Modelling and Pricing Weather Derivatives

Mengyu Huang, Haoran Wang, Jing Xu, Jingzhu Xiong, Xiaoyu Yu

April 14, 2017

1 Introduction

Weather derivatives are financial instruments based on indices that are calculated using observed temperature data. Organizations use these derivatives to reduce risks associated with weather [1]. One type of weather derivative is **weather future** that obligates the buyer to purchase the value of the underlying weather index - measured in heating degree days (HDD) or cooling degree days (CDD) - at a predetermined future date [2]. The settlement price of the underlying weather index is equal to the value of the relevant month's HDD/CDD multiplied by $20.

The main purpose of this paper is to apply stochastic processes to model temperatures and use Monte Carlo Simulation to price these futures. We choose New York as our sample city and price for New York weather derivatives. In addition, we detect regions temperature under El Nino influence is significantly different from the ones that are not affected by El Nino, so we pick Lima as our sample city that with El Nino and develope a different model. We hope that would provide a new perspective for weather derivative pricing in these cities.

In section 2, we model the temperature based on historical temperature data using stochastic processes. In section 3, we use the HDD and CDD as underlying index and run Monte Carlo Simulations to price the derivatives. In section 4, we expand our weather models to include El Nino effects by introducing triggers and modifying parameters.

2 Modelling temperature

In this section, we develop a model to fit historical daily temperature of New York City in the past 16 years from 1/1/2001 to 12/31/2016.

2.1 Mean Temperature

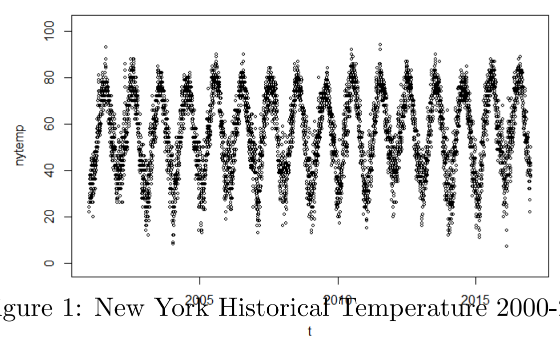

Figure 1: New York Historical Temperature 2000-2016

Intuitively, temperature patterns vary for summer and winter. After plotting the data, it is clear that the mean temperature follows a strong seasonal trend. According to the trend shown in Figure 1, we decide to use a sine-function to capture the seasonal characteristics.

The expression of process describing the temperatures with seasonal trend has the general form

$$sin(\omega t + \phi)$$

where t denotes time. For the convenience of data fitting, we instead use the form

$$Csin(\omega t) + Dcos(\omega t)$$

giving the exact same result without losing generality.

2.2 Adding the Global Warming Term

Besides the seasonal trend, we also find a positive trend in the data, which can be explained by the global warming effect: evidence has shown that the average temperature of the Earth raised from 0.4 to 0.8C during the past century.The rate of warming almost doubled over the past 50 years. Therefore, it is reasonable to add a term to show this warming trend. Here we assume the warming trend is approximately linear to time.

In summary, the deterministic model with data fitting in least squares method for the mean temperature has the form:

$$Tm = A + Bt + Csin(\omega t) + Dcos(\omega t)$$

Then adding the stochastic term, we get the complete model for temperature:

$$Tm = A + Bt + Csin(\omega t) + Dcos(\omega t) + \sigma W_t = T_m + \sigma W_t$$

Usually, people assume stochastic processes representing the temperature should have a mean-reversion property since the temperature cannot deviate too far away from the mean value. However, the temperature is probably not reverting to the mean in the long run, so we decide to utilize a revised model without assuming mean-reversion property. Therefore, the temperature is modelled as the solution of the following stochastic differential equation,

$$dT = \frac{dT_m}{dt} dt + \sigma dW_t$$

where $\frac{dT_m}{dt}$ is the drift term, and σ is the diffusion, W_t is Wiener Process.

2.3 Model fitting for New York temperature

To fit model for New York temperature, we scrape daily average temperature data of New York Laguardia Airport from 2001/01 to 2016/12.

Firstly, we fit a linear model taking the temperature mean as the response variable in R, starting with linear model fitting to get approximate estimates of all the coefficients. The result gives as below:

$$T_m = 54.2312 + 0.0004681t + 12.900sin(0.0168t) - 9.6925cos(0.0168t)$$

As our model is in a nonlinear form, it would not be accurate to use linear model for fitting it, but it does give us good estimations for all the parameters. Based on the approximate results, we use non-linear method to further refit this model. The nonlinear model fitting are using least square method to determine the parameter. Also, since the nonliner regression involves numerical computation for estimating the parameters, it requires the initial input as the start point so as to estimate the parameters. We pass the result from original linear fitting to be the start point of nonlinear method, then conducted non-linear fitting.

The model in non-linear method:

$$T_m = 54.92 + 0.0003336t - 8.459sin(0.0172t) - 20.28cos(0.0172t)$$

Figure 2 shows the two function with the temperature data. Red line shows the linear fitting result and the blue line shows the non-linear fitting result. It can be easily seen that the nonlinear method gives a better fitting.

In the complete model equation, a stochastic term $W_t$is added, and the diffusion term is to be estimated. We used the New York historical temperature data to estimate the volatility. The estimator is simply based on the quadratic variation of T_t.

$$\hat{\sigma} = \frac{1}{N} \sum_{j=0}^{N-1} (T_{j+1} - T_j)^2$$

And the estimated volatility equals 5.79 in this work.

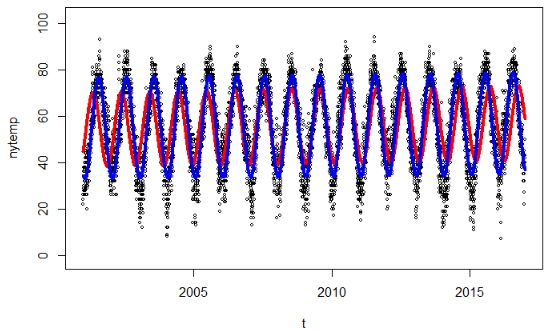

Figure 2: Fitted New York Temperature

3 Pricing weather derivatives

3.1 Derivative introduction

The first weather derivative market transaction took place in the US in 1997.[3]. Different models have been proposed for pricing structured weather derivatives, such as options, futures and swaps. Two commonly used indices, heating degree-days (HDDs) and cooling degree-days (CDDs), are underlying indices of derivatives in this report. HDDs are used to measure the demanded energy for heating in winter, while CDDs are used to measure the demanded energy for cooling in summer[4]. Measurement periods are either seasonal, i.e, they correspond to the summer season from May until September and the winter season from November until March, or monthly. April and October contracts are traded with both HDD and CDD, while all other winter and summer contracts are traded with HDD and CDD respectively. [5]

$$HDD_i = max(65 - T_i, 0)$$

$$CDD_i = max(T_i - 65, 0)$$

HDD_i is the number of HDDs for day i; CDD_i is the number of CDDs for day i; T_i is the average temperature; 65 F is the baseline temperature. The HDDs and CDDs are the number of degrees that the real temperature for a day deviates from the baseline temperature level, because people are inclined to use more energy to heat or cool the their homes as the temperatures go below or beyond the reference level 65 F.

$$H_n = \sum_{i=1}^{n} HDD_i, C_n = \sum_{i=1}^{n} CDD_i$$

Monthly or seasonally HDD/CDD index H_n/C_n is defined as the sum of HDDs during that period. In this study, we mainly focus on the weather future, whose contract value equals (undelying HDD/CDD indices) *(20 USD per points).

3.2 Derivative SDE Equation

The weather derivative market is a typical incomplete market, as the underlying asset, temperature is non-tradable [6]. Therefore, since price cannot be obtained from the objective measure, we need to consider the market price of risk under some specific risk measure. We simplify the market price of risk as a constant

λ.Then write SDE under the new martingale measure Q (Here we only consider HDD; CDD follows the same method) :

$$dT = (\frac{dT_m}{dt} - \lambda\sigma)dt + \sigma d\bar{W}_t$$

Combining with the formula of HDD index, we get:

$$H_n = \sum_{i=1}^{n} HDD_i = \sum_{i=1}^{n} max(65 - T_i, 0)$$

$$E^Q[H_n|F_t] = E^Q[\sum_{i=1}^{n} HDD_i|F_t] = E[e^{\int r(u)du} \sum_{i=1}^{n} HDD_i] = \int_0^{65} \sum_{i=1}^{n}(65 - T_i) * f(T_i)dT_i$$

To get the distribution of T_i, we solve for the explicit formula. However, as the Q measure depends on risk price λ, the HDD indices is highly dependent on how we choose λ, which would be hard to determine in practice. Alternatively, we use a numerical way to solve price, namely Monte Carlo simulation method.

3.3 Derivative Pricing Using Monte Carlo Simulation

According to the temperature model developed above, we use the SDE to run the Monte Carlo Simulation to predict future temperature paths. Here we set the number of paths to be 1000. With predicting the temperature, we can get the predicted HDD/ CDD index accordingly.

Simulating result is to predict the HDD index value for 2017 February and March, and CDD index value for 2017 August and September, and to compare simulation results to the average index values observed in the market. As the CDD Indices for 2017 August and September are not available yet, we list the market CDD index value and simulated index value for 2016 August and September for reference. The comparison table is listed as below.

Index Value	16/08 CDD	16/09 CDD	17/02 HDD	17/03 HDD	17/08 CDD	17/09 CDD
Market Index Value	516	307	807.8	715	NA yet	NA yet
Simulate Index Value	496	353	824	723	513	341

Table 1: result table

4 Temperature Modeling for El Nino Cities

4.1 Introduction of El Nino

Over the past 50 years, the average global temperature has increased at the fastest rate in recorded history. According to independent analyses by NASA and NOAA, 2016 recorded the warmest earth surface temperature since modern record keeping began, making 2016 the third consecutive year to set a new global record[7] [8]. El Nino events are occurring more frequently in recent years and the increasing frequency of extreme El Nino events is associated with rising global climate [9].

El Nino is a complex and naturally occurring weather pattern characterized by unusually warm ocean temperatures in the central and east-central equatorial Pacific. It can be seen in measurements of the sea surface temperature, and is associated with rising air temperature [10]. It has a tremendous impact on agriculture, with warm (El Nino) years unfavorable for fishing and marked by devastating floods, and cold years unfavored by farmers and marked by drought and crop failures [11]. The successful prediction of El Nino benefit strategic agriculture planning, the management of water resources, and reserves of grain and fuel oil, all of which have huge economic incentives.

In this section, we want to predict El Nino events using our stochastic temperature model. For the purpose of study, we choose to study temperatures of regions that are exposed to by El Nino. We use Perus capital Lima as our sample which is located in the east-central equatorial Pacific. We model temperature variations of Lima using local temperature from 2001 and 2016, and predict El Nino by adding a trigger condition.

Firstly, we used R to linearly fit the temperature mean to get corresponding coefficients.

$$T_m = 66.13 + 0.0004943t - 2.229sin(0.0201t) + 4.333cos(0.0201t)$$

As we discussed earlier, it would be more accurate to use nonlinear model. The non-linear model is:

$$T_m = 65.87 + 0.0005654t + 4.088sin(0.01717t) + 4.823cos(0.01717t)$$

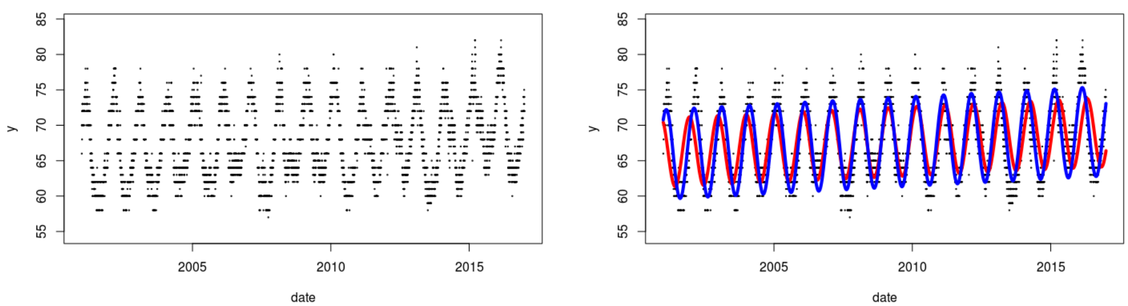

Figure 3: Lima's Temperature(Left) and Fitting Lima's Temperature(Right)

4.2 Setting trigger for El Nino

In our model, we assume that El Nino event is triggered if the average month temperature is higher than last year's average month temperature in the corresponding three months. The trigger will print out the date that the condition is met. The dates printed out are from 2014-06-12 to 2016-02-06, confirming that an intensive El Nino happened during 2015-2016. The following is the temperature from 2014-06-12 to 2016-02-06, the time period when El Nino happened, and we intend to model this period's temperature specifically.

The temperature mean is fitted in R. We firstly use linear model fitting to get approximate estimates of all the coefficients. The result gives as below:

$$T_m = 67.2785 + 0.0058093t - 3.7880sin(0.0201t) - 3.2132cos(0.0201t)$$

As our model is in nonlinear form, it is be more accurate to conduct nonlinear model fitting like we discussed earlier. The model using non-linear method:

$$T_m = 67.6138 + 0.0070450t - 5.8324sin(0.0171194t) - 0.5306cos(0.0171194t)$$

The red and blue line are the corresponding linear and non-linear fits. We can see that the blue line, the non-linear fit, is a better fit for the raw data. Compared to the period without El Nino, We can detect that the coefficient has a jump of 1.74, increases from 65.87 increases to 67.61. It also has a different osccillation, and the absolute of sine coefficient increases from 4.088 to 5.832. We concluded that the el Nino temperature model is significant different from the model without its influence.

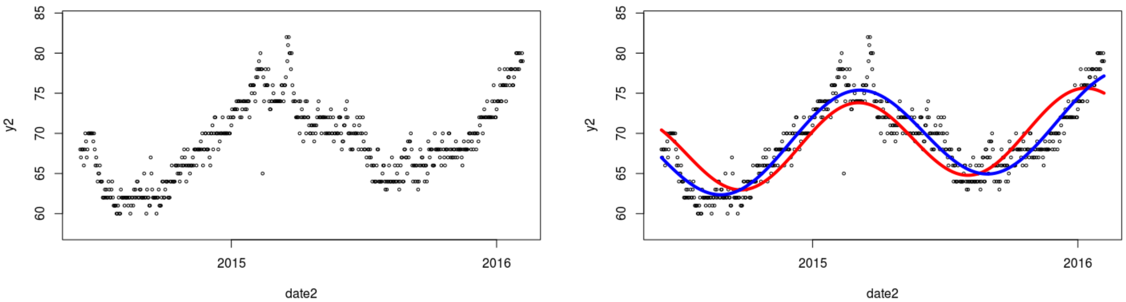

Figure 4: El Nino from 2014-2016 (Left) and Model fitting of El Nino Years (Right)

5 Conclusion

In this report, we applied techniques in stochastic processes to model weather temperature. Using New York Laguadias Airport temperature as our sample data, we derive our non-linear model,

$$T_m = 54.92 + 0.0003336t - 8.459sin(0.0172t) - 20.28cos(0.0172t)$$

It is the solution of a stochastic differential equation with sine-function allowing it to revert to the mean in the long run. Also, former reseachers have paid attention to pricing weather options in monte carlo method, we in contrast run Monte Carlo simulation to price the weather futures in this work. Comparing the simulated results with market indices, we conclude they are approximate with each other, indicating that our previous modeling for temperature is effective. Thus, the temperature model could be considered in further research and weather derivatives pricing.

When pricing model of the Peru temperature, we add an additional trigger term, since the original model is slightly inappropriate in this senario. As we can see from the two Limas data plot, with El Nino, the coefficient 65.87 increases to 67.61 (a jump of 1.74), also the absolute coefficient of sin increases from 4.088 to 5.832, indicating that the oscillations increase. Hence, it is tenable to conclude that the pricing model for weather derivatives is significantly different from the one for New York, which is not affected by the El Nino. Although few literatures are concerned with effects of El Nino on derivatives pricing, it is sensible to consider the effects on places where irregular and complex clamatic changes may take place, due to potential need for hedging natural catastrophe.

References

[1] G. Dorfleitner and M. Wimmer. The Pricing of Temperature Futures at the Chicago Mercantile Exchange. *Journal of Banking and Finance*, 34(6):1360–1370, 2010.

[2] M. Cao and J. Wei. Weather Derivatives Valuation and Market Price of Weather Risk. *The Journal of Futures Markets*, 24(11):1065–1089, 2004.

[3] G. Considine. Introduction to Weather Derivatives.

[4] P. Alaton, B. Djehiche, and D. Stillberger. On Modelling and Pricing Weather Derivatives. *Applied Financial Mathematics, 9, 1-20*, 9:1–20, 2002.

[5] S. Huang. Modelling Temperature and Pricing Weather Derivatives, 2014.

[6] E. Benth and J. SALTYT. Stochastic Modelling of Temperature Variations With A View Towards Weather Derivatives. *Applied Mathematical Finance*, 12:53–85, 2005.

[7] NASA. NOAA Data Show 2016 Warmest Year on Record Globally, 2017.

[8] G. Schmidt and D. Arndt. Annual Global Analysis for 2016, 2017.

[9] W. Cai. Increasing Frequency of Extreme El Nino Events Due to Greenhouse Warming. *Nature Climate Change*, 4:111–116, 2014.

[10] K. Trenberth. Evolution of El Nio Southern Oscillation and Global Atmospheric Surface Temperatures. *Journal of Geophysical Research*, 107 (D8):4065, 2002.

[11] T. Jones. Agricultrual Applications of Weather Derivatives. *International Business and Economics Research Journal*, 6(6):53–60, 2007.