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 Zapranis 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

\[
\begin{align*}
HDD(T) &= \max(65 - T, 0) \\
CDD(T) &= \max(T - 65, 0),
\end{align*}
\]

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

\[
P = 10,000 \cdot \max(\text{cCDD} - 360, 0).
\]

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

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 Zapranis 2013), investing in weather derivatives is one approach that an insurance company can take.
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 $T_{\text{max}}$ and minimum daily temperature $T_{\text{min}}$ 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 $d = 1, \ldots, D = 30$ 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

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

For each day, cooling degree days are computed as

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

and heating degree days are computed as

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

Cooling and heating degree days are most commonly used to construct cumulative indices over some time period of interest. Let $\{t \in T\}$ be the set of all days within the time period of interest $T$. Then a cumulative CDD index is simply

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

where $T$ 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 ($cCDD(T)$ or $cHDD(T)$) 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

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

The VaR is easily obtained since the loss $L$ is often a continuous random variable for larger values of $\alpha$, and therefore the distribution function $F_{\alpha}(l) = P(L \leq l)$ is strictly increasing and has inverse distribution func-
Incorporating Spatial Dependence and Climate Change Trends for Measuring Long-Term Temperature Derivative Risk

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.

Annual Heating Degree Days, 1946–2014

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 derivatives 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.

\[
CTE_{\alpha}(L) = \mathbb{E}(L|L > Q_{\alpha}),
\]

which is the expected loss conditional upon exceeding the VaR.

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

\[
CTE_{\alpha}(L) = \mathbb{E}(L|L > Q_{\alpha}),
\]

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 derivatives 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.

\[
CTE_{\alpha}(L) = \mathbb{E}(L|L > Q_{\alpha}),
\]

which is the expected loss conditional upon exceeding the VaR.
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 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/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
where \( \mu_X \) is the mean vector and \( \Sigma_X \) 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 \( \Sigma_X \) is easily invertible from a computational cost standpoint. The covariance is estimated with the usual unbiased estimator

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

The mean of the multivariate normal is estimated using

\[
\hat{\mu}_X = \frac{1}{n} \sum_i x_i. \tag{5}
\]

### 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 \( \mu_X \) and \( D \) covariance matrix \( \Sigma_X \), 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

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

<table>
<thead>
<tr>
<th>Index</th>
<th>City</th>
<th>Trend</th>
<th>( \bar{SE} ) (Trend)</th>
<th>tStat</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>cCDDs</td>
<td>Boston</td>
<td>0.771</td>
<td>0.798</td>
<td>0.966</td>
<td>0.337</td>
</tr>
<tr>
<td></td>
<td>New York</td>
<td>2.399</td>
<td>1.283</td>
<td>1.870</td>
<td>0.067</td>
</tr>
<tr>
<td></td>
<td>Las Vegas</td>
<td>12.878</td>
<td>1.624</td>
<td>7.931</td>
<td>4.2 ( \times 10^{-11} )</td>
</tr>
<tr>
<td></td>
<td>Tucson</td>
<td>9.710</td>
<td>1.720</td>
<td>5.644</td>
<td>4.2 ( \times 10^{-7} )</td>
</tr>
<tr>
<td>cHDDs</td>
<td>Boston</td>
<td>−1.927</td>
<td>1.830</td>
<td>−1.053</td>
<td>0.296</td>
</tr>
<tr>
<td></td>
<td>New York</td>
<td>−5.519</td>
<td>2.659</td>
<td>−2.075</td>
<td>0.043</td>
</tr>
<tr>
<td></td>
<td>Las Vegas</td>
<td>−11.932</td>
<td>1.459</td>
<td>−8.177</td>
<td>1.7 ( \times 10^{-11} )</td>
</tr>
<tr>
<td></td>
<td>Tucson</td>
<td>−6.334</td>
<td>1.422</td>
<td>−4.456</td>
<td>3.4 ( \times 10^{-6} )</td>
</tr>
</tbody>
</table>
derivative that pays $20 per contract per degree day, with the sign of the payment determined by an entry level. Mathematically, this payment is

\[ P_i = 20n_i (X_i - E_i) = -20n_i E_i + 20n_i X_i, \]

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

\[ c^T = (-20n_1, \ldots, -20n_{30}E_{30}), \]

and square matrix \( B \) as

\[ B = \begin{pmatrix}
20n_1 & 0 & \cdots & 0 \\
0 & 20n_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 20n_{30}
\end{pmatrix}. \]

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

\[ P \sim \mathcal{N} \left( c + B\mu_X, B\Sigma_X B^T \right). \]  

The quantity which is ultimately of interest is \( L = \Sigma P \), the total aggregate scalar loss. If we define vector \( b^T = (1, \ldots, 1) \) we can write \( L = b^TP \), which is again an affine transformation and therefore the distribution of \( L \) is normal as

\[ L \sim \mathcal{N} \left( b^T(c + B\mu_X), b^T B\Sigma_X B^T b \right) = \mathcal{N} \left( \mu_L, \sigma^2_L \right), \]

where \( \mu_L \) is the univariate mean of the loss, and \( \sigma^2_L \) 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 \( \mu_X \) and \( \Sigma_X \) 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

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

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

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

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

\[ \hat{\text{CTE}}_\alpha(L) = \hat{\mu}_L + \frac{\hat{\sigma}_L}{1 - \Phi \left( \frac{Q_\alpha - \hat{\mu}_L}{\hat{\sigma}_L} \right)} \left( \frac{Q_\alpha - \hat{\mu}_L}{\hat{\sigma}_L} \right), \]

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 approached. 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 \( P_i \) 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 \( L = \Sigma P_i \). 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 \( i = 1, \ldots, 30 \) define \( R'_i = \max (P'_i, 0) \). This is simply replacing negative elements of \( P' \) with zero.
3. Compute \( L' = \Sigma R'_i \)
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

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

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 \( J \to \infty \)). The conditional tail expectation can be estimated as

\[
\overline{CTE}_\alpha(L) = \frac{1}{|\tau|} \sum_{j} L_j \cdot I_{\{L_j > \overline{VaR}(L; \alpha)\}}
\]

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

5. Results

Estimated mean vectors \( \hat{\mu}_c \) and covariance matrixes \( \hat{\Sigma}_c \) 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 \( \hat{\mu}_c \) 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 $\hat{\mu}$ (from Equation 5) and covariance matrix $\hat{\Sigma}$ (Equation 4) for annual cooling degree totals. Each entry of $\hat{\Sigma}$, shown is expressed as a fraction of

<table>
<thead>
<tr>
<th>City</th>
<th>Mean</th>
<th>Covariance</th>
<th>Covariance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bos</td>
<td>0.60</td>
<td>0.21</td>
<td>0.19</td>
</tr>
<tr>
<td>Chi</td>
<td>0.71</td>
<td>0.25</td>
<td>0.09</td>
</tr>
<tr>
<td>Cin</td>
<td>0.78</td>
<td>0.23</td>
<td>0.08</td>
</tr>
<tr>
<td>Dal</td>
<td>0.85</td>
<td>0.17</td>
<td>0.05</td>
</tr>
<tr>
<td>Hou</td>
<td>0.90</td>
<td>0.13</td>
<td>0.03</td>
</tr>
<tr>
<td>Jac</td>
<td>0.93</td>
<td>0.11</td>
<td>0.02</td>
</tr>
<tr>
<td>Las</td>
<td>0.95</td>
<td>0.10</td>
<td>0.01</td>
</tr>
<tr>
<td>Lit</td>
<td>0.96</td>
<td>0.09</td>
<td>0.00</td>
</tr>
<tr>
<td>Min</td>
<td>0.97</td>
<td>0.08</td>
<td>0.00</td>
</tr>
<tr>
<td>Phi</td>
<td>0.98</td>
<td>0.07</td>
<td>0.00</td>
</tr>
<tr>
<td>Por</td>
<td>0.99</td>
<td>0.06</td>
<td>0.00</td>
</tr>
<tr>
<td>Ral</td>
<td>1.00</td>
<td>0.05</td>
<td>0.00</td>
</tr>
<tr>
<td>Sac</td>
<td>1.00</td>
<td>0.04</td>
<td>0.00</td>
</tr>
<tr>
<td>Sal</td>
<td>1.00</td>
<td>0.03</td>
<td>0.00</td>
</tr>
<tr>
<td>Tuc</td>
<td>1.00</td>
<td>0.02</td>
<td>0.00</td>
</tr>
<tr>
<td>Was</td>
<td>1.00</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>Cal</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Edm</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Win</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Variance Advancing the Science of Risk
### Table 3: Estimated mean vector $\hat{\mu}_X$ (Equation 5) and covariance matrix $\hat{\Sigma}$ (Equation 4) for annual heating degree totals. Each entry of $\hat{\Sigma}$ 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.

| City | Atl | Bos | Bal | Bos | Chi | Cin | Col | Dal | Des | Det | Hou | Jac | Kan | Las | Lit | Los | Min | New | Phi | Ral | Sac | Sal | Tuc | Was | Cal | Edm | Mon | Tor | Van | Win |
|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| $\hat{\mu}_X$ | 0.20 | 0.19 | 0.14 | 0.16 | 0.16 | 0.13 | 0.13 | 0.11 | 0.11 | 0.14 | 0.11 | 0.13 | 0.14 | 0.15 | 0.14 | 0.13 | 0.12 | 0.12 | 0.11 | 0.12 | 0.11 | 0.12 | 0.12 | 0.12 | 0.11 | 0.12 | 0.12 | 0.12 |
| $\hat{\Sigma}$ | 0.008 | 0.060 | 0.087 | 0.083 | 0.082 | 0.076 | 0.076 | 0.074 | 0.073 | 0.072 | 0.072 | 0.072 | 0.072 | 0.072 | 0.072 | 0.072 | 0.072 | 0.072 | 0.072 | 0.072 | 0.072 | 0.072 | 0.072 | 0.072 | 0.072 | 0.072 | 0.072 |

... (Continues with the rest of the table entries)...
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>
<thead>
<tr>
<th>α</th>
<th>Dependence</th>
<th>Independence</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.90</td>
<td>958,804</td>
<td>822,520</td>
</tr>
<tr>
<td>0.95</td>
<td>1,294,455</td>
<td>1,036,317</td>
</tr>
<tr>
<td>0.99</td>
<td>1,967,233</td>
<td>1,460,644</td>
</tr>
</tbody>
</table>
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


