# Precipitation Modeling and Contract Valuation: *A Frontier in Weather Derivatives*

MELANIE CAO, ANLONG LI, AND JASON WEI


MELANIE CAO
is an assistant professor in the Schulich School of Business at York University, Toronto, Canada.
mcao@schulich.yorku.ca

ANLONG LI
is a director of research at XL Weather & Energy Inc. in Stamford, CT.
ali@xlwe.com

JASON WEI
is an associate professor in the Joseph L. Rotman School of Management at the University of Toronto, Toronto, Canada.
wei@rotman.utoronto.ca


Weather derivatives debuted in the mid '90s when deregulation of the energy and utility industries started in the U.S. Growing competition and uncertainty in demand prompted energy and utility companies to seek for effective hedging tools to stabilize earnings. Price hedging alone was no longer adequate; a combined framework of price and volumetric risk management was called for. Since weather conditions are among the key factors determining the demand of energy, managing volumetric risk is tantamount to managing weather risk. Specifically, for the electricity and natural gas sectors, temperature is the key factor affecting the demand. According to the Weather Risk Management Association (WRMA), the total notional value of OTC weather contracts was around $4 billion by 2001; 80% of the contracts or 90% of the notional value were for temperature derivatives. Meantime, trading of temperature derivatives on the Chicago Mercantile Exchange (CME) has also been on the rise. Currently, the CME lists temperature futures for 15 U.S. cities and 5 European cities. The total number of contracts traded in 2003 was 14,234.

The second important category of weather derivatives is arguably precipitation contracts. In contrast to temperature derivatives, the development of the precipitation derivatives market is still in its infancy. According to the WRMA, deals based on rain and snow made up around 3% of the global weather market in the winter of 2000. The current proportion is not much higher judging by the sparsity of new contracts coming out.

The slow growth by no means reflects a lack of interest or demand. In fact, end-users such as farmers and hydroelectric power producers are very keen on precipitation contracts. The difficulty is in finding the counterparty who is willing or able to provide a reasonable quote. The hesitation on the part of financial institutions is in turn due to the difficulty and challenge in properly modeling precipitation. Modeling precipitation and valuing related derivative contracts are indeed a frontier in the field of weather derivatives.

This article makes an earnest attempt to fill this important gap in the literature and the industry. We propose, calibrate, and compare three precipitation models: a gamma distribution, a mixture of exponentials, and kernel density. Based on the data for Chicago Midway Airport (1950-2003), we find that the latter two models dominate the first model in terms of fit. In the remainder of the article, we first discuss the application of precipitation contracts by describing several deals; we then delineate the modeling, calibration, and related issues; and finally, we summarize and conclude the article.

## APPLICATION OF PRECIPITATION CONTRACTS

Precipitation, be it rain or snow, exerts a significant impact on the revenue of many businesses. Farming is an obvious example. Drought or flood can both adversely affect the crop yield. Operators of certain outdoor recreational services (e.g., golf and skiing) have much to dread about excess precipitation or the lack of it. Hydroelectric generators are very keen on the accumulative precipitation over any given time period. Lack of precipitation means a low water level in reservoirs, which in turn means a shortage of power supply. The reduced power output not only leads to a loss of revenue, sometimes it also means purchasing power at unfavorable prices from other generators in order to make up the supply shortage. Below, we describe three deals, two of which involve hydroelectric power generators.

***Case 1: Southern Hydro Partnership versus Credit Lyonnais Rouse Derivatives*** (Source: ARTEmis, http://www.artemis.bm/index.htm). Southern Hydro Partners (SHP) is a hydroelectric power generator in South East Australia, with most of its facilities located in Victoria and New South Wales. With rainfall levels being significantly below the historical average for several years, the company decided to enter into a precipitation contract with Credit Lyonnais Rouse Derivatives (subsequently becoming Calyon in May 2004) in 2003. SHP's primary goal was to stabilize cashflows and revenue. The precipitation contract was for a three-year period. To save hedging cost, the contract was structured as a collar whereby SHP would receive payments from Calyon should the rainfall be lower than a specified threshold level, and pay Calyon should the rainfall be above a certain level.

***Case 2: Sacramento Municipal Utility District versus Aquila*** (Source: *Environmental Finance*, October 2001). Sacramento Municipal Utility District (SMUD) is America's sixth largest community-owned electric utility in terms of customers served. The utility generates half of its electricity and buys the rest. The generated portion is primarily from hydroelectric and cogeneration power plants. In September 2000, SMUD entered into a five-year precipitation contract with Aquila (an energy trading firm based in Kansas City) to protect against low rainfall levels. Similar to the deal in Case 1, the contract was structured as a collar whereby Aquila would pay SMUD annually up to $20 million when the water flow through the hydro plants is below a certain amount, while SMUD would pay Aquila $20 million in years when precipitation is abundant. To further reduce the cost of hedging, the payments to Aquila was capped at $50 million.

***Case 3: Golf Course Operator versus Société Générale SA*** (Source: Bloomberg News and Commentary, http://www.bloomberg.com/news/index.html). Mr. Dieter Worms operates a golf course, Gut Apeldor Gold Club, in Hennstedt which is about 100 kilometers north of Hamburg in Germany. The business suffered in 2001 when the weather was mainly wet. In the subsequent year, Mr. Worms entered into a derivative contract with Société Générale SA, the third largest bank in France. The deal covered the period from May to September, within which Mr. Worms would receive compensation should the total number of rainy days go beyond 50. Specifically, once the number of days with rainfall of more than a millimeter passed 50, Société Générale SA would pay Mr. Worms for every wet day.

Finally, the Swedish Meteorological and Hydrological Institute and Energy-Koch Trading have teamed up to launch a Nordic Precipitation Index. This index is based on 17 stations, 9 in Norway, and 8 in Sweden. Concerned companies in the Nordic region can enter into contracts based on this index for their risk management.[1]

## PRECIPITATION MODELING AND CONTRACT VALUATIONS

### Modeling Daily Precipitations

As pointed out by Dischel [2000], in contrast to the modeling of temperature, modeling precipitation presents several challenges.[2] The first challenge is the accurate measurement of precipitation. Most techniques involve physically collecting raindrops and measure the precipitation amount accordingly. Factors such as local wind can affect the collection accuracy. Secondly, spatial correlation is an elusive measure. Unlike temperature which is highly correlated across nearby regions, rainfall can be very localized. A tremendous basis risk is present for any precipitation contracts when the measurement site (usually government operated) is far away from the site in question. The third challenge is in selecting a proper distribution to describe the precipitation data. Again, unlike temperature which can adequately be described by a simple distribution such as Gaussian, the statistical property of precipitation is far more complex and a more sophisticated distribution is called for.

Notwithstanding, some authors have attempted to model precipitation statistically. For instance, Sansó and Guenni [1999] proposed a model for tropical rainfall at a single location for a fixed period (e.g., 10 days). The amount of rainfall is modeled as a transformed normal variable with dynamic parameters, while the event of rain or dry is modeled by truncating the same normal variable—negative draws from the normal distribution correspond to dry days. Sansó and Guenni [2000] later extended their model to a multi-site setting. Wilks [1998], on the other hand, proposed a multi-site model for daily precipitation using a combination of two-state Markov process (for the rainfall occurrence) and a mixed exponential distribution (for the precipitation amount). He found that the mixture of exponential distributions offered a much better fit than the commonly used gamma distribution (e.g., Katz [1977]; Richardson and Wright [1984]; and Wilks [1989]).

In the following, we propose a single-site model in a spirit similar to Wilks [1998]. For comparison purposes, we examine three approaches of modeling the precipitation: a gamma distribution, a mixture of exponentials, and kernel density.

Let $X_t$ be a binary variable that takes a value of 1 if it rains on day $t$ and 0 otherwise. $X_t$ is an $n$-th order Markov chain if $X_t$ is independent of $X_{t-k}$ for all $k > n$. For simplicity, we will consider only first order Markov chains, i.e., $n = 1$. Let $p$ be the probability that day $t$ is wet. Then for the first order Markov chain, we have

$$p_t = p_{t-1}q_{11} + (1 - p_{t-1})q_{01} \tag{1}$$

where $q_{11}$ and $q_{01}$ are the one-step transition probabilities.

Conditional on $X_t = 1$ (i.e., a wet day), the amount of rainfall $Y_t$ can be modeled as a random variable that follows a particular distribution.[3] When $Y_t$ follows a gamma distribution, the probability density function is

$$f_{gamma}(y) = \frac{y^{a-1}e^{-y/b}}{b^a\Gamma(a)} \tag{2}$$

where $a$ and $b$ are distribution parameters which can be estimated using the maximum likelihood method. When $Y_t$ follows a mixture of two exponential distributions, the probability density function and cumulative density function are

$$f_{mixture}(y) = \frac{\alpha}{\beta_1}e^{-y/\beta_1} + \frac{1-\alpha}{\beta_2}e^{-y/\beta_2} \tag{3}$$

and

$$F_{mixture}(y) = \alpha e^{-y/\beta_1} + (1-\alpha)e^{-y/\beta_2} \tag{4}$$

respectively. The parameters $\alpha$, $\beta_1$, and $\beta_2$ can also be estimated using the maximum likelihood method.

Seasonalities can easily be built into the framework. For instance, we can allow the one-step transition probability $q_{11}$ to vary with the time of the year as follows:

$$q_{11} = c_0 + \sum_{i=1}^{m}\left[c_{1i}\sin\left(\frac{2\pi i}{365}t\right) + c_{2i}\cos\left(\frac{2\pi i}{365}t\right)\right] \tag{5}$$

where $t$ is time, $m$ is a small integer to be set by the user, and $c_0$, $c_{1i}$, and $c_{2i}$ ( $i = 1, 2, \ldots, m$) are parameters to be estimated.

The seasonal feature for the conditional daily precipitation amounts can be handled in a similar manner. Specifically, the mean conditional precipitation for day $t$ can be modeled as

$$\overline{Y}_t = d_0 + \sum_{i=1}^{m}[d_{1i}\sin(\frac{2\pi i}{365}t) + d_{2i}\cos(\frac{2\pi i}{365}t)] \tag{6}$$

where $m$ serves the same purpose as in (5), and $d_0$, $d_{1i}$, and $d_{2i}$ ( $i = 1, 2, \ldots, m$) are parameters to be estimated.

## Model Calibration, Estimation, and Comparisons

If the transition probabilities are constant, then the maximum likelihood estimate for $q_{01}$ ($q_{11}$) is the ratio of (a) the total number of raining days where the previous day is dry (wet) and (b) the number of dry (wet) days. To estimate seasonal transition probabilities as defined in Equation (5), we estimate the parameters using the following linear regression model for those days where the previous day rains:

$$I_t(1) = \hat{c}_0 + \sum_{i=1}^{m}[\hat{c}_{1i}\sin(\frac{2\pi i}{365}t) + \hat{c}_{2i}\cos(\frac{2\pi i}{365}t)] + \varepsilon_t$$

where $I_t(1) = 1$ if day $t$ rains and 0 otherwise. The term $\varepsilon_t$ is the regression error. The conditional probability $q_{01}$ can be estimated in the same way. In general, the larger the number $m$, the richer the seasonality pattern the model can capture. However, a larger $m$ will also reduce the estimation accuracy. Based on our experiences, we usually set $m = 5$. Exhibit 1 shows the estimated transition probability $q_{01}$ and $q_{11}$ for Chicago Midway Airport using data from 1950 to 2003.

Similarly, the mean conditional precipitation amount can also be estimated using linear regression from Equation (6). The result for Chicago Midway is shown in Exhibit 2.

Next, we need to estimate the conditional distribution of the precipitation amount. To dampen the seasonality effect, we normalize the precipitation amount $Y_t$ by the seasonal mean $\overline{Y}_t$ in Equation (6). The parameter estimates for the mixed exponential distribution are presented in Exhibit 3. (Note: precipitation is measured in inches throughout the article.)

The parameter estimates for the gamma distribution are shown in Exhibit 4.

To see how good the fits are, we make the following observation: Let $F$ be the true probability distribution of $Y_t/\overline{Y}_t$ and $M$ be the standard normal distribution; then $\Phi^{-1}(F(Y_t/\overline{Y}_t))$ will be a standard normal random variable. If $\hat{F}$ is a good estimated of $F$, then $Z_t = \Phi^{-1}(\hat{F}(Y_t/\overline{Y}_t))$ should be very close to a standard normal variable too.

We tested the above for both the gamma and the mixture of two exponential distributions. For sanity check, we also used kernel density estimate of the distribution $F$. As expected, the kernel density estimate from the data set did pass the Kolmogorov-Smirnov test for normality. On the other hand, for both the gamma and the mixed exponential distributions, the sample $Z_t$ failed the Kolmogorov-Smirnov test. However, they are not far from normal as shown by the first four moments of $Z_t$ in Exhibit 5. Note that the numbers outside (inside) the parentheses are for the condition where it was dry (wet) the day before.

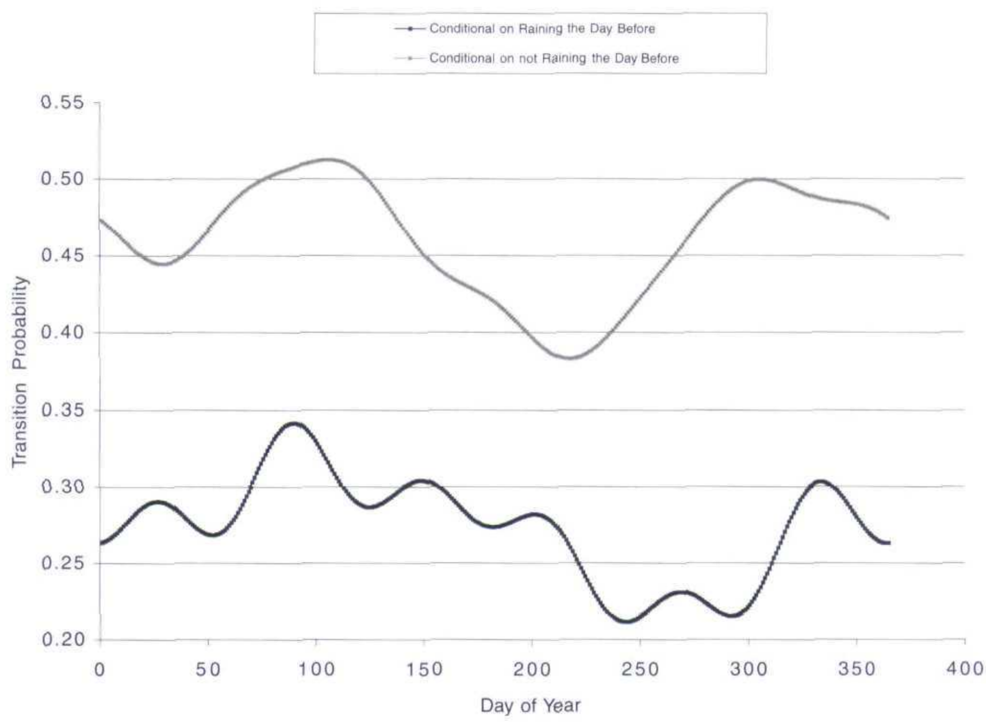
A casual examination of the first four moments suggests that the mixture of exponentials provides a slightly better fit than the gamma distribution. This is consistent with previous findings (e.g., Wilks [1998]). Kernel density is in turn superior to the exponential mixture.

### A Valuation Example

For precipitation derivatives, since the underlying is not traded, the non-arbitrage option pricing theory developed in the financial markets is not applicable. Pricing is typically a result of risk-return analysis. In other words,

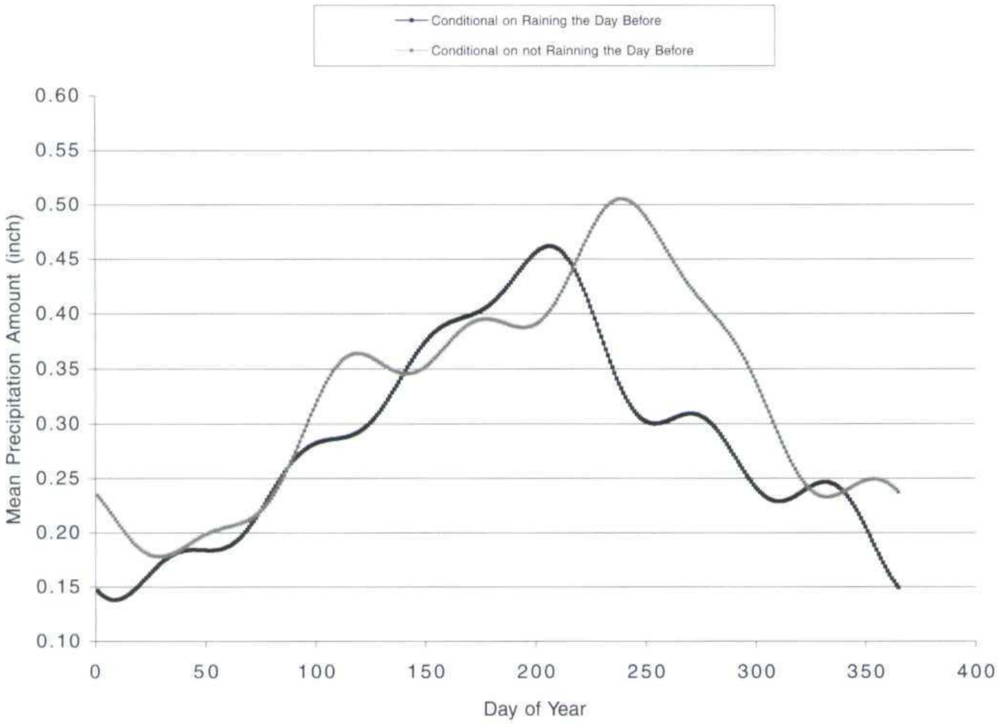
## EXHIBIT 1

**Seasonality in Transition Probabilities**




*Sample period: 1950–2003, Location: Chicago Midway Airport.*

## EXHIBIT 2
**Seasonality in Conditional Precipitation**




*Sample period: 1950–2003, Location: Chicago Midway Airport.*

## EXHIBIT 3

| Previous day's state | $\alpha$ | $\beta_1$ | $\beta_2$ | mean | Stdev |
|---|---|---|---|---|---|
| Dry | 0.447 | 0.295 | 1.573 | 1.001 | 1.345 |
| Rain | 0.429 | 0.249 | 1.563 | 1.000 | 1.358 |

## EXHIBIT 4

| Previous day's state | a | b | Mean | Stdev |
|---|---|---|---|---|
| Dry | 0.720 | 1.390 | 1.001 | 1.391 |
| Rain | 0.682 | 1.465 | 1.000 | 1.465 |

price equals the sum of expected payout plus a risk premium. Therefore, the key in valuation is the accurate projection of the payout distribution.

The daily precipitation model we have presented can be used to value most of the precipitation contracts. Since closed-form solutions are difficult to derive due to the complexity of the model, we resort to Monte Carlo simulations.

Let's consider a one-inch put option on the cumulative rainfall for the month of March at Chicago Midway Airport. We simulate 10,000 daily precipitations using each of the three first-order seasonal Markov chain models estimated in the last section. From that we can derive the simulated cumulative precipitation for the month of March. Note that all the simulation paths start with a rainy day on December 31 and end on March 31. The snow amount is converted to rainfall equivalent.

Exhibit 6 presents the precipitation results for the month of March from the historical simulation (of the last 54 years) and the model-based simulation. The last column of the table shows the value of the option (defined by the mean value of the payoff).

From the table we can see that all three models underestimate the value of the put option relative to the historical average. Most of the undervaluation is attributed to the underestimation of the standard deviation of the cumulative precipitation for the month of March. We find that this is true for almost all the months. This bias can be corrected by using a higher order Markov chain such as those discussed in Dubrovsky, Buchtele, and Zalud [2004].

Consistent with our estimation results, the performance of the exponential mixture and kernel density is superior to that of the gamma distribution, and kernel density is the best. This is not surprising in that the kernel density approach is non-parametric and therefore is the most faithful to the data. Of course, every benefit comes with a cost.

## SUMMARY AND CONCLUSION

As the overall market for weather derivatives grows, contracts on precipitation are gradually making their way into the scene. However, compared with temperature derivatives, the market share of precipitation derivatives is very minimal. The slow growth is not due to a lack of interest or demand. In fact, end-users such as farmers and

## EXHIBIT 5

| Distribution | Mean | Standard Dev. | Skewness | Kurtosis |
|---|---|---|---|---|
| Gamma | -0.025 (-0.251) | 0.986 (0.985) | 0.677 (0.652) | 3.290 (3.180) |
| Mixture of Exp. | 0.0128 (0.009) | 0.963 (0.970) | 0.306 (0.290) | 2.861 (2.870) |
| Kernel Density | -0.001 (-0.003) | 0.996 (0.998) | 0.017 (0.005) | 3.163 (3.186) |

## EXHIBIT 6

| Distribution | Mean | Standard Dev. | Skewness | Kurtosis | Put Value |
|---|---|---|---|---|---|
| History | 2.664 | 1.352 | 0.255 | 2.034 | 0.027 |
| Gamma | 2.574 | 1.154 | 0.683 | 3.603 | 0.016 |
| Mixture of Exp. | 2.629 | 1.272 | 0.779 | 3.704 | 0.020 |
| Kernel Density | 2.629 | 1.298 | 0.8559 | 3.971 | 0.021 |

hydroelectric power producers are very keen on precipitation contracts. The key culprit is the difficulty and challenge in properly modeling precipitation. When financial institutions do not have a good handle on modeling, they hesitate to provide quotes. So far, there isn't any literature on the valuation of precipitation derivatives.

The current article fills this important gap in the literature and the industry. We propose, calibrate, and compare three precipitation models: a gamma distribution, a mixture of exponentials, and kernel density. Our analyses show that the latter two models dominate the first.

Undoubtedly, many issues are still pending and further research is required. For instance, spatial correlation may be very important for certain contracts. If the contract site and the measurement site are far away, it is imperative that a proper gauge of correlation be in place. Another issue is the modeling of extreme events: prolonged period of draught or sudden, severe flood. Although most contracts concern cumulative precipitations over a particular period, the same amount of cumulative precipitation can have quite different consequences. For instance, a three-inch precipitation over a month can either occur evenly throughout or in the form of a downpour in 30 minutes. This will have quite different consequences for a farmer. Future research needs to focus on such important issues.

## ENDNOTES

Cao and Wei gratefully acknowledge the financial support from the Social Sciences and Humanities Research Council of Canada. The views expressed here are solely the authors' and are not those of the XL Weather & Energy Inc..

[1]See http://www.entergykoch.eu.com/businessgroups/SMHI_INDEX.pdf for details.

[2]Several authors have proposed various valuation models for temperature derivative. Examples include Brody et al. [2002], Campbell and Diebold [2003], Cao and Wei [2000], and Cao and Wei [2004]. The list of challenges was from Dischel [2000].

[3]We model precipitation as a year-round variable. Snowfall is converted to rainfall equivalent.

## REFERENCES

Brody, D.C., J. Syroke, and M. Zervos. "Dynamic Pricing of Weather Derivatives." *Quantitative Finance*, 2 (2002), pp. 189-198.

Campbell, S., and F.X. Diebold. "Weather Forecasting for Weather Derivatives." NBER Working Paper No. W10141, 2003.

Cao, M., and J. Wei. "Pricing the Weather." *Risk*, May 2000 pp. 67-70.

——. "Weather Derivatives Valuation and Market Price of Weather Risk." *Journal of Futures Markets*, forthcoming, 2004.

Dischel, B. "Seeding a Rain Market." *Environmental Finance*, September 2000, pp. 2-4.

Dubrovsky M., J. Buchtele, and Z. Zalud. "High-Frequency and Low-Frequency Variability in Stochastic Daily Weather Generator and Its Effect on Agricultural and Hydrologic Modeling." *Climatic Change*, Vol. 63, Nos. 1-2 (2004), pp. 145-179.

Katz, R.W. "Precipitations as a Chain-Dependent Process." *Journal of Applied Meteorology*, 16 (1977), pp. 671-676.

Richardson, C.W., and D.A. Wright. "WGEN: A Model for Generating Daily Weather Variables." U.S. Department of Agriculture, Agricultural Research Service, ARS-8, 1984.

Sansó, B., and L. Guenni. "A Stochastic Model for Tropical Rainfall at a Single Location." *Journal of Hydrology*, 214 (1999), pp. 64-73.

——. "A Non-stationary Multi-site Model for Rainfall." *Journal of American Statistical Association*, 95 (2000), pp. 1064-1089.

Wilks, D.S. "Conditioning Stochastic Daily Precipitation Models on Total Monthly Precipitation." *Water Resources Research*, 23 (1989), pp. 1429-1439.

——. "Multisite Generalization of a Daily Stochastic Precipitation Generation Model." *Journal of Hydrology*, 210 (1998), pp. 178-191