Dept. of Math./CMA University of Oslo
Statistical Research Report no 3
ISSN 0806–3842 December 2011

PRICING OF TEMPERATURE INDEX INSURANCE

CHE MOHD IMRAN CHE TAIB AND FRED ESPEN BENTH

Abstract. The aim of this paper is to study pricing of weather insurance contracts based on temperature indices. Three different pricing methods are analysed: the classical burn approach, index modelling and temperature modelling. We take the data from Malaysia as our empirical case. Our results show that there is a significant difference between the burn and index pricing approaches on one hand, and the temperature modelling method on the other. The latter approach is pricing the insurance contract using a seasonal autoregressive time series model for daily temperature variations, and thus provides a precise probabilistic model for the fine structure of temperature evolution. We complement our pricing analysis by an investigation of the profit/loss distribution from the contract, in the perspective of both the insured and the insurer.

1. Introduction

Weather index insurance is a class of products targeted to households in developing countries (see Barnett et al. [1, 2, 9], Sakurai and Reardon [14] and Skees [16, 17, 18]). Such insurance contracts have close resemblance with weather derivatives, since the claim is tied to the value of a weather index measured in a specific location. In classical weather-related insurance contracts, the insured must prove that a claim is justified based on damages. The weather index contracts refer to an objective measurement, like for instance the amount of rainfall or the temperature in a specific location. As such, weather index insurance accommodates a transfer of risk for droughts or flooding, say, from households in rural areas in Africa to insurance companies. The premium to pay for buying a weather index insurance is our focus.

Consider a weather index insurance written on a temperature index. For example, we may consider a contract giving protection against unusually high temperatures over a given period in the season for growing crop. If temperatures are above a given limit, then there may be a significant risk of dry conditions leading to a bad harvest. The limit or predefined threshold is the point where payments start. Once the threshold is exceeded, then the payment is calculated as how much it goes beyond the limit. We may construct

Date: December 5, 2011.

Fred Espen Benth acknowledges financial support from the project "Energy markets: modelling, optimization and simulation" (EMMOS), funded by the Norwegian Research Council under grant 205328/v30.

a contract which pays out a certain amount of money according to an index value over the period, for example, based on a CDD-type (cooling-degree day) index.

Suppose, for further concreteness, a contract paying

$$X(\tau_1, \tau_2) = k \times \sum_{s=\tau_1}^{\tau_2} \max(T(s) - c, 0), \tag{1.1}$$

that is, an amount k times the CDD index over the time period τ_1 to τ_2, where c is some threshold. Both k and c are positive constants, and c measures the critical temperature level, and k is the conversion factor transforming the weather index into money. Both k and c are contractual parameters, while the temperature is measured at some agreed station. The insured will receive the amount $X(\tau_1, \tau_2)$ of money at time τ_2, against paying a premium for the insurance at time $t \leq \tau_1$. We note that the index and actual experienced losses by the insured are obviously not perfectly correlated. There is a possibility of the insured receiving no indemnity even though experiencing a loss, and in contrast, the insured may also receive the indemnity with having a loss.

Given a reliable temperature model, the insured may assess the distributional properties of $X(\tau_1, \tau_2)$ and determine if the contract provides the protection sought for. A major problem is of course the spatial risk incurred by the location of the temperature measurement station relative to the location where the insured is living. The insurance company offering the protection will most likely wish to settle the contract against an index calculated from an official measurement station, which typically are existing only in major cities. The spatial risk may be significant, and the insurance contract may only provide partial protection against the real temperature risk. In Šaltytė Benth et al. [15] a spatial temperature model is presented, and in Barth et al. [3] questions concerning hedging of spatial risk using weather derivatives are analyzed. The spatial consideration in the pricing of rainfall insurance has been done by Turvey [19].

Important in the assessment of the contract is the premium charged by the insurance company. The standard approach to pricing the contract is by finding the expected value of the claim size $\mathbb{E}[X(\tau_1, \tau_2)]$, adjusted for risk (sometimes called the risk loading), and discount it by some interest rate to obtain the present value,

$$P(t, \tau_1, \tau_2) = \exp(-r(\tau_2 - t))\mathbb{E}[X(\tau_1, \tau_2) \,|\, \mathcal{F}_t] \,. \tag{1.2}$$

Since the money is paid at the end of the measurement period of the index, at time τ_2, we discount by $\exp(-r(\tau_2 - t))$ to get the present value, with $r > 0$ being the discount rate assumed to be a constant. We use continuously compounding discount rates in our analysis. The filtration $\mathcal{F}_t$ denotes all the available information in the market up to time t, which the insurance company will take into account in its pricing.

Our concern is to study the price $P(t, \tau_1, \tau_2)$ using three different approaches. The first approach is the so-called *burn analysis* advocated in Jewson and Brix [10] as the classical method to price weather derivatives. The burn analysis is based on the empirical distribution of the payoff $X(\tau_1, \tau_2)$ within the sample data collected. The price is calculated using the mean value of the observations. Next, we apply the slightly more sophisticated *index modelling* approach presented in Jewson and Brix [10], which amounts in fitting a distribution to the historically observed claims $X(\tau_1, \tau_2)$, and price the contract based on the expected value of the distribution. Lastly, we propose a very detailed modelling approach, where the daily temperature dynamics is modelled by a time series. Based on empirical findings, an *autoregressive*, AR(p), model turns out to be highly suitable for describing the dynamics of temperature evolution. Using this model, one may compute the index $X(\tau_1, \tau_2)$, and find the expectation for pricing the insurance contract. There are several advantages with this approach. First, we obtain a consistent framework for pricing insurance contracts for various given time periods $[\tau_1, \tau_2]$ without having to re-estimate a distribution (as in the index approach), or collecting data (as in the burn analysis). Furthermore, we can use current information on the temperature to price the insurance contract. The index modelling and burn analysis will not use any dynamical model, and hence we cannot take into account current information when pricing. This is an important aspect for the insurance company in their risk assessment of the contracts. As a final remark, a detailed time series model of the temperature dynamics is likely to capture the statistical properties better that simply looking at the historical data (burn analysis), or fit an "arbitrary" distribution to the historical data (index modelling). The two latter approaches also suffer from little available data compared to the situation for the temperature modelling approach, where the amount of information is far better.

We will analyse the profit/loss distribution for both the insurer and the insured. The two factors determining the profit/loss distribution are of course the index $X(\tau_1, \tau_2)$ which settles the payoff, and the price of the contract. The insurance company wants to stay solvent, and thus charges an additional premium on the "fair" expected value. However, the higher premium, the less attractive will these insurance contracts become. We emphasis here that the index-based contracts are very different from traditional insurance, as the insurance company cannot diversify its risk by attracting many clients. In fact, the weather index-based contracts are very similar to financial derivatives. For example, to illustrate matters by a simple case, an insurance company issuing one contract on a temperature index in a given city, will have to pay $X(\tau_1, \tau_2)$. However, if it issues 100 contracts, it must pay 100 times this amount. In traditional non-life insurance, the risk of paying out insurance claims are distributed among the clients, and the company would on average every year have to pay an *expected claim size amount* when having 100 clients, that is, only a fraction of the insured will make a claim. With temperature index contracts, one will risk that *all* clients claim the insurance one year, whereas the next year *none* will claim a payoff. This is parallel to how options functions in financial markets. Hence, in the case of weather index insurance, the insurance company has bigger variations in their claim payoffs as in traditional insurance. This means bigger risk, and higher prices as a consequence thereof.

On the other hand, the insurance company may hedge their risk using financial weather derivatives traded, for example, on the Chicago Mercantile Exchange. Such weather derivatives are not yet being traded on temperature indices in the developing part of the world, like Africa say, but there may be a demand for such products with an increase of weather insurance contracts. The insurance company may also hedge their risk by exploiting weather correlation. From a careful analysis of temperatures in different locations, the insurance company may identify places with independent or negatively correlated weather patterns. This would offer possibilities to spread risk for the insurance company, and thereby lower prices of the contracts. In our analysis, we shall not analyse such hedging opportunities in more detail, but focus on the effect of charging a risk loading on the contracts. In order to compare the different pricing approaches, we use the a risk loading which is based on the quantiles of the index $X(\tau_1, \tau_2)$, since this is observable in all three approaches.

We use daily temperature data from Malaysia in our analysis. Agriculture is the main economical activity in Malaysia. 15.3 percent of the work force was employed in agriculture 2000, contributing approximately 8.9 percent of the national GDP (Prime Minister's Department [12]). The favourable Malaysian tropical climate spur the production of various crops including the main export articles rubber, palm oil and cocoa. Almost 24 percent of the whole area in Malaysia is allocated for agricultural activity.

The paper is organized as follows. In Section 2, we present the temperature data from Malaysia and price empirically the weather insurance contracts based on the three approaches. The next Section 3 is devoted to some risk analysis, seen both from the insured and the insurers point of view. Finally, the conclusion is given in Section 4.

2. Pricing temperature index insurance contracts

In this section we investigate three methods of pricing of a single temperature index insurance contract. The methods include the classical way of pricing weather derivatives using classical burn approach and index modelling (see Jewson and Brix [10]), and a third method of pricing based on a dynamical temperature model adopted from Benth et al. [4, 5, 6].

For all pricing methods, we analyse the CDD-index $X(\tau_1, \tau_2)$ defined in (1.1). As we want to perform an empirical analysis of the performance of the different pricing methods, we must choose a threshold level c. Obviously, since these insurance contracts are mainly targeted for farmers, the threshold c will be dependent on the crop grown. For the temperature insurance contract to be attractive for the farmer, the level c must be so that it reflects the harmful threshold for his or her crop. For the sake of illustration, we choose $c = 28$ in our studies, imagining that temperatures above this threshold may imply drought, harming the growth of a specific crop. The money factor k is fixed to RM50 per unit of the contract. We set a low value of the money factor just for ease rather than dealing with any substantial amount. We further concentrate our analysis to January, that is, the time span $[\tau_1, \tau_2]$ means the month of January.

In our analysis of prices, we shall first compute "fair prices", based on the expected payoff of $X(\tau_1, \tau_2)$, appropriately discounted to present values. However, such prices will not compensate the insurer for taking on the risk, and a risk loading will be added in the real pricing of the contract. We will also add a risk loading, based on adding 5% of the 95% quantile of the payoff $X(\tau_1, \tau_2)$. For the insurance company, the 95% quantile of $X(\tau_1, \tau_2)$ means a size of loss which is exceeded with 5% probability.

2.1. Description of the data.

We have obtained daily average temperature (DATs) data measured in degrees of Celsius from the Malaysian Meteorological Department, which are observed over the period ranging from 1 January 1971 to 31 December 2010. The data have been collected in Petaling Jaya, Malaysia, as the nearest station to the capital city Kuala Lumpur. A number of 14,610 records covering 40 years of DATs data are observed, however, the amount of data is reduced to 14,600 after removing the measurements on February 29 in each leap year to synchronize the length of the years to 365 days in the analysis. A small amount of defective or missing values have been detected in the data, constituting only 0.48% of the total sample. These data observations have been corrected by using the average temperature between the previous and following day observations. The time series of the average DATs is plotted in Figure 1. For the purpose of illustration, we just show the last 10 years of the time series data.

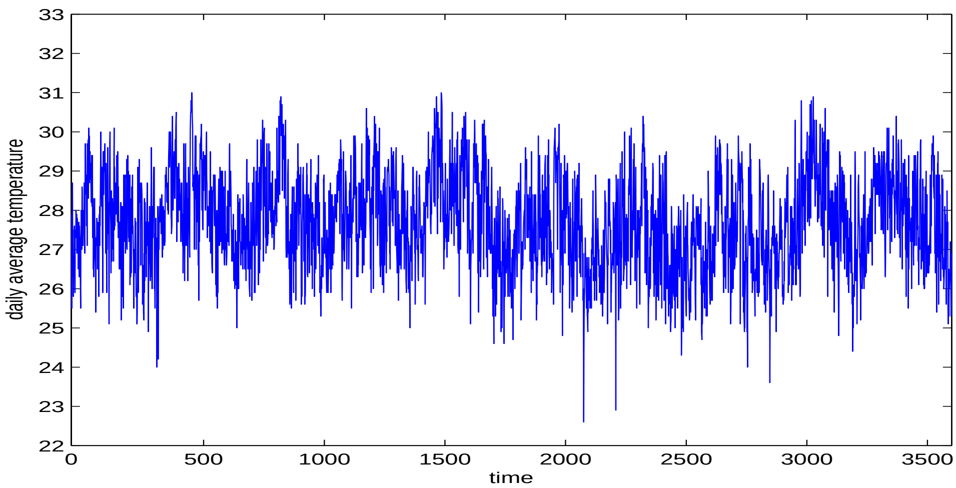

FIGURE 1. Petaling Jaya DATs for the period starting 1 January 2001 to 31 December 2010.

Note that the lowest and highest temperature recorded in the data set are 22.3 and 31.2, respectively, with the mean being equal to 27.4. We observe a rather small variation in the data. In Figure 2, we present the histogram of DATs in Petaling Jaya. The skewness coefficient is 0.010, indicating nearly symmetric data. The kurtosis is 3.011, showing that the data are not normally distributed.

2.2. **Burn analysis.** Burn analysis is a very simple method that is traditionally used for pricing a weather derivatives contracts (see Jewson and Brix [10]). It simply uses the empirically computed mean value of the observed index $X(\tau_1, \tau_2)$. In our case, we computed the index $X(\tau_1, \tau_2)$ for January each year in the data sample, altogether yielding 40 samples of $X(\tau_1, \tau_2)$. Figure 3 shows the histogram of $X(\tau_1, \tau_2)$. Based on the 40 observed values, we compute the empirical mean, and discount it by $\exp(-r(\tau_2 - t))$ in order to get the present value at time t of the contract. In the calculation of the premium, we do not include any information $\mathcal{F}_t$ on present or historical temperatures, since this is naturally *not* entering into this pure data driven approach. The conditional expectation

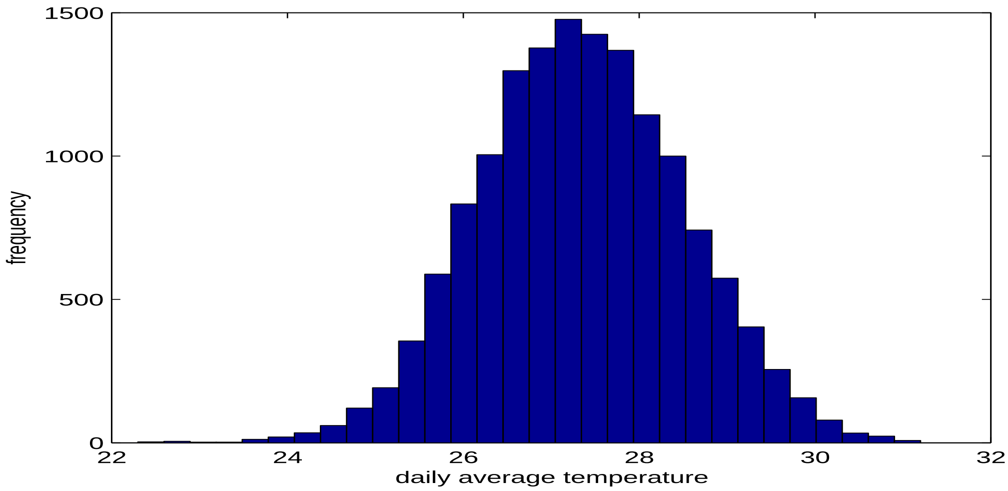

FIGURE 2. Histogram of daily average temperature in Petaling Jaya.

is transformed into a standard expectation. As $t \to \tau_1$, the price will converge to the expected claim size $X(\tau_1, \tau_2)$ as estimated from data. The empirical mean was estimated to be 118.25.

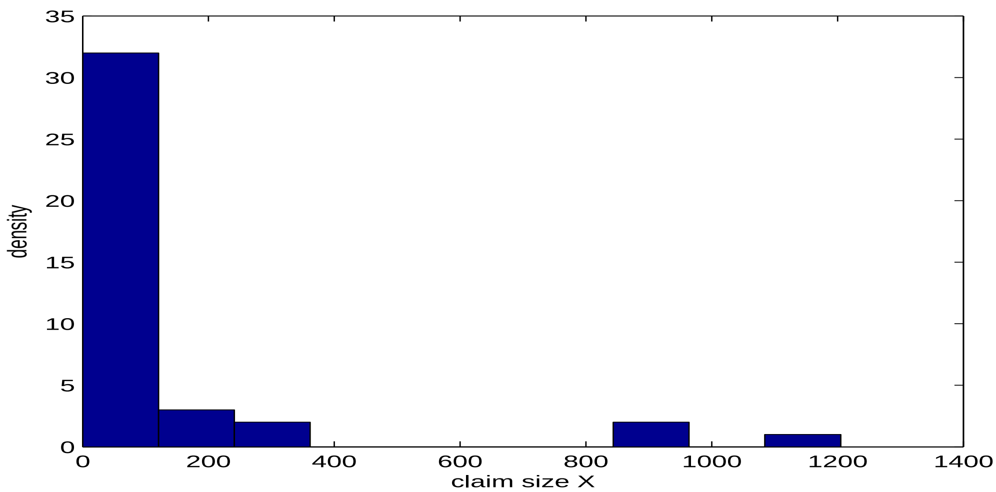

FIGURE 3. Histogram of the claim size X.

2.3. **Index modelling.** Index modelling is a pricing method for weather derivatives based on a distribution statistically modelling the claim size. To the observed claim sizes $X(\tau_1, \tau_2)$, one selects and fits a distribution, and computes the mean of this distribution to find the expected value of $X(\tau_1, \tau_2)$. The advantage of this method is that we may derive statistically information of the claims outside the range of the observed data values, and can make assessments of the probability of extreme events happening. In particular, we may estimate quantiles of the claim $X(\tau_1, \tau_2)$ outside the range of observed data. However, we note that the data backing up the estimation of the distribution is the same as for the burn analysis, which in our case of Malaysian data amounts in only 40 values.

From the histogram of the claim size $X(\tau_1, \tau_2)$ in Figure 3, one may propose an exponential distribution in modelling the claims. Recall the density with parameter μ of the exponential distribution as

$$f_{\exp}(x;\mu) = \frac{1}{\mu}\exp\left(\frac{-x}{\mu}\right). \tag{2.1}$$

We apply the maximum likelihood (ML) method to estimate the parameter of distribution. We estimated μ to be $\hat{\mu} = 118.25$. As the parameter μ of the exponential distribution is the mean, we find the expected claim size to be 118.25. This estimate coincides with the result of the burn analysis, not unexpectedly as the maximum likelihood estimation in this case will be based on the mean value of the data. We show the empirical density of the claims with the fitted exponential distribution in Figure 4. The confidence interval

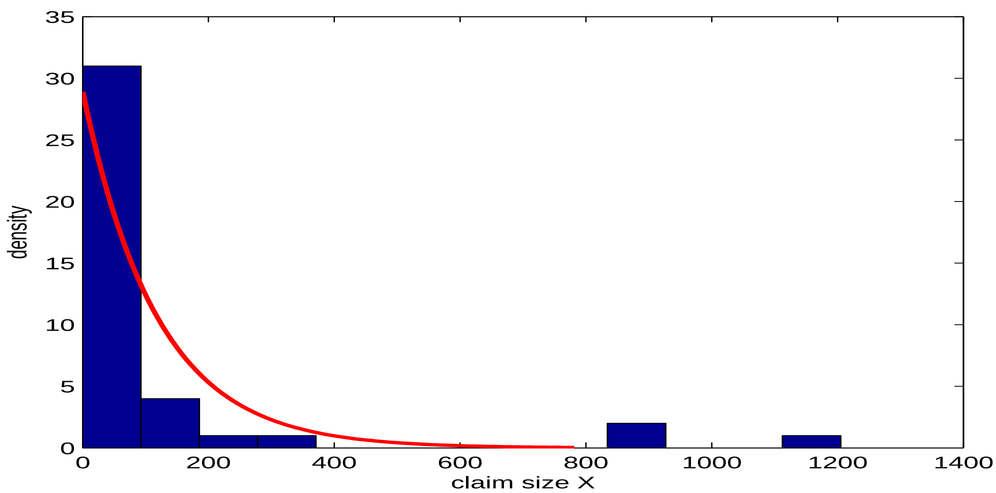

FIGURE 4. The empirical density together with fitted exponential.

for the estimate $\hat{\mu}$ has upper and lower limit 81.33 and 184.87, respectively, at the 1% significance level. This is a very wide confidence interval, demonstrating clearly the huge statistical uncertainty in this approach to model the claims. The reason is obviously that we have only 40 data points available. We note that the mean claim size estimated for the burn analysis will be infected by the same uncertainty. This is a drawback with these two methods. In a real world application, an insurance company is likely to charge a premium also for this uncertainty, leading to more expensive insurance contracts.

We also note that for the index modelling approach, the historical temperature records up to current time t, $\mathcal{F}_t$, do not play any role in the pricing. We do not create a dynamical model of the index, and hence there is no natural definition of the filtration. Therefore, we also in this case compute an expectation rather than a conditional one when assessing the insurance premium of the contract.

2.4. **Temperature dynamical modelling.** As an alternative to the burn analysis and index modelling approach, we propose to model the time dynamics of the temperature evolution. Taking into account 40 years of *daily* data, the amount of information available for such a model is significantly more advantageous than the poor 40 data points for the burn and index model methods.

Suppose that the temperature $T(t)$ at time $t \geq 0$ is given as follows

$$T(t) = S(t) + Y(t), \tag{2.2}$$

where $S(t)$ is a deterministic seasonal mean function and

$$Y(t) = \sum_{i=1}^{p} \alpha_i Y(t-i) + \epsilon(t)\,. \tag{2.3}$$

Here, $\epsilon(t)$ are i.i.d normally distributed noise with mean zero. The AR(p)-process $Y(t)$ models the random fluctuations around the seasonal mean, or, in other words, the dynamics of the deseasonalized temperatures, $T(t) - S(t)$.

The seasonal mean function $S(t)$ is defined as

$$S(t) = a_0 + a_1 t + a_2 \sin\left(\frac{2\pi(t-a_3)}{365}\right)\,. \tag{2.4}$$

The constants a_0 and a_1 describe the average level of temperature and slope of a linear trend function, respectively, while the amplitude of the mean is represented by a_2. A constant a_3 is referred to as the phase angle.

From this temperature dynamics, we may compute the temperature index and subsequently the payoff $X(\tau_1, \tau_2)$. Thereafter, we may compute the expected payoff. The advantage now is that we get a very precise model for the temperature and a price which takes current temperature knowledge into account. Moreover, the model is flexible in pricing contracts settled on various periods of the year without having to perform a statistical re-estimation like in the burn approach and index modelling. We efficiently exploit 40 years of daily data to get a detailed statistical description of the claim size distribution.

A look at the ACF of the DATs in Figure 5 shows that there are clear seasonal effects in the data, but also apparent signs of mean-reversion. The latter is observed from the decaying ACF for moderate lags. This indicates that our model is appropriate. To estimate it to data, we use a step-by-step procedure, where first we estimate the trend and seasonal component in $S(t)$, and next find the best AR(p) model fitting the deseasonalized temperature data.

We start by checking for the presence of a linear trend. By simple least squares, we obtain the slope equal to 0.0001 and intercept being 26.8. Although the trend slope seems

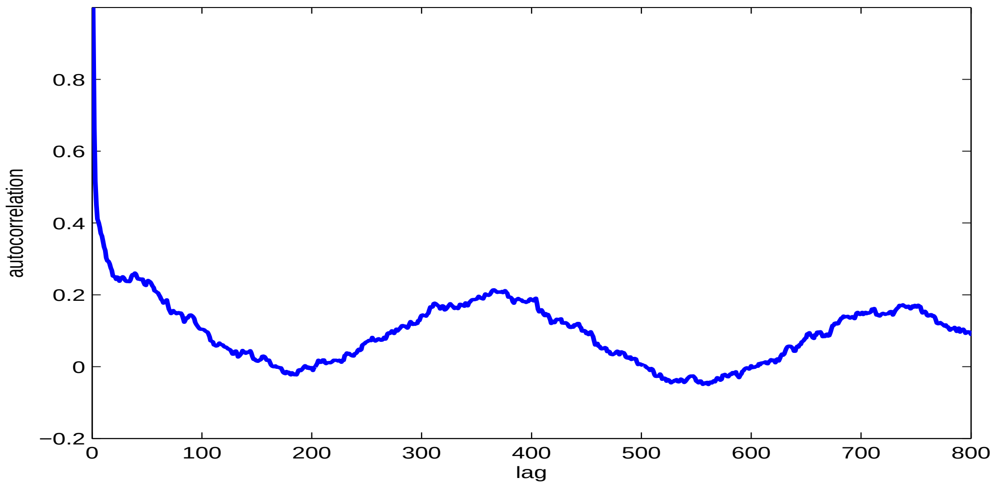

Figure 5. Empirical ACF of daily average temperature in Petaling Jaya.

Table 1. Estimated parameters for seasonal fitting

a_0	a_1	a_2	a_3
26.8373	0.0001	0.5673	56.1070

to be very small, the P-value of 0.0000 validates that it is significant. Next, we fit the data with the complete seasonal function $S(t)$ given by (2.4) using least squares method. The estimated parameters are presented in Table 1. We find an R^2 value of 18.4%, indicating that there is not much explanatory power in the seasonality function $S(t)$. The DATs for the last 5 years is plotted in Figure 6 together with the fitted seasonal function. We conclude that there is not a very pronounced seasonal variation in the data. This is unlike the observations in most of the European countries, which have high temperatures in summer and low in winter. Temperature analysis for Stockholm, Sweden, shows a clear seasonality (see Benth et al. [4, 5]), similar to the findings in USA and Lithuania (see Campbell and Diebold [7] and Šaltytė Benth et al. [15] respectively). Malaysia on the other hand experiences two seasons in general, with a dry season usually ranging from March to October and rainy from November to February whereby June and July are recorded as the driest months of the year. Nevertheless, we find the existence of a small seasonal variation which we include for further analysis.

Next, we eliminate the linear trend and seasonal components by subtracting the estimated $S(t)$ from the original observations and plot the autocorrelation (ACF) of residuals as in Figure 7. It shows the positive strong autocorrelations which rapidly decays toward zero. Noteworthy is that we do not observe any strong seasonal variation anymore in the ACF, which, despite a small seasonality pattern $S(t)$, has a clear impact on the ACF. By inspecting the partial ACF (PACF) plot in Figure 8, we observe a very high spike in lag

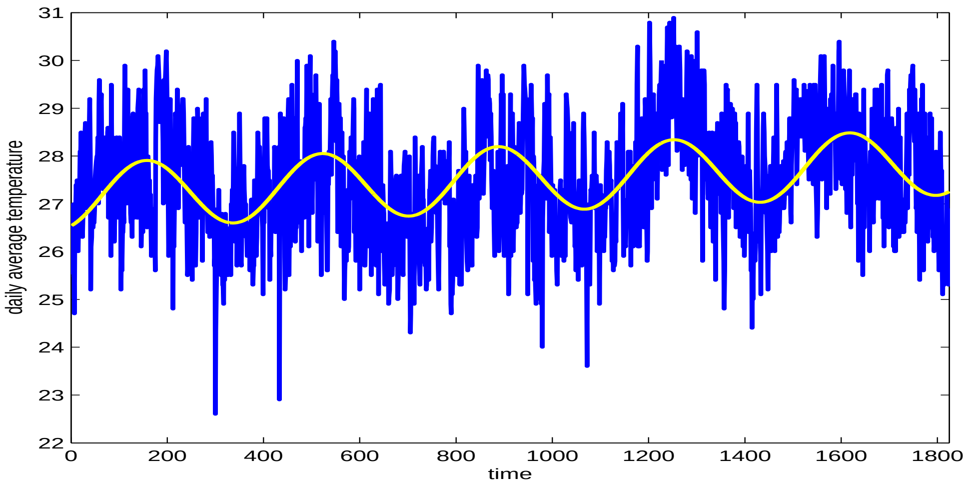

FIGURE 6. DATs in Petaling Jaya with fitted seasonal function.

1, thus suggesting AR(1) to be the most preferable model explaining the evolution of the time series.

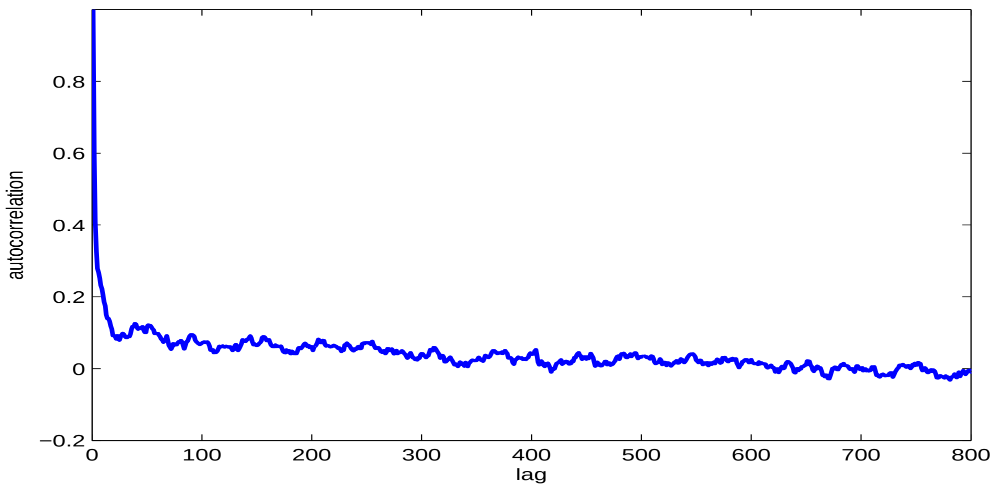

FIGURE 7. The ACF of the residuals of DATs after removing linear trend and seasonal component.

We estimate the parameter α_1 for the AR(1) process by using a simple linear regression of the detrended and deseasonalized data and found $\alpha_1 = 0.5895$. The α_1 value corresponds to the mean reversion rate where temperature reverts back to its long term mean at this speed. This indicates that the speed of mean reversion is rather fast. As suggested by Clewlow and Strickland [8], the half life of Ornstein-Uhlenbeck process with mean reversion β driven

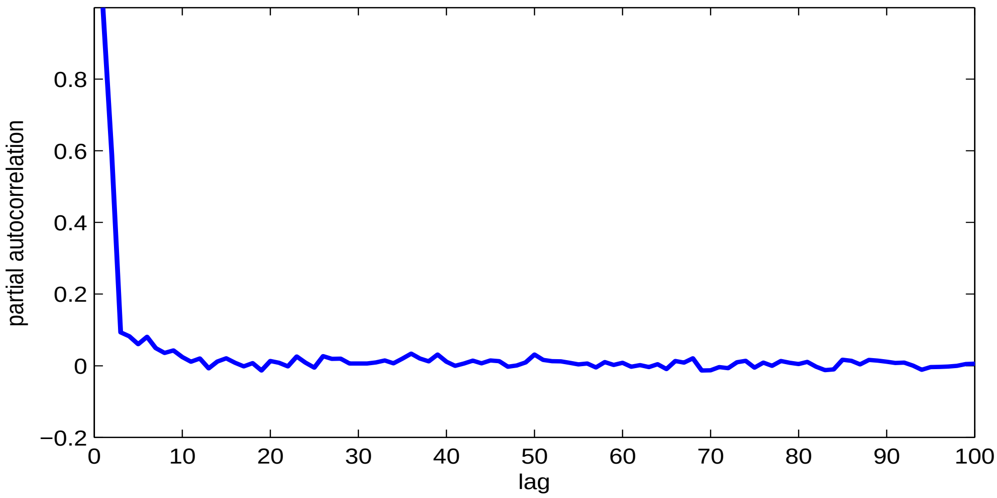

FIGURE 8. The PACF of the residuals of DATs after removing linear trend and seasonal component.

by a Brownian motion is given by

$$T_\beta = \frac{\ln(2)}{\beta}. \tag{2.5}$$

Converting the speed of mean reversion of our time series dynamics into a continuous time mean reversion, we find $\beta = -\ln(\alpha) \approx 0.5284$. This implies an estimate for the half life of temperature dynamics being $T_\beta = 1.18$, meaning that on average the temperature takes 1.18 days to revert half-way back to its long term level.

Looking at the histogram of residuals in Figure 9, we may say that it follows the normal distribution. But the Kolmogorov-Smirnov statistics of 0.022 is significant at the 1% level, and we cannot reject the hypothesis of nonnormality of data. However, taking the large number of data into account, it is very hard to pass through a normality test. We find the normal distribution a satisfactory choice for the residuals. The residuals and squared residuals for the last 10 years are plotted in Fig. 10. Looking at the squared residuals, one may suspect the existence of clustering, indicating some pattern in the residuals that may be modelled using a seasonal variance in line with Benth et al. [5], or stochastic volatility (see Benth and Saltyte Benth [6]). However, the effects seem to be minor, and we decided not to increase the level of sophistication of our model. The estimated standard deviation for the noise term $\epsilon(t)$ is 0.9571.

2.4.1. *Temperature dynamics insurance pricing.* From (1.2), we have the price of the contract with the conditional expected value, using the filtration generated by the time series $Y(t)$ in the dynamics of $T(t)$. We can simulate this conditional expectation by appealing to the Markov property of $Y(t)$. Thus, to find the price at time t, $P(t, \tau_1, \tau_2)$, we first simulate a path of the temperature $T(s)$ for times $s \leq t$ (which means to simulate $Y(s)$, and then to

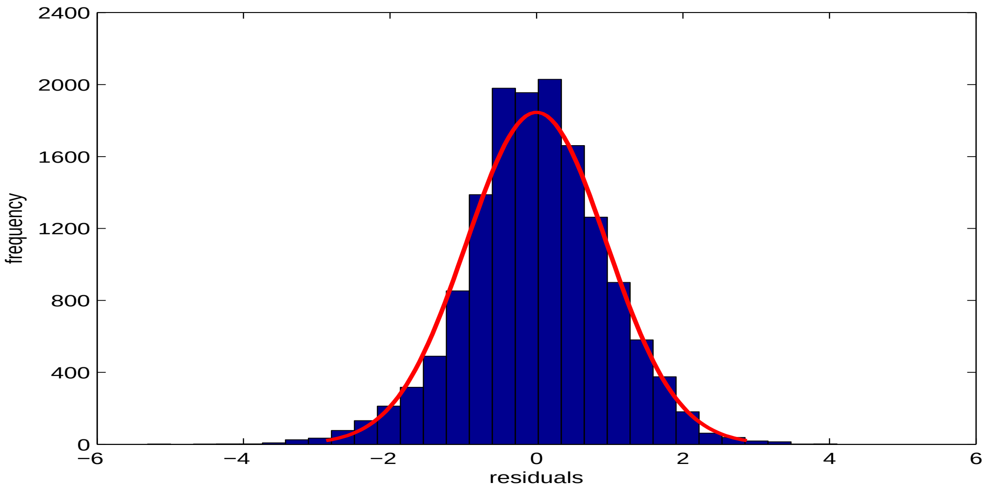

FIGURE 9. Histogram of the residuals of DATs after removing linear trend, seasonal component and AR(1).

add the seasonality function). Given this $T(t)$, we simulate N paths of $T(s)$ for $t \leq s \leq \tau_2$, and compute the index $X(\tau_1, \tau_2)$ for each path. Averaging over all the N realizations of $X(\tau_1, \tau_2)$, we obtain an estimate of $P(t, \tau_1, \tau_2)$. In this way, we have a mechanism which allows the insurance company to take current information about the weather into account, and thereby yielding a more accurate and detailed pricing technology.

Figure 11 illustrates the price evolution of the contract for January. To obtain this price path, we have conditioned on the actual observed temperatures $T(t)$ at the dates in question. Starting off the path simulations from these observed temperatures, we find the price paths which are wiggling rather than smooth curves. We started at 1 December 2010 and by simulating $n = 10,000$ paths of temperature dynamics, we obtained n indexes of $X(\tau_1, \tau_2)$. The indexes were then averaged and discounted in order to get its present value $P(t, \tau_1, \tau_2)$. This procedure was done for the following dates until 31 December 2010. We used a discounting factor $r = 0.00014$ (corresponding to 5% annual interest rate). This is at the level of the current interest rate in Malaysia. Since the price is calculated with regards to the current information about temperatures, the evolvement is no longer smoothly increasing.

We have plotted the obtained price path together with the price derived from the burn approach for comparison. We can clearly see that $P(t, \tau_1, \tau_2)$ is significantly higher than the prices obtained from the burn analysis. The price at 1 December 2010 computed by temperature modelling is higher than burn approach with a difference of 164.89. At 31 December 2010, the price obtained by temperature modelling deviates about 194.72 above burn. It is to be noted that the premia computed for the temperature approach is prone to Monte Carlo error. We have estimated this by repeating the simulation of premia 100 times for 1 December, 15 December and 31 December 2010 in order to find a numerical estimate

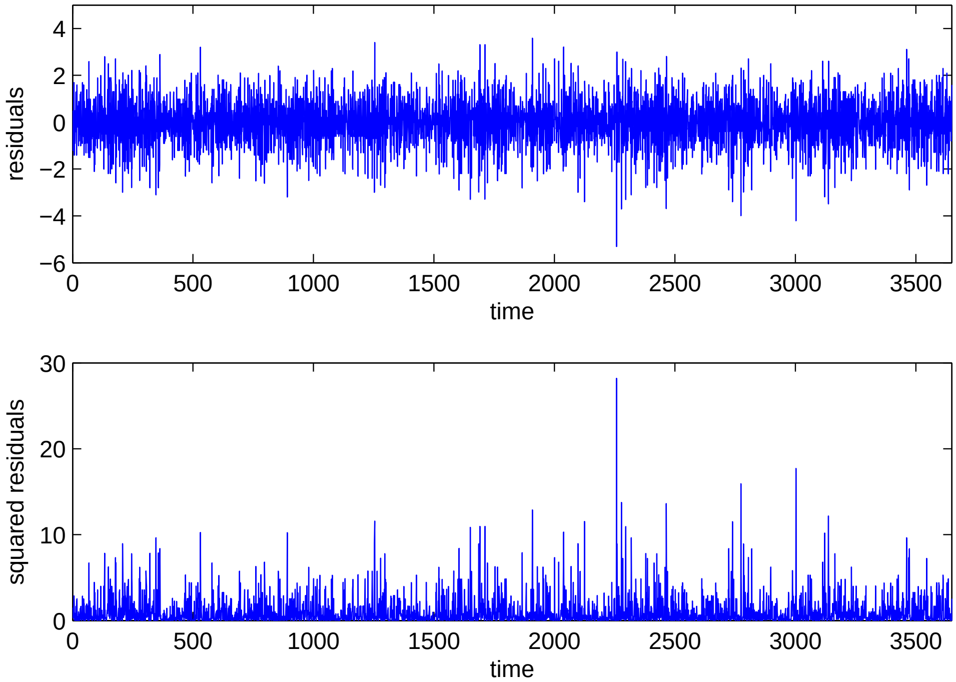

FIGURE 10. Residuals and squared residuals of DATs for the last 10 years after removing linear trend, seasonality component and AR(1).

of the confidence interval. We found very narrow confidence intervals of $[282.51, 282.77]$, $[294.68, 294.88]$ and $[324.64, 325.21]$ for the respective dates, significant at 5% confidence level.

The histogram of the claim size under temperature modelling is plotted in Figure 12 together with the empirical one obtained from burn approach. It is apparent that the claim size distribution resulting from temperature modelling has a mode, and that the exponential distribution seems to be a bad choice. One would rather imagine a lognormal distribution to be more appropriate. This is a first clear sign of the superiority of the temperature modelling approach, since it is able to reveal a much more detailed description of the claim size distribution. The histogram for the claims resulting from the burn analysis has some real big values (around 1000), and a collection of many small. The small values have too big probability compared to with the distribution from the temperature modelling approach, which results in a far lower insurance price. The temperature modelling approach produces an accurate view on the distribution of claims, and is not prone to a high degree of uncertainty. We get far better information on the tail probabilities of the claims, enabling us to get a probabilistic grip on extreme events. Due to the little data supporting the burn analysis estimates, one should put more trust into the temperature modelling.

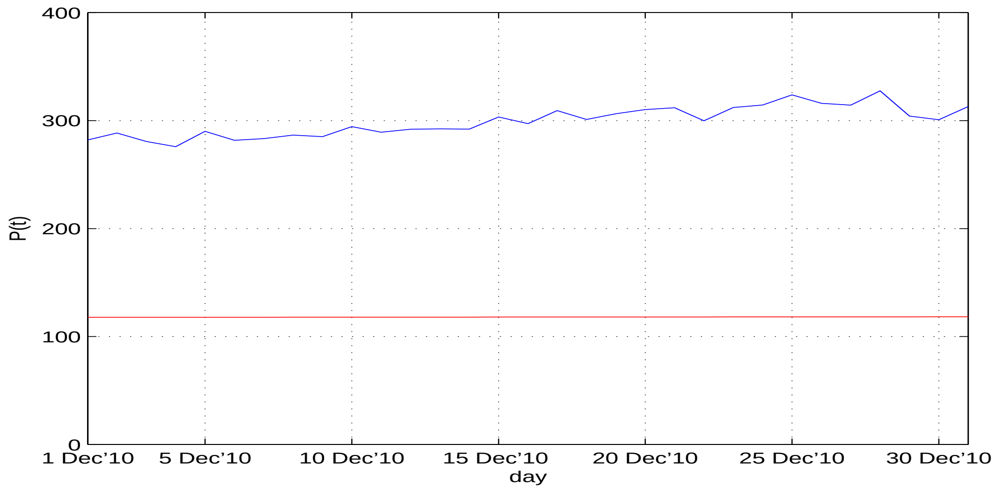

FIGURE 11. The movement of the price P for contract in January. The blue and red curve respectively represent price calculated by burn approach and temperature modelling.

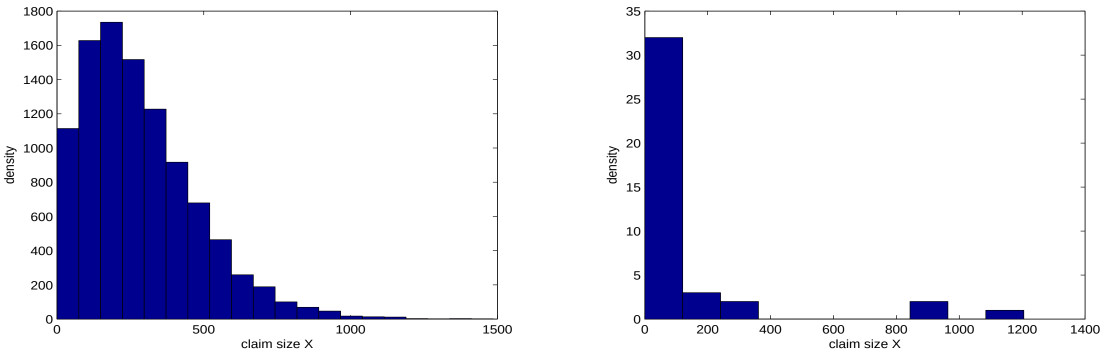

FIGURE 12. Claim size distribution from temperature modelling (left) and burn analysis (right).

3. What's in it for the farmer?

Consider a farmer who seeks to insure his crops in a period $[\tau_1, \tau_2]$ against the impact of extreme temperature levels. The decision to buy a weather index insurance will be based upon the size of the premium compared with the actual protection given by the insurance contract. In this section we analyse this protection as a function of the premium.

Based on the prices obtained in the previous section, we can imagine a farmer who has bought the insurance for a certain amount of premium. The premium is an expense that he must pay in advance to the insurer for obtaining the weather index protection. We will look at the distribution for his protection, and the probabilities of getting back money. In parallel, we also wish to find the probability of the money will exceed the premium paid.

The profit (total loss or gain) for the insured is denoted by L and defined by the difference between the total claim sizes $X(\tau_1, \tau_2)$ and the price $P(t, \tau_1, \tau_2)$ at time t. For simplicity, we suppose that $t = 0$, and recall that the premium we have to pay to purchase the insurance today will be $P(0, \tau_1, \tau_2) = \mathrm{e}^{-r\tau_2}\mathbb{E}[X]$. We can express the profit in a more convenient way by

$$L = X(\tau_1, \tau_2) - P(t, \tau_1, \tau_2).$$

We use the same definition for total loss or gain to the insurer but with opposite sign. The loss to the insured can be considered as gain to the insurer and vice versa. We will discuss loss or gain for an individual insured with a single coverage provided by a single contract. In principle, the fair premium holds

$$\mathrm{e}^{-r(\tau_2 - t)}\mathbb{E}[X(\tau_1, \tau_2) - P(t, \tau_1, \tau_2)|\mathcal{F}_t] = 0,$$

which implies that L is equal to zero in expectation. In reality, the insurance company will charge an additional risk loading on the premium, which will make the profit/loss function L have negative expectation (that is, the farmer will on average lose on the insurance contract). The reason is that the insurance contract will add a safety loading to the premium as a compensation for bearing the risk. We have designed some cases to investigate the probability distribution of L under various pricing regimes.

3.1. **Insurance calculations.** Suppose a farmer who wishes to protect his crops for adverse temperature events in the period of January 2011. He is entering a CDD-based weather index insurance contract for January 2011 ”today”, which we let be 1st of December 2010. At the end of January where the time equals τ_2, the claim for the particular month will be calculated. We use the initial price (the price on the 1st of December 2010) of $P_{burn}(\tau_1, \tau_2) = 117.75$ for burn approach and index modelling and $P_{temp}(\tau_1, \tau_2) = 282.64$ for temperature modelling. These are the ‘fair prices’ or the prices with no risk loading. With having $n = 10,000$ indexes of expected payoffs, we obtain the distribution of profit as shown in Figure 13. Clearly, the distributions have negative values meaning that there is a positive probability for loss.

The probability of loss for the farmer entering the insurance using the burn approach is

$$P\left(X(\tau_1, \tau_2) < P_{burn}(\tau_1, \tau_2)\right) = P\left(X(\tau_1, \tau_2) < 117.75\right) = 0.7958\,.$$

Hence, with an 80% chance, the farmer will receive nothing or less than what he has paid. With only approximately 20% chance he will actually receive more, meaning that if he renews his insurance contract every year, he will only in 1 out of 5 years receive more than he paid in premium that year. In the case of temperature modelling, the probability of loss becomes

$$P\left(X(\tau_1, \tau_2) < P_{temp}(\tau_1, \tau_2)\right) = P\left(X(\tau_1, \tau_2) < 282.64\right) = 0.5642\,.$$

Thus, there is only 56% chance of a loss, and approximately in 3 out of 7 years the farmer will receive a profitable income from the contract. The premium and the loss distribution are much more favourable for him. These considerations emphasize once more

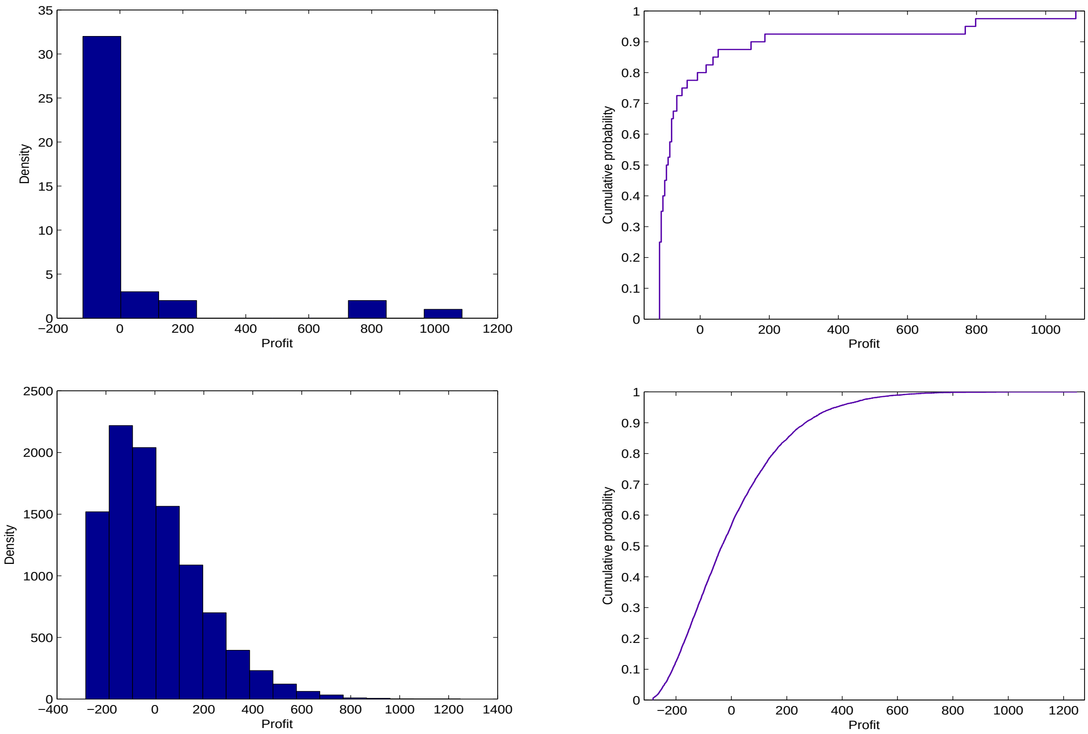

FIGURE 13. Top: Profit distribution and cumulative density for burn approach and index modelling. Bottom: Profit distribution and cumulative density for temperature modelling.

the differences in an approach resting on information from temperature series, and the burn analysis which is grounded on very little data. Of course, we need to contrast these probabilities with the actual losses incurred in order to get the true picture of the value of this insurance for the farmer.

As already indicated, an insurance company will naturally incur a risk loading to the 'fair premium' as we have calculated above. One may say that the insurance company wants to "insure themselves", and add on to the fair premium such that they can control their risk of having to pay excessive amounts in claims. A standard way to do this is to introduce a safety loading which controls a quantile of their loss distribution. This resembles the value-at-risk concept in finance, and entails in putting a premium on the insurance contract such that their loss distribution from claims is within an acceptable probability.

To be more specific, we find the value-at-risk at a given significance level of the loss distribution for the insurance company under 'fair premium' of the contract. Given this level, we suppose that the insurance company charge a risk loading resembling a certain fraction of this level. Letting the significance level be 5%, we search a level γ such that

$$P(\text{Claim} \geq \gamma) = 0.05\,.$$

The level γ will only be exceeded with 5% probability. For burn approach, we obtain $\gamma_{burn} = 900.00$ (the same value for index modelling) while for temperature modelling profit distribution, γ_{temp} is equal to 660.77. The insurance company will now add 5% of the estimated γ as risk loading to the fair price (risk loaded premium). With this new price, the probabilities of loss and gain for the farmer are presented in Table 2.

Table 2. Probability of loss and gain for 5% of risk loading

Method	γ	Price $P(\tau_1, \tau_2)$	$P(X < P(\tau_1, \tau_2))$	$P(X > P(\tau_1, \tau_2))$
Burn analysis	900.00	162.75	0.8478	0.1510
Temperature	660.77	315.68	0.6253	0.3729

A risk loaded premium increases the probability of loss for the farmer, naturally. With the burn approach, the farmer will only receive a gain from the contract with 15% probability, or approximately once every 7 year. With the temperature modelling approach, the farmer receives a profit with 37% probability, which is slightly less than with the 'fair premium'. It is interesting to note that although the premium is significantly bigger with the temperature modelling approach, the contract seems more attractive than if we solely base our assessment on the burn analysis.

4. Conclusion

We have analysed three different pricing approaches for weather index insurance contracts in this paper. Weather index insurance has gained some attention in recent years as a way for farmers, say, in developing coutries to protect their crop. We have focused on contracts settled on temperature indexes in given months, using data collected in Malaysia as our empirical case, and we analyse cooling-degree indexes, which measures excessive temperatures which may harm the crop of a farmer.

As the claims in such weather index contracts are settled on a measurable objective index, and not on actual losses incurred, one may view the insurance contracts as weather derivatives. Motivated by theory on weather derivatives, we have considered the three pricing approaches *burn analysis*, *index modelling* and *temperature modelling*. The first two give similar results, and are based directly on computing historical values of the index in question. We argue that such an approach will rest on very thin data material. The temperature modeling approach, on the other hand, is based on the time series properties of temperature, from which one can compute the index. Usually, one has very rich sets of data for temperature.

We fitted a simple autoregressive time series model with seasonality for Malaysian daily average temperatures, based on 40 years of data. This gives us a very precise information on the dynamics of temperature, and hence, a very detailed description of the temperature index in a given period. In fact, simulating from the model, we can obtain a very detailed probabilistic information on the index, and therefore assess correctly the premium and probabilities of loss and profits.

The burn analysis and index modelling approaches have a high degree of uncertainty in their premium estimates. The premium estimated from the temperature modelling approach is prone to Monte Carlo error, on the other hand. Controlling for this, we find big differences in premia between the approaches. The temperature modelling approach has the additional advantage that it can account for current information of the weather situation, while this is not the case for the two other approaches.

Finally, we analysed the chances of receiving a profit from a weather index insurance contract. The chances are not very high using the burn analysis approach, but far better in the temperature modelling case. This is obviously a result of the very poor data set backing up the results in the burn analysis method. The temperature modelling methodology provides better foundation for making assessments on the probabilities of loss and profit, and for our case the insurance becomes more advantageous for the farmer despite the much higher premium.

As is standard in an insurance contract, the insured must pay a premium upfront in order to obtain a protection. Since farmers in developing countries are rather poor, this may seem like an unfair deal, since information and knowledge of these contracts naturally would be biased towards the insurance company. Since the weather index insurance contracts are closely related to temperature futures traded on the Chicago Mercantile Exchange (CME), the insurance companies may in principle hedge their risk. For the time being, no temperature futures are traded for cities in developing countries. However, one may imagine places where one could partially hedge the risk using other locations (see Barth et al. [3] for spatial hedging of temperature derivatives). This would reduce the risk for the insurance company, and thus the premium, and making such products more attractive. But still, one is left to wonder whether it is better for the farmers to actually save the same amount as the premium in a bank account, controlling the money themselves, rather than entering into swap deals as is the case here.

References

[1] Barnett, B.J., Barrett, C. B., and Skees, J. R. (2008). Poverty traps and index based risk transfer products. *World Development*, **36**, 1766–1785.

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

[3] Barth, A., Benth, F. E., and Potthoff, J. (2011). Hedging of spatial temperature risk with market-traded futures. *Applied Mathematical Finance*, **18:2**, 93–117.

[4] Benth, F. E., Šaltytė Benth J. and Koekebakker S. (2007). Putting a price on temperature. *Scandinavian Journal of Statistics*, **34**, 74–767.

[5] Benth, F. E., Šaltytė Benth J. and Koekebakker S. (2008). *Stochastic modelling of electricity and related markets*. World Scientific.

[6] Benth, F. E. and Šaltytė Benth J. (2011). Weather derivatives and stochastic modelling of temperature. *International Journal of Stochastic Analysis*, Vol. 2011, 21 pages. (electronic publication)

[7] Campbell, S. D. and Diebold, F. X. (2005). Weather forecasting for weather derivatives. *Journal of American Statistical Association*, **100**(469), 6–16.

[8] Clewlow, L. and Strickland, C. (2000). *Energy Derivatives. Pricing and Risk Management*. Lacima Publishers.

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

[10] Jewson, S. and Brix, A. (2005). *Weather Derivative Valuation: The Meteorological, Statistical, Financial and Mathematical Foundations*. Cambridge University Press.

[11] Kliger, D., and Levikson, B. (1998). Pricing insurance contracts - an economic viewpoint. *Insurance: Mathematics and Economics*, **22**, 243–249.

[12] Prime Minister's Department (2006). Ninth Malaysian Plan 2006–2010. The Economic Planning Unit: Author. Retrieved from http://www.epu.gov.my/html/themes/epu/html/rm9/html/english.htm.

[13] Reiss, R.D., and Thomas, M. (2007). *Statistical Analysis of Extreme Values.* Birkhäuser, Basel.

[14] Sakurai, T., and Reardon, T. (1997). Potential demand for drought insurance in Burkina Faso and its determinants. *American Journal of Agricultural Economics*, **79**, 1193–1207.

[15] Šaltytė Benth J., Benth, F. E. and Jalinskas P. (2007). A spatial-temporal model for temperature with seasonal variance. *Journal of Applied Statistics*, **34**, 823–841.

[16] Skees, J.R. (1999). Opportunities for improved efficiency in risk sharing using capital markets. *American Journal of Agricultural Economics*, **81**, 1228–1233.

[17] Skees, J.R. (2002). Examining the feasability of livestock insurance in Mongolia. *The World Bank Policy Research Working Paper 2886.*

[18] Skees, J.R. (2008). Innovations in index insurance for the poor in lower income countries. *Agricultural and Resource Economics Review*, **37/1**, 1–15.

[19] Turvey, C.G. (2001). Weather derivatives for specific event risks in agriculture. *Review of Agricultural Economics*, **23**, 333–351.

Che Mohd Imran Che Taib: Center of Mathematics for Applications, University of Oslo, PO Box 1053 Blindern, N-0316 Oslo, Norway, and, Universiti Malaysia Terengganu, 21030 Kuala Terengganu, Terengganu, Malaysia.

E-mail address: imran@umt.edu.my

Fred Espen Benth: Center of Mathematics for Applications, University of Oslo, PO Box 1053 Blindern, N-0316 Oslo, Norway, and, University of Agder, School of Management, Serviceboks 422, N-4604 Kristiansand, Norway.

E-mail address: fredb@math.uio.no

URL: http://folk.uio.no/fredb/