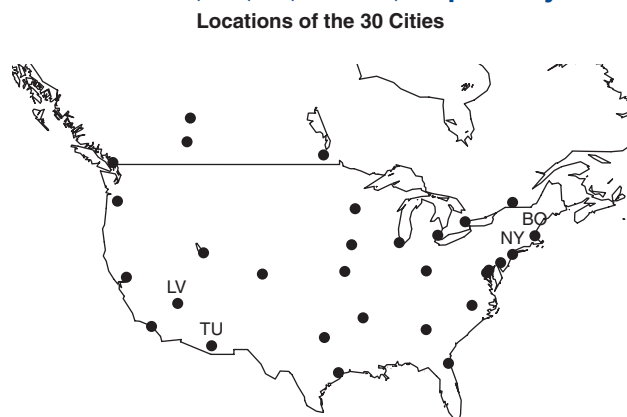
In this paper, we develop a methodology for measuring the risk of holding an arbitrary portfolio of

CME-traded temperature derivatives with a specific eye towards spatial dependence of locations. We take the point of view of a company which holds a portfolio of temperature derivatives, some of which will involve payments triggered by weather events. Whether the company directly sold the derivatives to various buyers or came to hold the obligations through trading is immaterial. What matters is the best estimate of combined risk from holding such a portfolio. Therefore the paper will focus on estimating risk measures for losses, and will not focus on the premium/revenue the company earned by accepting the risk in the first place.

Further, we explore quantifying the risk from temperature derivatives with an eye towards trends due to climate change. The fifth assessment report from the Intergovernmental Panel on Climate Change has been released (IPCC 2013), and it predicts global increases in average temperatures. It also states that “it is extremely likely that human influence has been the dominant cause of the observed warming since the mid-20th century.” Computer model experiments under different carbon emissions scenarios give projections of future temperature rise in line with future greenhouse gas emissions. The *Bulletin of the American Meteorological Society* called for “very long-term hedging contracts” to help manage climate risk (Dutton 2002). It seems likely that increasing concerns over climate change will continue to have an impact on temperature derivative markets, and climate trends must be incorporated into their study. The methodology described in this paper incorporates long-term trends in temperature due to climate change.

Weather risk markets have the potential to impact the way some insurance companies operate with regard to weather-related risks. Mills (2005) discussed the scope of climate change risk to the insurance industry. Erhardt and Smith (2014) explored a connection between weather derivatives, extremes, and insurance. Since the financial outcomes of weather derivatives are minimally correlated with other financial outcomes facing the insurance industry (Alexandridis and Zaprana 2013), investing in weather derivatives is one approach that an insurance company can take

Figure 1. Locations of the 30 North American cities whose temperature derivatives are listed on the Chicago Mercantile Exchange. The four cities, Boston, New York, Las Vegas, and Tucson (which are highlighted throughout this paper), are labeled BO, NY, LV, and TU, respectively



to maximize returns for a given risk level. Additionally, since the events which trigger a weather derivative payment arise from scientific laws governing the weather, the construction of probabilistic models for an *actuarial* method of pricing and measuring risk is possible (Jewson, Brix, and Ziehmann 2005).

2. The data and risk measures

2.1. Data

Daily temperature data were freely obtained from the National Climate Data Center (<http://www.ncdc.noaa.gov/cdo-web/>). Specifically, we obtained data for the 30 North American cities whose temperature derivatives are listed on the Chicago Mercantile Exchange. Locations of the cities are shown in Figure 1. At each location, measurements include the maximum daily temperature T_{max} and minimum daily temperature T_{min} in degrees Fahrenheit for the period January 1, 1945, through August 30, 2014. Although SI units are common in much of the world, the Chicago Mercantile Exchange lists all North American temperature derivatives in degrees Fahrenheit, so we follow this convention.

Let $d = 1, \dots, D = 30$ index the 30 cities, and let t be an index for time. Following the conventions used in weather derivative pricing (Jewson, Brix, and Ziehmann 2005; Alexandridis and Zapranis 2013), we define the daily observed temperature as

$$T_{obs}(d, t) = \frac{1}{2}(T_{min}(d, t) + T_{max}(d, t)).$$

For each day, cooling degree days are computed as

$$CDD(d, t) = \max(T_{obs}(d, t) - 65, 0),$$

and heating degree days are computed as

$$HDD(d, t) = \max(65 - T_{obs}(d, t), 0).$$

Cooling and heating degree days are most commonly used to construct cumulative indices over some time

period of interest. Let $\{t \in \mathcal{T}\}$ be the set of all days within the time period of interest $\mathcal{T}$. Then a cumulative CDD index is simply

$$cCDD(d, \mathcal{T}) = \sum_{t \in \mathcal{T}} CDD(d, t),$$

where $\mathcal{T}$ is a time period of interest—most commonly one of the calendar months, a set of adjacent months, or one of the six-month periods April 1–Sep 30 or October 1–March 31 (Jewson, Brix, and Ziehmann 2005). Six-month contracts are commonly traded, and have been studied in other papers (Campbell and Diebold 2005; Erhardt 2014).

2.2. Actuarial risk measures

It can be useful to imagine a company that sells derivatives to various buyers for some premium, and then holds these obligations as a risk—the hope is that the realization of this risk is less than the total revenue brought in at the sale. Call L the random loss that the institution will ultimately have to pay. The exact realization of L comes after the last settlement date has passed. The specific mathematical relationship linking this loss L to the degree-day settlement index ($cCDD(\mathcal{T})$ or $cHDD(\mathcal{T})$) will be outlined in later sections. Here it suffices to see that L is simply a function based on the random variable temperature, T , and that one can therefore build up distributions of L by constructing distributions of temperature T and degree-day indices.

We will estimate values for various actuarial risk measures as defined in Kaas et al. (2008). Definitions for risk measures in the actuarial literature are not universally agreed upon, so care should be taken. The value at risk is defined as

$$VaR(L; \alpha) = Q_\alpha = \min\{Q : P(L \leq Q) \geq \alpha\}. \quad (2)$$

The VaR is easily obtained since the loss L is often a continuous random variable for larger values of α , and therefore the distribution function $F_L(l) = P(L \leq l)$ is strictly increasing and has inverse distribution func-

tion F_L^{-1} . Then we simply have $Q_\alpha = F_L^{-1}(\alpha)$. Kaas et al. (2008) define the conditional tail expectation as

$$CTE_\alpha(L) = \mathbb{E}(L | L > Q_\alpha), \quad (3)$$

which is the expected loss conditional upon exceeding the VaR.

3. Long-term modeling of CDD and HDD indices

Here we will always refer to cumulative CDDs for the time period April 1–Sep 30, and cumulative HDDs for the time period October 1–March 31. These are two of the more commonly traded temperature deriv-

atives on the CME (Campbell and Diebold 2005). Summing degree days over six months allows one to use the normal approximation, since the length of summation is substantially longer than the length of temporal dependence in the data. The methodology is therefore substantially simpler than that needed for modeling shorter time periods.

Figure 2 shows the October–March cHDDs for the four representative cities. For each city, the marginal distribution is roughly normal. Neighboring pairs Boston–New York and Las Vegas–Tucson each show strong positive dependence in cHDDs in the respective scatterplots, while non-neighboring pairs show little to no dependence, as expected. Figure 3 reinforces this point by plotting the linear correlation

Figure 2. October–March cumulative heating degree days for four representative cities. The neighboring pairs of Boston–New York and Las Vegas–Tucson show strong positive dependence, while pairs of cities much farther apart show little dependence

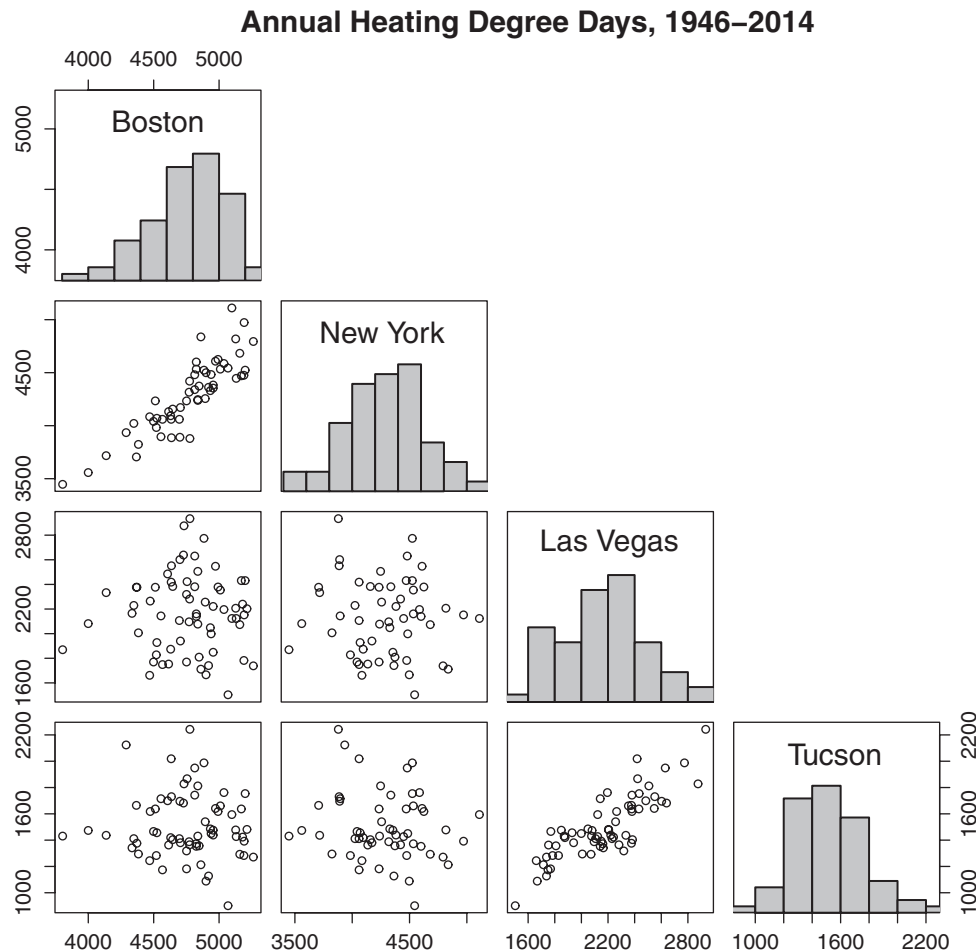
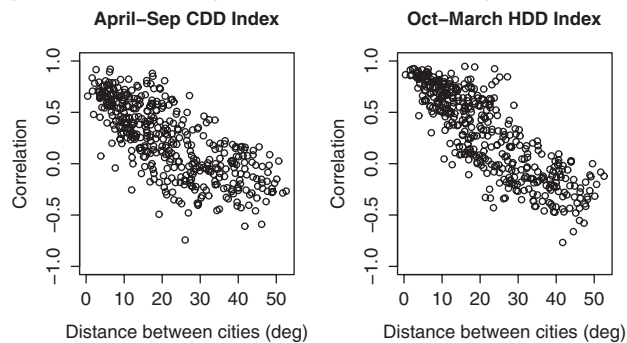


Figure 3. Linear correlation in the cCDDs and cHDDs as a function of the distance between the pairing of cities (measured in degrees)



between two cities as a function of their geographic distance from one another (measured in degrees Lat-Lon). The general trend for both heating and cooling is high correlation at small distances, which drops off as distance increases.

Figures 4 and 5 highlight positive trends in cCDDs and negative trends in cHDDs over time at the four representative cities. These trends are a result of hotter summers and milder winters since the mid-twentieth century. For all 30 cities whose linear trends were

Figure 4. Linear trends in the April–September cumulative cooling degree days for four representative cities. Slopes for Las Vegas and Tucson are significantly different from zero at the 0.05 level

Trends in April–September Cooling Degree Days

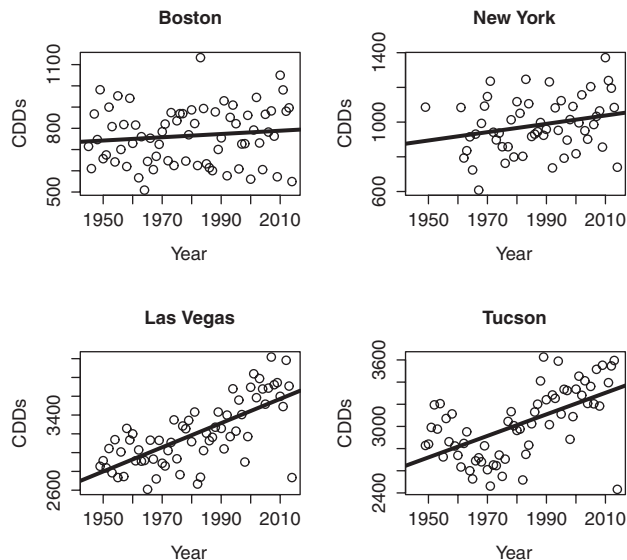
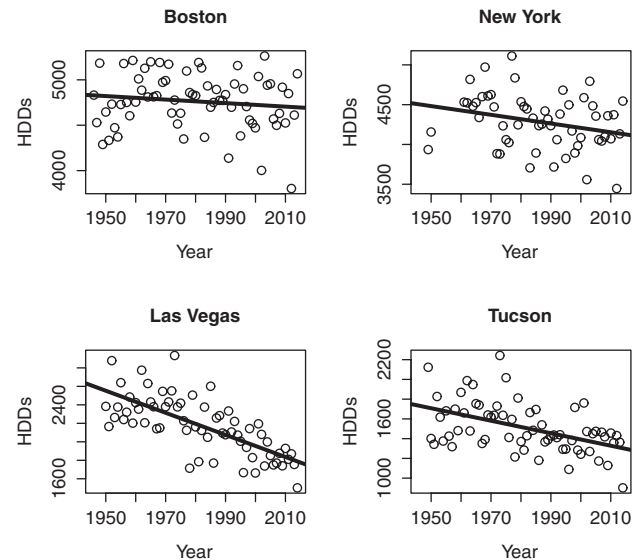


Figure 5. Linear trends in the October–March cumulative heating degree days for four representative cities. Slopes for Las Vegas, New York, and Tucson are significantly different from zero at the 0.05 level

Trends in October–March Heating Degree Days



significant at the 0.1 level, cCDDs and cHDDs were de-trended and put on a common year 2014 level. We did not de-trend data for a city whose trend estimates were not significant at the 0.1 level. Some statistical output for trend estimation is shown in Table 1.

Readers may wonder about possible long distance spatial and/or temporal dependence introduced by climatic episodes such as El Niño, La Niña, the North Atlantic Oscillation, and other large-scale climatic phenomena. We investigated the possibility that historical cCDD and cHDD totals for some of the 30 cities should be adjusted using the Oceanic Niño Index (ONI) (http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml). Specifically, we calculated the April–October average ONI for the period 1950–2014, and compared this to the historical cCDDs for each of the 30 cities over the same time period. Similarly, we computed November–March average ONI values of the ONI and compared these to the cHDDs for each of the 30 cities from the period 1951–2014. We found no statistically significant trends or relationships that

Table 1. Annual trend estimates for cumulative heating and cooling degree day indices for the four cities highlighted in Figures 4 and 5.

Index	City	$\widehat{Trend}$	$\widehat{SE}(\widehat{Trend})$	t-Stat	p-Value
cCDDs	Boston	0.771	0.798	0.966	0.337
	New York	2.399	1.283	1.870	0.067
	Las Vegas	12.878	1.624	7.931	$4.2 \cdot 10^{-11}$
	Tucson	9.710	1.720	5.644	$4.2 \cdot 10^{-7}$
cHDDs	Boston	-1.927	1.830	-1.053	0.296
	New York	-5.519	2.659	-2.075	0.043
	Las Vegas	-11.932	1.459	-8.177	$1.7 \cdot 10^{-11}$
	Tucson	-6.334	1.422	-4.456	$3.4 \cdot 10^{-5}$

suggested an adjustment. This does not suggest that future research will not uncover statistical value in using climate indices such as the ONI as covariates. For example, it may be that there is a lag of several months between a high ONI value and its impact on North American temperature, and perhaps this temporal lag may differ for the 30 cities in North America based on their geographic location. All we state here is that the spatial dependence among locations is the largest omission to current actuarial temperature derivative pricing, and that further adjustments based on the ONI and similar indices are outside the scope of this paper. To recap, then, we have de-trended data for long-term climate trends but not any climate indices, and so the data x_i discussed in the next section always refer to the de-trended temperature values.

3.1. The model for cCDDs and cHDDs

The longest temporal autocorrelations observed for daily temperature data were about 14 days long. None of the cities showed any remaining temporal autocorrelation in annual indices from one year to the next. Since cumulative totals were computed by summing CDDs or HDDs over roughly 180+ days, these “long-term” indices were sufficiently long to produce normally distributed margins without autocorrelation by year. The natural choice for a model is to let X_i be the $D = 30$ dimensional random vector of cCDD (or cHDD) indices for year $i = 1, \dots, n$, and to model X_i as multivariate normal,

$$X \sim \mathcal{N}(\mu_X, \Sigma_X),$$

where μ_X is the mean vector and Σ_X is the covariance matrix. Readers familiar with spatial interpolation and kriging should recognize that our goal is only to model the dependence at the particular locations of the 30 cities listed on the CME; therefore, it isn't necessary to choose a Gaussian process model, interpolate to unobserved locations, and so forth. The 30 by 30 dimension covariance matrix Σ_X is easily invertible from a computational cost standpoint. The covariance is estimated with the usual unbiased estimator

$$\hat{\Sigma}_X = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})^T. \quad (4)$$

The mean of the multivariate normal is estimated using

$$\hat{\mu}_X = \frac{1}{n} \sum_i x_i. \quad (5)$$

4. Models for losses

4.1. Affine models

An affine transformation of random vector X is any function of the form $a + BX$. Since our model assumes that cumulative degree days X follow a multivariate Gaussian distribution with D dimensional mean vector μ_X and D by D covariance matrix Σ_X , then any affine transformations of X will similarly be Gaussian with only a change in the mean and variance. We begin by considering a simple weather

derivative that pays \$20 per contract per degree day, with the sign of the payment determined by an entry level. Mathematically, this payment is

$$P_i = 20n_i(X_i - E_i) = -20n_iE_i + 20n_iX_i,$$

where n_i is the number of contracts for city i , X_i is the settlement value of the cumulative degree day index, and E_i is the *entry level* for city i . For the affine model of payments, the entry level refers to the point at which the sign of payments changes from positive to negative. For these weather derivatives there is always a payment; we will consider the more common case where payments are either positive or exactly zero in the next subsection. The amount \$20 is selected since standard contracts traded on the CME pay \$20 per degree per contract. As a simple example, if a company held 100 Atlanta cCDD contracts for the April 1–October 31 period with entry level 2000 cCDDs, but the index ultimately settled at 2200 cCDDs, the company would lose $100 \cdot 20 \cdot (2200 - 2000) = 400,000$.

Define vector c^T as

$$c^T = (-20n_1E_1, \dots, -20n_{30}E_{30}),$$

and square matrix B as

$$B = \begin{pmatrix} 20n_1 & 0 & \cdots & 0 \\ 0 & 20n_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 20n_{30} \end{pmatrix}$$

Then we can express the payment vector $P = c + BX$, which is an affine transformation of vector X , which has a known multivariate normal distribution. The distribution of P is also multivariate normal, as

$$P \sim \mathcal{N}(c + B\mu_X, B\Sigma_X B^T). \quad (6)$$

The quantity which is ultimately of interest is $L = \Sigma_i P_i$, the total aggregate scalar loss. If we define vector $b^T = (1, \dots, 1)$ we can write $L = b^T P$, which is again

an affine transformation and therefore the distribution of L is normal as

$$L \sim \mathcal{N}(b^T(c + B\mu_X), b^T B\Sigma_X B^T b) = \mathcal{N}(\mu_L, \sigma_L^2),$$

where μ_L is the univariate mean of the loss, and σ_L^2 is the variance of the loss. Observe that one model fit to the CDD and HDD indices can be used to estimate financial outcomes for any arbitrary collection of temperature derivatives. Different collections will involve different choices of the user-selected quantities b , c , and B , but the terms μ_X and Σ_X will be the same. Estimates of the mean and variance of the financial payments L are simply

$$\hat{\mu}_L = b^T(c + B\hat{\mu}_X)$$

and

$$\hat{\sigma}_L^2 = b^T B\hat{\Sigma}_X B^T b.$$

The Value at Risk (defined in Equation 2) is estimated by

$$\widehat{VaR}(L; \alpha) = \hat{Q}_\alpha = \min \left\{ Q : \Phi \left(\frac{Q - \hat{\mu}_L}{\hat{\sigma}_L} \right) \geq \alpha \right\},$$

where $\Phi(\cdot)$ is the cumulative distribution function for the standard normal. A closed-form expression for the estimator of $CTE_\alpha(L)$ (which was defined in Equation 3) can be derived by recognizing it is simply the expectation of a truncated normal (Green 2003),

$$\widehat{CTE}_\alpha(L) = \hat{\mu}_L + \frac{\hat{\sigma}_L}{1 - \Phi \left(\frac{Q_\alpha - \hat{\mu}_L}{\hat{\sigma}_L} \right)} \phi \left(\frac{Q_\alpha - \hat{\mu}_L}{\hat{\sigma}_L} \right),$$

where $\phi(\cdot)$ is the standard normal density function.

4.2. Extending to strike values

While the previous subsection provided some nice mathematical results and a simple framework for estimating some risk measures, nearly all weather deriva-

tives have a payment that is non-linear in relationship to degree day indices, and typically there are no negative payments. Instead, there is a large probability of a zero dollar payment which occurs when the strike is not exceeded. Here we show how the previous mathematical results can be used with a computational approach to consider weather derivatives whose payments are identically zero *unless* the entry level is exceeded by the degree day index. Mathematically, payments are

$$P_i = 20n_i \max(X_i - E_i, 0)$$

where n_i is the number of contracts for city i , X_i is the value the cumulative degree day index settles at, and E_i is the *strike level* for city i —the amount that, once exceeded, dictates that *non-zero* payments are triggered. The key difference from the affine model is that now, in the event that $X_i \leq E_i$, the payment is zero instead of negative. Hence it is the distribution of $P_i | P_i > 0$ that is a continuous distribution, and the distribution of P_i itself is a mixed discrete-continuous distribution with a point mass at $P_i = 0$.

Recall that our ultimate interest is in the quantity $L = \sum_i P_i$. The nonlinearity of the payment P_i makes closed-form distribution functions of L difficult to compute, so instead we turn to a computational approach:

1. Simulate a realization P' from the Gaussian distribution shown in Equation 6.
2. For $i = 1, \dots, 30$ define $R'_i = \max(P'_i, 0)$. This is simply replacing negative elements of P' with zero.
3. Compute $L' = \sum_i R'_i$
4. Repeat steps (1) - (3) J times.

The result is a collection of J realizations of the random variable L , and from this distribution one can compute empirical estimates of all desired risk measures. In general J should be a very large number, in the hundreds of thousands or millions, which costs very little in terms of computational power. This output can be used to estimate risk measures as

$$\widehat{VaR}(L; \alpha) = w_1 \cdot L_{[j]} + w_2 \cdot L_{[j+1]} \quad (7)$$

where $j/J < \alpha < (j+1)/J$, $L_{[j]}$ is the j^{th} order statistic, and $w_1 + w_2 = 1$ are the weights (whose relevance vanishes as $J \rightarrow \infty$). The conditional tail expectation can be estimated as

$$\widehat{CTE}_\alpha(L) = \frac{1}{|\tau|} \sum_j L_j \cdot I_{\{L_j > \widehat{VaR}(L; \alpha)\}}, \quad (8)$$

where $|\tau|$ is the number of losses above $\widehat{VaR}(L; \alpha)$ and $I_{\{L_j > \widehat{VaR}(L; \alpha)\}}$ is the indicator function which takes value 1 when the argument holds, and 0 otherwise.

5. Results

Estimated mean vectors $\hat{\mu}_x$ and covariance matrixes $\hat{\Sigma}_x$ for both cCDDs and cHDDs are shown in Tables 2 and 3. Since these tables can be used with any choice of vectors c and matrix B , they contain all information needed to estimate aggregate losses for a collection of April–September cCDD and October–March cHDD temperature derivatives. Spatial dependence is naturally incorporated through the off-diagonal elements of $\hat{\Sigma}$. Recall that since the data were de-trended to 2014 levels, some recognition of climate change trends was also incorporated.

Here we demonstrate the approach to computing the risk measures for the aggregate loss on a portfolio of cCDDs using the strike model. Suppose an insurance company wishes to hold 100 contracts at each of the four cities highlighted in this paper (Boston, Las Vegas, New York, and Tucson). For each of these four cities, the strike level E_i will be 100 CDDs higher than average cCDD values taken from $\hat{\mu}_x$ in Table 2. Thus, the four strike levels are 865, 3723, 1148, and 3442. We have selected strike values above the means for these four cities to demonstrate how zero dollar payments arise when none of the cities exceeds their respective strike values. To clarify, the company will pay amounts $20 \cdot 100 \cdot \max(X_i - E_i, 0)$. The total loss is obtained by adding the individual losses, but a closed-form density for L is not easily obtained so we will use the simulation based approach in Section 4.2, beginning with

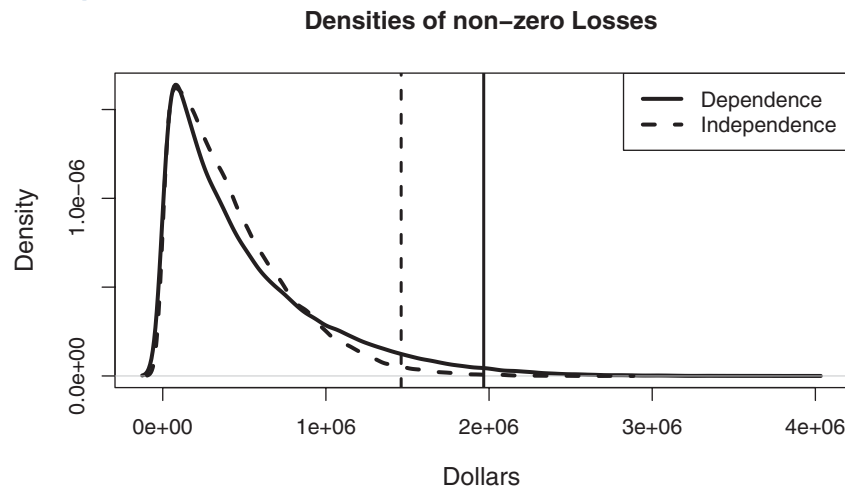
Table 2. Estimated mean vector $\hat{\mu}_x$ (from Equation 5) and covariance matrix $\hat{\Sigma}_x$ (Equation 4) for annual cooling degree totals. Each entry of $\hat{\Sigma}_x$ shown is expressed as a fraction of the largest estimated covariance 106,851.5 and rounded to three decimal places, so the largest entry on the table is 1.000.

μ	Atl	Bal	Bos	Chi	Cin	Col	Dal	Des	Det	Hou	Jac	Kan	Las	Lit	Los	Min	New	Phi	Por	Ral	Sac	Sal	Tuc	Was	Cal	Edm	Mon	Tor	Van	Win
Σ	Atl	Bal	Bos	Chi	Cin	Col	Dal	Des	Det	Hou	Jac	Kan	Las	Lit	Los	Min	New	Phi	Por	Ral	Sac	Sal	Tuc	Was	Cal	Edm	Mon	Tor	Van	Win
Atl	.608	.251	.097	.159	.232	.119	.431	.176	.098	.192	.286	.139	.165	.428	-.104	.110	.173	.243	-.060	.354	-.156	.004	.309	.145	-.014	-.008	.076	.134	-.005	.070
Bal	.251	.301	.150	.204	.226	.072	.156	.194	.180	.120	.181	.271	.013	.228	-.091	.177	.195	.257	-.039	.230	-.046	-.006	.064	.135	-.005	.000	.120	.164	-.011	.089
Bos	.097	.150	.170	.139	.099	.070	.075	.154	.149	.035	.034	.119	-.005	.098	-.004	.142	.141	.159	-.035	.101	.006	.036	.025	.093	.003	.005	.114	.136	-.014	.120
Chi	.159	.204	.139	.376	.236	.111	.166	.279	.219	.046	.092	.352	-.089	.252	-.105	.246	.171	.208	-.082	.141	-.056	.006	-.015	.180	.000	.003	.162	.236	-.019	.208
Cin	.232	.226	.099	.236	.273	.050	.069	.193	.157	.044	.079	.279	-.083	.212	-.114	.128	.134	.175	-.046	.176	-.066	-.064	-.033	.113	-.001	.002	.072	.128	-.014	.089
Col	.119	.072	.070	.111	.050	.155	.215	.153	.084	.077	.007	.138	.104	.175	-.023	.130	.087	.077	-.059	.076	-.040	.070	.091	.014	.002	-.001	.064	.087	-.015	.151
Dal	.431	.156	.075	.166	.069	.215	1.000	.244	.110	.480	.249	.261	.222	.673	-.106	.160	.119	.129	-.078	.321	-.221	.054	.304	.060	-.012	-.011	.085	.155	-.005	.372
Des	.176	.194	.154	.279	.193	.153	.244	.354	.193	.129	.046	.388	.021	.287	-.061	.263	.161	.171	-.069	.143	-.036	.077	.061	.124	.006	-.002	.135	.176	-.024	.277
Det	.098	.180	.149	.219	.157	.084	.110	.193	.235	.049	.042	-.020	-.030	.160	-.040	.178	.139	.177	-.062	.128	-.050	.009	-.012	.173	-.003	-.003	.124	.166	-.028	.206
Hou	.192	.120	.035	.046	.044	.077	.480	.129	.049	.514	.267	.159	.183	.326	-.086	.064	.120	.087	-.011	.190	-.062	.043	.185	-.041	.000	-.004	.042	.053	.000	-.013
Jac	.286	.181	.034	.092	.079	.138	.261	.388	-.020	.159	-.018	.468	-.110	.335	-.126	.360	.249	.235	-.043	.194	-.007	-.014	.093	.142	.028	.006	.125	.195	-.001	.000
Kan	.139	.271	.119	.352	.279	.138	.261	.388	-.020	.159	-.018	.468	-.110	.335	-.126	.360	.249	.235	-.043	.194	-.007	-.014	.093	.142	.028	.006	.125	.195	-.001	.000
Las	.165	.013	-.005	-.089	-.083	.104	.222	.021	-.030	.183	.073	.110	.582	.121	.035	.033	.004	.029	.014	.092	.085	.254	.436	-.070	.005	.000	.020	.037	.016	-.031
Lit	.428	.228	.098	.252	.212	.175	.673	.287	.160	.326	.245	.335	.121	.663	-.090	.176	.163	.205	-.075	.318	-.186	.026	.222	.108	-.012	-.006	.099	.165	-.004	.267
Los	-.104	-.091	-.004	-.105	-.114	-.023	-.106	-.061	-.040	-.086	-.143	-.126	.035	-.090	.408	-.066	-.054	-.101	.048	-.149	.147	.066	-.020	-.180	.011	.009	-.040	-.078	.023	.072
Min	.110	.177	.142	.246	.128	.130	.160	.263	.178	.064	.040	.360	.033	.176	-.066	.309	.142	.179	-.067	.130	-.015	.084	.056	.146	.005	.001	.168	.215	-.024	.280
New	.173	.195	.141	.171	.134	.087	.119	.161	.139	.120	.090	.249	.004	.163	-.054	.142	.220	.184	-.038	.141	-.022	.006	.009	.104	.006	.007	.092	.139	-.008	.077
Phi	.243	.257	.159	.208	.175	.077	.129	.171	.177	.087	.167	.235	.029	.205	-.101	.179	.184	.307	-.045	.236	-.059	-.015	.095	.104	-.008	-.005	.125	.181	-.016	.062
Por	-.060	-.039	-.035	-.082	-.046	-.059	-.078	-.069	-.062	-.011	.003	-.043	.014	-.075	.048	-.067	-.038	-.045	.088	-.061	.065	.037	-.023	-.108	.010	.008	-.034	-.047	.025	-.020
Ral	.354	.230	.101	.141	.176	.076	.321	.143	.128	.190	.239	.194	.092	.318	-.149	.130	.141	.236	-.061	.381	-.146	.329	.089	.121	-.014	-.002	.097	.155	-.002	-.055
Sac	-.156	-.046	.006	-.056	-.066	-.040	-.221	-.036	-.050	-.062	-.078	-.007	.085	-.186	.147	-.015	-.022	-.059	.065	-.146	.329	.089	.000	-.101	.023	.015	.020	-.033	.019	-.043
Sal	.004	-.006	.036	.006	-.064	.070	.054	.077	.009	.043	.003	-.014	.254	.026	.066	.084	.006	-.015	.037	-.034	.089	.279	.181	-.029	.020	.020	.038	.051	.010	.160
Tuc	.309	.064	.025	-.015	-.033	.091	.304	.061	-.012	.185	.184	-.093	.436	.222	-.020	.056	.009	.095	-.023	.198	.000	.181	.631	-.013	-.013	-.003	.021	.067	.009	-.030
Was	.145	.135	.093	.180	.113	.014	.060	.124	.173	-.041	.108	.142	-.070	.108	-.180	.146	.104	.104	-.108	.121	-.101	-.029	-.013	.207	.000	.012	.103	.159	-.031	.000
Cal	-.014	-.005	.003	.000	-.001	.002	-.012	.006	-.003	.000	.000	.028	.005	-.012	.011	.005	.006	-.008	.010	-.014	.023	.020	-.013	.000	.009	.008	.005	.002	.003	.012
Edm	-.008	.000	.005	.003	.002	-.001	-.011	-.002	-.003	-.004	.010	.006	.000	-.006	.009	.001	.007	-.005	.008	-.002	.015	.020	-.003	.012	.008	.010	.006	.009	.004	.016
Mon	.076	.120	.114	.162	.072	.064	.085	.135	.124	.042	.055	.125	.020	.099	-.040	.168	.092	.125	-.034	.097	.020	.038	.021	.103	.005	.006	.158	.190	-.009	.104
Tor	.134	.164	.136	.236	.128	.087	.155	.176	.166	.053	.092	.195	.037	.165	-.078	.215	.139	.181	-.047	.155	-.033	.051	.067	.159	.002	.009	.190	.264	-.008	.214
Van	-.005	-.011	-.014	-.019	-.014	-.015	-.005	-.024	-.028	.000	.013	-.001	.016	-.004	.023	-.024	-.008	-.016	.025	-.002	.019	.010	.009	-.031	.003	.004	-.009	-.008	.016	-.057
Win	.070	.089	.120	.208	.089	.151	.372	.277	.206	-.013	-.032	.000	-.031	.267	.072	.280	.077	.062	-.020	-.055	-.043	.160	-.030	.000	.012	.016	.104	.214	-.057	.228

Table 3. Estimated mean vector $\hat{\mu}_x$ (Equation 5) and covariance matrix $\hat{\Sigma}_x$ (Equation 4) for annual heating degree totals. Each entry of $\hat{\Sigma}_x$ shown is expressed as a fraction of the largest estimated covariance 603,778.8 and rounded to three decimal places, so the largest entry on the table is 1.000.

μ	Atl	Bal	Bos	Chi	Cin	Col	Dal	Des	Det	Hou	Jac	Kan	Las	Lit	Los	Min	New	Phi	Por	Ral	Sac	Sal	Tuc	Was	Cal	Edm	Mon	Tor	Van	Win
2512	4139	4755	5145	4378	5190	2115	5451	5301	1162	1269	4726	1786	2903	958	6533	4124	4035	3419	3018	2201	4680	1301	3879	6798	7954	6868	5809	3943	7845	
Σ	Atl	Bal	Bos	Chi	Cin	Col	Dal	Des	Det	Hou	Jac	Kan	Las	Lit	Los	Min	New	Phi	Por	Ral	Sac	Sal	Tuc	Was	Cal	Edm	Mon	Tor	Van	Win
Atl	.200	.139	.090	.144	.166	.058	.119	.136	.136	.113	.125	.095	-.008	.136	-.046	.107	.133	.162	-.026	.168	-.026	-.001	.009	.126	-.012	-.030	.067	.116	-.048	.152
Bal	.139	.181	.134	.157	.166	.028	.094	.140	.180	.087	.092	.151	-.030	.118	-.056	.144	.168	.185	-.031	.148	-.038	-.034	-.022	.134	-.025	-.012	.111	.148	-.057	.371
Bos	.090	.134	.152	.134	.136	.031	.063	.127	.168	.050	.051	.136	-.017	.086	-.035	.140	.143	.155	-.010	.108	-.022	-.019	-.015	.088	.005	.009	.138	.153	-.016	.463
Chi	.144	.157	.134	.288	.245	.093	.139	.291	.261	.110	.098	.272	-.026	.153	-.040	.296	.180	.194	.037	.151	-.045	-.004	-.014	.088	.114	.075	.136	.199	.017	.558
Cin	.166	.166	.136	.245	.233	.067	.135	.213	.229	.122	.107	.198	-.047	.162	-.048	.211	.170	.190	.010	.178	-.041	-.028	-.031	.097	.052	.024	.113	.179	-.018	.287
Col	.058	.028	.031	.093	.067	.153	.095	.133	.055	.057	.024	.135	.063	.090	.018	.095	.008	.040	.033	.053	.014	.099	.066	.026	.052	.013	.008	.041	.025	.045
Dal	.119	.094	.063	.139	.135	.095	.147	.145	.111	.114	.078	.164	.009	.128	-.033	.100	.076	.111	.001	.117	-.020	.018	.022	.093	-.004	-.050	.048	.088	-.023	-.045
Des	.136	.140	.127	.291	.213	.133	.145	.349	.241	.095	.076	.308	.005	.155	-.008	.334	.146	.171	.063	.137	-.026	.039	.014	.055	.217	.174	.116	.185	.053	.607
Det	.136	.180	.168	.261	.229	.055	.111	.241	.249	.084	.087	.261	-.037	.140	-.052	.258	.193	.199	.004	.156	-.055	-.032	-.029	.125	.067	.041	.174	.213	-.012	.553
Hou	.113	.087	.050	.110	.122	.057	.114	.095	.084	.106	.080	.107	.081	.084	-.043	.056	.092	.106	-.022	.114	-.025	-.019	.009	.077	-.049	-.073	.038	.073	-.041	-.006
Jac	.125	.092	.051	.098	.107	.024	.078	.076	.087	.080	.107	.081	-.013	.084	-.043	.056	.092	.106	-.022	.114	-.025	-.019	.007	.098	-.041	-.059	.038	.079	-.043	.008
Kan	.095	.151	.136	.272	.198	.135	.164	.308	.261	.099	.081	.290	.007	.145	.026	.217	.168	.157	.100	.149	.010	.038	.034	.125	.063	-.012	.074	.121	.031	.000
Las	-.008	-.030	-.017	-.026	-.047	.063	.009	.005	-.037	-.004	-.013	.007	.079	-.001	.034	-.015	-.052	-.030	.017	-.017	.042	.089	.064	-.017	.021	.026	-.019	-.019	.026	-.039
Lit	.136	.118	.086	.153	.162	.090	.128	.155	.140	.106	.084	.145	-.001	.145	-.037	.131	.110	.134	-.005	.132	-.028	.014	.008	.121	.008	-.022	.078	.111	-.024	.197
Los	-.046	-.056	-.035	-.040	-.048	.018	-.033	-.008	-.052	-.039	-.043	.026	.034	-.037	.077	-.008	-.055	-.057	.042	-.057	.048	.066	.038	-.086	.080	.073	-.060	-.058	.052	-.030
Min	.107	.144	.140	.296	.211	.095	.100	.334	.258	.060	.056	.217	-.015	.131	-.008	.401	.158	.177	.056	.117	-.036	.004	-.017	.033	.256	.222	.163	.220	.067	.787
New	.133	.168	.143	.180	.170	.008	.076	.146	.193	.075	.092	.168	-.052	.110	-.055	.158	.183	.177	.017	.143	-.048	-.045	-.045	.108	-.019	-.007	.121	.162	-.046	.361
Phi	.162	.185	.155	.194	.190	.040	.111	.171	.199	.099	.106	.157	-.030	.134	-.057	.177	.177	.225	-.027	.173	-.036	-.034	-.016	.140	-.004	.003	.136	.190	-.043	.378
Por	-.026	-.031	-.010	.037	.010	.033	.001	.063	.004	-.015	-.022	.100	.017	-.005	.042	.056	-.017	-.027	.095	-.028	.026	.070	.009	-.066	.122	.087	-.017	-.009	.088	.022
Ral	.168	.148	.108	.151	.178	.053	.117	.137	.156	.105	.114	.149	-.017	.132	-.057	.117	.143	.173	-.028	.180	-.036	-.029	-.004	.140	-.019	-.043	.095	.134	-.048	.129
Sac	-.026	-.038	-.022	-.045	-.041	.014	-.020	-.026	-.055	-.017	-.025	.010	.042	-.028	.048	-.036	-.048	-.036	.026	-.036	.077	.067	.039	-.075	.019	.043	-.046	-.038	.030	-.013
Sal	-.001	-.034	-.019	-.004	-.028	.099	.018	.039	-.032	.004	-.019	.038	.089	.014	.066	.004	-.045	-.034	.070	-.029	.067	.216	.080	-.054	.017	-.003	-.045	-.032	.059	.158
Tuc	.009	-.022	-.015	-.014	-.031	.066	.022	.014	-.029	.009	-.007	.034	.064	.008	.038	-.017	-.045	-.016	.009	-.004	.039	.080	.079	-.034	.000	.011	-.029	.020	.012	-.050
Was	.126	.134	.088	.088	.097	.026	.093	.055	.125	.077	.098	.125	-.017	.121	-.086	.033	.108	.140	-.066	.140	-.075	-.054	-.034	.132	-.133	-.152	.087	.095	-.084	.000
Cal	-.012	-.025	.005	.114	.052	.052	-.004	.217	.067	-.049	-.041	.063	.021	.008	.080	.256	-.019	-.004	.122	-.019	.019	.017	.000	-.133	.631	.608	.037	.068	.181	.274
Edm	-.030	-.012	.009	.075	.024	.013	-.050	.174	.041	-.073	-.059	-.012	.026	-.022	.073	.222	-.007	.003	.087	-.043	.043	-.003	.011	-.152	.608	.714	.016	.055	.149	.552
Mon	.067	.111	.138	.136	.113	.008	.048	.116	.174	.038	.038	.074	-.019	.078	-.060	.163	.121	.136	-.017	.095	-.046	-.045	-.029	.087	.037	.016	.228	.198	.000	.553
Tor	.116	.148	.153	.199	.179	.041	.088	.185	.213	.073	.079	.121	-.019	.111	-.058	.220	.162	.190	-.009	.134	-.038	-.032	-.020	.095	.068	.055	.198	.239	.005	.573
Van	-.048	-.057	-.016	.017	-.018	.025	-.023	.053	-.012	-.041	-.043	.031	.026	-.024	.052	.067	-.046	-.043	.088	-.048	.030	.059	.012	-.084	.181	.149	.000	.005	.113	.086
Win	.152	.371	.463	.558	.287	.045	-.045	.607	.553	-.006	.108	.000	-.039	.197	-.030	.787	.361	.378	.022	.129	-.013	.158	-.050	.000	.274	.552	.553	.573	.086	1.000

Figure 6. Empirical density estimates of non-zero losses using the strike value method described in section 4.2. The dependence distribution was built recognizing dependence among the four cities, while the independence distribution simulates independent losses for the four cities. Vertical dashed and solid lines show the 99th percentiles for the two cases



simulations of P . Specifically, the formulation of the multivariate normal distribution of P is:

- $n_i = 100$ for $i = 3, 13, 17$, and 23 , and 0 otherwise;
- $E_3 = 865$; $E_{13} = 3723$; $E_{17} = 1148$; $E_{23} = 3442$; $E_i = 0$ for all i not equal to $3, 13, 17, 23$
- The vector $c = \{c_i, i = 1, \dots, 30\}$ has non-zero entries only for $c_3 = -20 \cdot n_3 \cdot E_3 = -1,730,000$; $c_{13} = -7,446,000$; $c_{17} = -2,296,000$; $c_{23} = -6,884,000$.
- The matrix B has non-zero entries only for $B_{3,3} = B_{13,13} = B_{17,17} = B_{23,23} = 20 \cdot 100 = 2000$.
- $\hat{\mu}_X$ and $\hat{\Sigma}_X$ are as shown in Table 2.

Thus, we have fully specified the multivariate Gaussian in Equation 6. We use this distribution to simulate $J = 100,000$ realizations, and for each realization we replaced negative elements with zero, producing what we termed R' in Section 4.2. Total losses were the sum of the payments for each of the 100,000 realizations. In these simulations, fully 36.1% of payments were identically equal to zero, meaning that none of the four cities exceeded the respective strike values. The density of the remaining 63.9% of non-zero payments is shown in Figure 6 as the solid line. For comparison, we repeated the methodology ignoring spatial dependence (i.e., we set all off-diagonal terms of $\hat{\Sigma}_X$ to zero),

and termed this the *independence* case. Here 24.2% of payments were identically zero, and the dashed line shows the density of the remaining 75.8% of non-zero payments. As expected, the variability and tail thickness of the dependent solid line is larger than for the dashed independent line. The 99th percentiles are also shown, and as expected the dependent case shows a much larger high percentile. Risk measures are shown in Table 4.

6. Discussion

In this paper we fit models for cumulative heating and cooling degree days for 30 North American cities. Through the use of a multivariate normal distribu-

Table 4. Estimated VaR and CTE values from Equations 7 and 8 for a few high values of α , using both the dependence model and the independence model which ignores spatial dependence. Numbers are computed empirically as quantiles and means for values above the quantile.

α	Dependence		Independence	
	$\widehat{Var}$	$\widehat{CTE}$	$\widehat{Var}$	$\widehat{CTE}$
0.90	958,804	1,406,799	822,520	1,229,679
0.95	1,294,455	1,702,666	1,036,317	1,533,535
0.99	1,967,233	2,293,387	1,460,644	2,183,935

tion, we incorporated spatial dependence among the 30 cities. Through affine transformations of cCDDs and cHDDs, we constructed normal distributions for payments and demonstrated how this distribution could be used to estimate some common risk measures. For the far more commonly traded derivatives based on a strike value, we demonstrated how a computational approach can simulate large numbers of losses and associated risk measures.

We demonstrated that historical data on degree day indices shows evidence of climate change consistent with scientific papers documenting the changes, and found statistically significant positive trends for cCDDs over the summer and statistically significant negative trends for cHDDs over the winter. Statistical models which use data going back several decades should explicitly recognize these trends and de-trend all data to the common time period of interest. Otherwise, models run the risk of slightly underestimating cCDDs and overestimating cHDDs. As mentioned earlier in the paper, we briefly investigated the need to adjust historical data based on El Niño and other climate oscillations but did not identify statistically significant relationships. We leave it as an open question to what degree incorporating external covariates based on climate indices would improve modeling.

A financial interest for the insurance industry comes about by recognizing that weather is minimally correlated with financial and insurance markets, and as such, the financial performance of weather derivatives is largely independent from the financial performance of other assets an insurer holds. This allows diversification and risk control. A business interest is simply that weather products can be new insurance products. Some property and casualty insurers are already in the business of selling insurance products to cover losses triggered from environmental or weather causes (crop, flood, and catastrophe insurance, to name only three). Sophisticated ratemaking procedures can incorporate past weather data and/or seasonal weather forecasting.

A natural extension of this expertise might be found in weather products.

The most notable limitation of the method in this paper is its reliance on sufficiently long time periods to allow for normality in the degree day sums. We do not recommend this approach for modeling cumulative degree days over time periods too short to allow the normal approximation to hold. Specifically, we do not recommend this approach for a month or even a few months. To model degree day totals for shorter time periods, a fundamentally different approach must be used. An obvious choice would be to first model temperature at the daily level, and then use these models to construct distributions cCDDs and cHDDs. Closed-form solutions would likely be hard to come by given the model complexity required for daily temperatures, so a simulation-based approach may be needed. Such approaches are beyond the scope of this paper, but we mention them not only to highlight the limitation of the approach presented here for shorter time scales, but also to suggest future avenues of promising research.

Acknowledgments

The author wishes to thank the Research Grants Task Force of the Casualty Actuarial Society and the Committee on Knowledge Extension Research of the Society of Actuaries for their generous financial support of this project. The authors also wish to thank the two anonymous referees and the editorial staff of *Variance* for their helpful suggestions which improved the paper.

References

- Alexandridis, A., and A. Zaprani, *Weather Derivatives: Modeling and Pricing Weather-Related Risk*, New York: Springer, 2013.
- Benth, F. and J. Saltyte Benth, *Modeling and Pricing in Financial Markets for Weather Derivatives*. Hackensack, NJ: World Scientific, 2013.
- Campbell, S.D. and F.X. Diebold, "Weather Forecasting for Weather Derivatives," *Journal of the American Statistical Association* 100 (no. 469), 2005, pp. 6–16.

- Dischel, R., *Climate Risk and the Weather Market: Financial Risk Management With Weather Hedges*, London: Risk Publications, 2002.
- Dutton, J.A., "Opportunities and Priorities in a New Era for Weather and Climate Services," *Bulletin of the American Meteorological Society* 83, 2002, pp. 1303–1311.
- Erhardt, R., "Mid-Twenty-First-Century Projected Trends in North American Heating and Cooling Degree Days," *Environmetrics* 26(2), 2014, pp. 133–144.
- Erhardt, R., and R. Smith, "Weather Derivative Risk Measures for Extreme Events," *North American Actuarial Journal* 18, 2014, pp. 379–393.
- Geman, H., *Insurance and Weather Derivatives: From Exotic Options to Exotic Underlyings*, London: Risk Publications, 1999.
- Greene, W.H., *Econometric Analysis*, 5th ed., New York: Pearson Education, 2003.
- IPCC (Intergovernmental Panel on Climate Change), "The Fifth Assessment Report (AR5) of the United Nations Intergovernmental Panel on Climate Change (IPCC)," Climate Change 2013: The Physical Science Basis, IPCC WGI AR5. Tech. rep.
- Jewson, S., A. Brix, and C. Ziehmann, *Weather Derivative Valuation. The Meteorological, Statistical, Financial and Mathematical Foundations*, New York: Cambridge University Press, 2005.
- Kaas, R., M. Goovaerts, J. Dhaene, and M. Denuit, *Modern Actuarial Risk Theory*, 2nd ed., New York: Springer, 2008.
- Kunreuther, H., and E. Michel-Kerjan, *At War with the Weather: Managing Large-Scale Risks in a New Era of Catastrophes*, Cambridge, Mass.: MIT Press, 2009.
- Mills, E., "Insurance in a Climate of Change," *Science* 309, 2005, pp. 1040–1044.
- Richards, T., M. Manfredo, and D. Sanders, "Pricing Weather Derivatives," *American Journal of Agricultural Economics* 86, 2004, pp. 1005–1017.
- Saltyte Benth, J., F. Benth, and P. Jalinskas, "A Spatial-Temporal Model for Temperature with Seasonal Variance," *Journal of Applied Statistics* 34, 2007, pp. 823–841.
- Saltyte Benth, J., and L. Saltyte, "Spatial-Temporal Model for Wind Speed in Lithuania," *Journal of Applied Statistics* 38, 2011, pp. 1151–1168.