# Extracting Risk Neutral Densities For Weather Derivatives Pricing Using The Maximum Entropy Method

Antonis K. Alexandridis[1*], Henryk Gzyl[2], Enrique ter Horst[3] and German Molina[4]

[1]Kent Business School, University of Kent, U.K.
A.Alexandridis@kent.ac.uk
[*]corresponding author

[2]Center for Finance, Instituto de Estudios Superiores de Administración, Venezuela
henryk.gzyl@iesa.edu.ve

[3]School of Business, Colegio de Estudios Superiores de Administración, Colombia
enriqueterhorst@gmail.com

[4]Idalion Capital Group, U.S.A.
german@alumni.duke.edu

**Abstract**

In this study we propose the use of the maximum entropy method to extract the risk neutral probabilities directly from the weather market prices. The proposed methodology is computationally fast, model free, non-parametric and can overcome the data sparsity problem that governs the weather market. We infer consistent risk neutral probabilities along with their densities from the market price of temperature options. The risk neutral probabilities inferred from a smaller subset of the data are consistent in the sense that they reproduce the other prices and can be used to value accurately all other possible derivatives in the market sharing the same underlying asset. We examine two sources of the out-of-sample valuation error. First, we use different sets of possible physical state probabilities that correspond to different levels of expertise of the trader. Then, we apply our methodology under three scenarios where the available information in the market is based on historical data, meteorological forecasts or both. Our results indicate that different levels of expertise can affect the accuracy of the valuation. When there is a mix of information, non-coherent sets of prices are observed in the market.

**Keywords**: Finance, weather derivatives pricing, temperature options, maximum entropy

# 1 Introduction

Weather fluctuations affect the economy both directly and indirectly. Even minor weather changes often have significant impact on the day-to-day operations and revenues of many businesses in sectors such as agriculture, energy, tourism, transportation and construction, [7, 18, 32].

A dry period could destroy farmers crops, and warm winters could cost millions to energy companies, due to reduced energy consumption for heating. In autumn 2014, the temperature in the UK was higher than normal by just 1.5°C. It was reported that this seemingly minor change led to financial losses of £700m in September and October alone, and due solely to reduced demand for winter clothing[1]. The insurance group RSA estimated that the problems caused by the snow in 2013 cost the country's economy about £470m a day, due to the massive travel disruption[2]. Similarly, local councils across the UK have overspent on their budgets by up to 100% in the recent cold years[3].

Due to the recent economic crisis and the increased weather volatility caused by climate change, the need for efficient and effective weather risk management is evident, [44]. Advances in meteorology and weather forecasting have created the opportunity for researchers in operations research (OR) and management science to create valuable tools for effective weather risk management, [42]. Such examples can be found in [11, 22, 24, 38, 39, 44]. Retailers have used weather-linked promotions, such as weather rebates, to protect against adverse financial outcomes due to unfavourable weather, [14, 19]. However, the outcome of this strategy is unknown and can lead to significantly volatile returns.

The necessity to hedge adverse weather effects and unseasonal weather resulted in the creation of a new class of financial assets called weather derivatives (WD). In general, WD are designed to cover non-catastrophic weather events, i.e. high probability, low risk events. Non-catastrophic weather risk is gaining importance as climate change becomes more pronounced, [44]. Weather derivatives were developed to hedge volume or quantity risk, rather than the price risk, [10]. The pay-off of a WD depends on the measurement

[1]http://www.telegraph.co.uk/finance/newsbysector/retailandconsumer/11711835/Wet-British-weather-has-devastatingimpact-on-retail-sales.html. Last accessed: 10 February 2017.

[2]http://www.bbc.com/news/business-21121219.

[3]http://www.theguardian.com/business/2012/jan/27/winter-weather-councils-derivatives-gamble.

of the underlying weather index. Clearly, weather derivatives can provide superior hedging opportunities. The majority of WD are written on temperature, and more precisely on the Heating Degree Day Index (HDD)[4]. Hence, in this study we focus on temperature derivatives. The weather market is a fast developing market. According to the Weather Risk Management Association the market grew by 20% in 2010-2011, to a total notional value of $11.8 billion, and further understanding of the uncertainty about weather trends or increased variability will potentially enhance even further its expansion.

The weather derivatives market is a classic incomplete market, because the underlying weather variables have no value, cannot be traded or stored. Hence, the replicating portfolio cannot be constructed and the classical Black-Scholes pricing approach cannot be directly applied. In addition, the payoffs of weather derivatives are determined by indices that are average quantities, [25]. Despite the increasing number of studies in weather derivatives pricing [2, 3, 5, 6, 16, 21, 48] the market still lacks a general accepted pricing framework. Furthermore, the market is characterised by a lack of liquidity, which increases its level of incompleteness.

An agent in this market is interested in determining the risk neutral measure used by traders for option valuation. These options are European calls and puts written on futures on the HDD index. In addition, the agent may be interested in knowing whether all options are priced consistently, at least by the same trader.

In order to derive the option prices, often stochastic differential equations are used to describe the dynamics of the daily average temperatures (DAT), [2–6, 45, 48]. The continuous processes used for modelling DAT usually take a mean-reverting form, which has to be discretised in order to estimate its various parameters. Once the process is estimated, any contingent claim (futures and options) can be valued by taking expectation of the discounted future payoff. Alternatively, regime switching, [23], Autoregressive Moving Average, [15, 29] and Neural Networks, [17, 48], have been proposed in the literature to estimate the temperature process.

However, modelling the DAT is not a straightforward process and strong assumptions about the temperature model and the noise-generating process are made. Also, agents face

[4]The HDD index is described in section 2.

the issue of model risk, since the estimation of the market price of risk depends on the assumed model. Small misspecifications in the DAT model can lead to large mispricing errors, [48]. Furthermore, it is usually difficult to solve the stochastic differential equation in order to price the financial weather products, and often it is impossible to find closed-form solutions for the pricing equations. Additionally, one is faced with the problem of illiquidity and data sparsity that characterises this market. Finally, due to data unavailability, previous studies model the underlying temperature process but do not proceed on testing the accuracy of the forecasted contract prices against real market data.

In this study, we propose an alternative methodology where we extract the risk neutral probabilities directly from the market prices of temperature options, which avoids all the aforementioned drawbacks. To our knowledge only in [33] a similar approach has been followed for weather derivatives. More precisely, in [33] the Bayesian quadrature method has been applied in order to derive the state price densities. In this paper we follow a different approach. Our proposed methodology is based on the maximum entropy, [8, 34, 49]. This is a very powerful technique which needs only few market data points in order to extract accurately the risk neutral probabilities, and can be used for density reconstruction, [9], and for the valuation of the prices of other option contracts traded in the market. Hence, the maximum entropy approach naturally overcomes the problem of data sparsity. The advantage of extracting the risk neutral probabilities directly from the market prices is that volatility and other moments can easily be calculated independently of any particular model, [33]. The concept of entropy has been used with great success in various finance related applications, [1, 20, 26–28, 30, 36, 41, 43]. In this study we extend the framework proposed in [31] to incomplete markets.

The main contributions of our approach are the following: we present a model-free, non-parametric approach to determine numerically the state prices, or equivalently the risk neutral probabilities, for any prior physical probabilities. Hence, an advantage of the method is that we do not have to calibrate for model parameters. In addition, the maximum entropy based procedure allows us to determine which data points (call or put contracts) are more informative. To numerically implement the maxentropic methodology we propose a market driven, systematic and intuitive discretisation procedure, in which any prior information consisting of the physical probabilities of the market is integrated in a natural way in the methodology. Furthermore, we decompose the out-of-sample valuation error in two components. The first one shows the proportion of the error that arises from the level of available information in the market. We capture this using three levels of information: (1) derived by a naive agent using only historical data; (2) derived by a semi-informed agent, using meteorological forecasts available for few days but not up to the maturity day; and (3) derived by an informed agent, using meteorological forecasts available up to the maturity day. The second component shows the proportion of error that arises from the level of expertise of the trader. We capture this by using prior physical probabilities derived: by (1) a naive; (2) a semi-advanced; and (3) a state of the art model for the underlying temperature process.

Our results show that the proposed method can extract the risk neutral probabilities overcoming the data sparsity problem, and can be used for accurate valuation of option contracts. The contracts used for fitting, the level of information in the market, as well as the prior probabilities, can play an important role in the accurate valuation.

Due to the attractive characteristics of weather derivatives, it is clear that their use can have significant managerial implications to a firm. As it is demonstrated in [39] the use of weather derivatives can lead to higher firm value, investments, and leverage. In [13] the importance of production planning under seasonal demand is highlighted. Similarly, weather derivatives can be used to create profitable investment portfolios due to diversification purposes, [3, 29, 35]. In [12] it is shown that weather derivatives have the potential to substantially affect and alter farm plans. Furthermore, as it is discussed in [40] weather derivatives can alter the actions of government agencies even in the absence of explicit incentives and monitoring mechanisms.

Weather derivatives can prove a valuable tool in the decision making process. As a result, the optimal utilisation of weather derivatives by risk managers can provide a competitive advantage in the marketplace and increased profits, [19]. Hence, an accurate tool for valuation of weather derivatives and for identification of inconsistent set of market prices is essential for efficient and effective weather risk management. The framework presented in this study might be of interest to government bodies and businesses related to agriculture, energy, tourism, transportation and construction, whose revenues are affected by the weather as well as to investors in the weather market.

The remainder of the paper is organized as follows. Section 2 briefly presents the weather market. Section 3 is devoted to methodological aspects. First, in section 3.1 we describe how the market model is built, while in section 3.2 the maximum entropy method is presented. Lastly, in section 3.3 the boundedness of the dual entropy function is discussed, where we explain the boundaries for the method. Section 4 is devoted to the numerical implementation of the maxentropic procedure. The data are described in section 4.1, while the numerical experiments performed are described in section 4.2. Our empirical results are discussed in section 4.3. Finally, in section 5 we present our concluding remarks.

## 2 The Weather Market

Temperature derivatives are settled in three main temperature indices: the HDDs, the Cooling Degree Days (CDDs) and the Cumulative Average Temperature (CAT).

The CAT index is the sum of the DATs over the contract period. HDD is the number of degrees by which the daily temperature is below a base temperature, and CDD is the number of degrees by which the daily temperature is above the base temperature. The base temperature is usually 65 degrees Fahrenheit (or 18$^{o}$C). HDDs and CDDs are accumulated over a period, usually over a month or a season. The value of the three indices for a measurement period in the time interval $[\tau_1, \tau_2]$ is given by the following expressions:

$$CAT(\tau_1, \tau_2) = \int_{\tau_1}^{\tau_2} T(s)ds \tag{2.1}$$

$$HDD(\tau_1, \tau_2) = \int_{\tau_1}^{\tau_2} \max\big(65 - T(s), 0\big)ds \tag{2.2}$$

$$CDD(\tau_1, \tau_2) = \int_{\tau_1}^{\tau_2} \max\big(T(s) - 65, 0\big)ds \tag{2.3}$$

where the DAT, $T(t)$, is the average of the daily maximum and minimum temperature, $T(t) = \big(T_t^{max} + T_t^{min}\big)/2$.

A trader is interested in finding the price of a temperature contract written on a specific temperature index. The price of a futures contract written in a temperature index under the risk neutral probability $Q$ and the filtration $\mathcal{F}_t$ at time $t \leq \tau_1 < \tau_2$ is

$$e^{-r(T-t)}\mathbb{E}_Q\Big[Index - F_{Index}(t, \tau_1, \tau_2) \mid \mathcal{F}_t\Big] = 0$$

where $Index$ is the CAT, HDD or CDD; $F_{Index}$ is the price of a futures contract written on the specific index; $r$ is the risk-free interest rate; and $\mathcal{F}_t$ is the history of the process until time $t$. Since $F_{Index}$ is $\mathcal{F}_t$-adapted, we derive the price of the futures contract to be

$$F_{Index}(t, \tau_1, \tau_2) = \mathbb{E}_Q\Big[Index \mid \mathcal{F}_t\Big] \tag{2.4}$$

Consequently, the European temperature call option price written on the futures price with a strike price $K$ is defined as:

$$C(K) = e^{-r(T-t)} \int \max\big(F_{Index}(t, \tau_1, \tau_2) - K, 0\big) f(x)dx \tag{2.5}$$

The majority of WDs are written on a temperature and specifically on the HDD index. In this study we focus only on this class of WDs, although our methodology can be easily adapted and applied in any WD.

# 3 Methodology

## 3.1 The market model

In this study we build on the methodology proposed in [31] and we extend it to incomplete markets. More precisely we focus on temperature derivatives written on the HDD index. The first step is to discretise the temperature index into $K$ non-overlapping intervals:

$$[X_0, X_1), \ldots, [X_{K-2}, X_{K-1}), [X_{K-1}, X_K)$$

There will be a market state associated to each temperature interval. The values $X_0$ and $X_K$ can be chosen using historical data, and we assume that the probability of the temperature index to be greater than $X_K$ or lower than $X_0$ is zero.

Second, we need to estimate the reference (physical) probabilities

$$p_{K-j+1} = P(X_{j-1} \leq X < X_j), \quad j = 1, 2, \ldots, K. \tag{3.1}$$

where $X$ is the temperature index at the time horizon of a chosen model. The probabilities can be computed using any preferred model. Next, we choose appropriate "levels" $\hat{X}_j$ such that $\hat{X}_j$ occurs when the actual temperature index is $X_{j-1} \leq X < X_j$. The simplest choice is to take the mid point of the temperature range. As the temperature ranges are homothetically related to the HDD or the CDD, this will amount to choosing the mid point of the range as the value of the asset in the corresponding market state.

Fixing "today" as $t = 0$, we consider all the options (written on futures) that have the same maturity day, $t = T$. We consider the mid points of the intervals as the possible outcomes of the HDD index for the measurement period $[\tau_1, \tau_2]$, and denote as $\hat{S}_j$ the possible values of the basic assets in the market.

Table 1 describes the characteristics of the data. The first row lists the market states. In the second row the asset level that characterizes each state is presented, while in the third row we list the corresponding ranges. Finally, in the last row the physical probabilities are listed. As shown in Section 4, we use three different sets of physical probabilities indicating different levels of information.

Table 1: Market states for the discrete model

| State | $\omega_1$ | $\omega_2$ | $\ldots$ | $\omega_K$ |
|---|---|---|---|---|
| Level | $\hat{S}_1$ | $\hat{S}_2$ | $\ldots$ | $\hat{S}_K$ |
| Range | $[S_0, S_1)$ | $[S_1, S_2)$ | $\ldots$ | $[S_{K-1}, S_K)$ |
| Phys. Prob. | $p_1$ | $p_2$ | $\ldots$ | $p_k$ |

In this market there exist both European call and put options, where the underlying asset is the price of a future on the HDD temperature index. Our aim is to determine the risk neutral probabilities that are used by traders to value these options. For this purpose, the maximum entropy method will be used.

Suppose for the sake of definiteness we consider $M < K$ European call options with strike prices $K_m = S_{j_m}$, where $\{j_1, \ldots, j_M\} \subset \{1, \ldots, K\}$.

Thus, the problem that we want to solve consists of determining the probabilities $\{q_j, | j = 1, \ldots, K\}$ such that

$$\pi_m = e^{-rT} \sum_{j=1}^{K} q_j O\Big(\hat{S}_j, K_m\Big) \quad m = 1, \ldots, M \tag{3.2}$$

where $M$ is the number of option prices used, $O\left(\hat{S}_j, K_m\right)$ is the payoff of the $m-$th option (which will be either a call or a put of European type), and $\pi_m$ is its observed price.

### 3.2 The maximum entropy method

In this section we describe the maximum entropy method that we use in order to solve (3.2). First, we rewrite (3.2) as

$$\pi_m = \sum_{j=1}^{K} \rho_j O_m\left(\hat{S}_j, K_m\right) p_j, \quad m = 1, \ldots, M \tag{3.3}$$

where $\rho_j$ is the density of the risk neutral probability $\boldsymbol{q}$ with respect to the probability $\boldsymbol{p}$. Note, that the discount factor in (3.2) is dropped in (3.3) for simplicity. We can assume that the discount factor has been made part of the option prices by replacing $\pi_m$ with $e^{rT}\pi_m$.

The constraint for $\rho_i$, $i = 1, ..., K$ to be a density is:

$$\sum_{i=1}^{K} \rho_i p_i = 1 \tag{3.4}$$

Consider the set $\mathcal{D} = \{\rho_i : i = 1, \ldots, K\}$ satisfying (3.3) and (3.4). Then, $\mathcal{D}$ is a closed and convex (in $\mathbb{R}^K$) set. We can then select a point from within that set by solving the following optimization problem:

Define the "entropy" function: $\mathcal{D} \to \mathbb{R}$ as

$$H(\boldsymbol{\rho}) = -\sum_{j=1}^{K} \rho_j \ln \rho_j p_j. \tag{3.5}$$

It is easy to verify that this function is continuous and strongly concave on $\mathcal{D}$. When a value of $\rho_j$ satisfying the constraints (3.3) and (3.4) exists, the entropy maximization yields the standard representation:

$$\rho_j^* = \frac{1}{Z_0(\boldsymbol{\lambda}^*)} \exp\left(-\sum_{m=1}^{M} \lambda_m^* O_m\left(\hat{S}_j, K_m\right)\right). \tag{3.6}$$

The normalization factor is given by the (log-convex) function

$$Z(\boldsymbol{\lambda}) = \sum_{j=1}^{K} \exp\left(-\sum_{m=1}^{M} \lambda_m O_m\left(\hat{S}_j, K_m\right)\right) p_j$$

defined on $\mathbb{R}^M$ and the vector $\boldsymbol{\lambda}^*$ is obtained by minimizing the convex dual entropy function

$$\Sigma(\boldsymbol{\lambda}, \boldsymbol{\pi}) = \ln Z(\boldsymbol{\lambda}) + \sum_{m=1}^{M} \lambda_m \pi_m. \tag{3.7}$$

We explain how this comes about at the end of the section. It is at this point that we can assert whether a solution to the entropy maximization problem exists. There are two alternatives:

1. The function $\Sigma(\boldsymbol{\lambda}, \boldsymbol{\pi})$ is bounded below and a minimum is reached at $\boldsymbol{\lambda}^* \in \mathbb{R}^M$

2. The function $\Sigma(\boldsymbol{\lambda}, \boldsymbol{\pi})$ is not bounded below[5].

The second case depends on whether there exits a $\boldsymbol{\lambda}$ such that $\sum_{m=1}^{M} \lambda_m O_m \left(\hat{S}_j, K_m\right) > \sum_{m=1}^{M} \lambda_m \pi_M$ for every $j = 1, \ldots, K$. Then, moving along this $\boldsymbol{\lambda}$ it is easy to verify that $\Sigma(g\boldsymbol{\lambda}, \boldsymbol{\pi}) \to -\infty$ as $g \to \infty$.

In the case that $\Sigma(\boldsymbol{\lambda}, \boldsymbol{\pi})$ is bounded below the minimizer $\boldsymbol{\lambda}^*$ can be obtained. Then equation (3.6) can be used for the computation of any derivative with pay-off $f(\hat{S}_j)$ at time $T$ by computing the expected values

$$\pi(f) = \sum_{j=1}^{K} \rho_j^* f(\hat{S}_j) p_j. \tag{3.8}$$

### 3.2.1 The duality argument

To understand where (3.6) comes from and the connection between the maximization of $H(\boldsymbol{\rho})$ and the minimization of (3.7), we begin by observing that for any two densities $\boldsymbol{\rho}(1)$ and $\boldsymbol{\rho}(2)$, an application of Jensen's inequality yields

$$\sum_{i=1}^{K} \rho_i(1) \ln \left( \frac{\rho_i(1)}{\rho_i(2)} \right) p_i \geq 0,$$

with equality occurring whenever $\boldsymbol{\rho}(1) = \boldsymbol{\rho}(2)$. Now, let $\boldsymbol{\rho}(1) = \boldsymbol{\rho}$ satisfy the constraints (3.3)-(3.4), and for any $\boldsymbol{\lambda} \in \mathbb{R}^M$, let $\boldsymbol{\rho}(\boldsymbol{\lambda})$ be given by (3.6). An application of the inequality yields

$$\sum_{i=1}^{K} \rho_i \ln \left( \frac{\rho_i}{\rho_i(\boldsymbol{\lambda})} \right) p_i = -H(\boldsymbol{\rho}) + \Sigma(\boldsymbol{\lambda}, \boldsymbol{\pi}) \geq 0.$$

[5] Detailed proof can be found on Section 2.2 of [8]

To finish, note that first order condition on $\boldsymbol{\lambda}^*$ to be a minimizer of $\Sigma(\boldsymbol{\lambda}, \boldsymbol{\pi})$ amounts to say that $\rho(\boldsymbol{\lambda}^*)$ satisfies the constraints, and therefore to maximize the entropy. Furthermore $H(\boldsymbol{\rho}(\boldsymbol{\lambda}^*)) = \Sigma(\boldsymbol{\lambda}^*, \boldsymbol{\pi})$.

### 3.3 On the boundedness of the dual entropy function

In this section we present more analytically the case where $\Sigma(\boldsymbol{\lambda}, \boldsymbol{\pi})$ can be unbounded below. To understand this, first we write the dual entropy function as

$$\Sigma(\boldsymbol{\lambda}, \boldsymbol{\pi}) = \ln\left(\sum_{j=1}^{K} e^{-\langle \boldsymbol{\lambda}, (\boldsymbol{O}\boldsymbol{e}_j - \boldsymbol{\pi})\rangle}\right)$$

where $\boldsymbol{e}_j$ are the standard basis vectors for $\mathbb{R}^K$, and $\boldsymbol{O}$ is the payoff matrix where $O_{m,j}$ denotes the payoff of the $m-$th option in the $j-$th state.

**Lemma 3.1** *If there exists a $\boldsymbol{\lambda}_c \in \mathbb{R}^M$ such that $\langle \boldsymbol{\lambda}_c, (\boldsymbol{O}\boldsymbol{e}_j - \boldsymbol{\pi})\rangle > 0$ for all $j = 1, \ldots, K$, then the dual entropy function is unbounded below.*

*Proof* Suppose that such a $\boldsymbol{\lambda}_c$ exists and let $m = \min\{\langle \boldsymbol{\lambda}_c, (\boldsymbol{O}\boldsymbol{e}_j \boldsymbol{\pi})\rangle \,|\, j = 1, \ldots, K\} > 0$. Then for $g > 0$, clearly

$$\Sigma(g\boldsymbol{\lambda}_c, \boldsymbol{\pi}) \leq \ln\left(Ke^{-gm}\right) = \ln K - gm$$

which clearly tends to $-\infty$ as $g \to \infty$. Therefore, along $\lambda_c$ the dual entropy is not bounded below and the minimum of $\Sigma(\boldsymbol{\lambda}, \boldsymbol{\pi})$ on $\mathbb{R}^M$ does not exist. □

The following remarks are related to another technical issue. If we think of $\boldsymbol{O}$ as (the matrix of) a linear mapping $\boldsymbol{O} : \mathbb{R}^K \to \mathbb{R}^M$, and any probability $\boldsymbol{p}$ as a point in the simplex $\mathbb{P} := \{\boldsymbol{p} \in \mathbb{R}^K | \sum p_j = 1\}$, then the possible price vectors are in the points in $\boldsymbol{O}(\mathbb{P})$. If Lemma 3.1 holds, then $\boldsymbol{\pi} \notin \boldsymbol{O}(\mathbb{P})$. This is because $\boldsymbol{\pi}$ and $\boldsymbol{O}(\mathbb{P})$ are separated by the hyperplane defined by $\boldsymbol{\lambda}_c$, as indicated in the Lemma. In this case we can say that some prices have not been consistently assigned.

## 4 Empirical Analysis

This section is separated into two parts. In the first we describe the data collected, and in the second we present the results of applying the maxentropic technique.

## 4.1 The market data

We use CME options and futures daily closing prices, written on the temperature HDD index as measured in the meteorological station in New York. We focus on New York since it is the biggest temperature market with volume around 20% of the total market volume. The data ranges from 6/3/2006 to 30/11/2010.

For this study we are interested in examining whether our maximum entropy approach can extract accurately the risk neutral probabilities and whether we can use these probabilities in order to value new (unobserved) contracts. Additionally, we want to examine the effectiveness of the method depending on the level of the information in the market. More precisely, we want to examine whether the trader's level of information has an impact on the extraction of the risk neutral probabilities.

There are two ways in which the level of information can affect the pricing. First, meteorological forecasts can be used in order to approximate the temperature index. Second, when temperature derivatives are traded before the measurement period, and meteorological forecasts are not available, advanced statistical models can be used for forecasting the DAT and hence the HDD index. We assume different scenarios in order to examine the uninformed, semi-informed and informed cases.

First, since meteorological forecasts are not accurate for more than 10 days, [3, 46, 47], we examine the performance of the maximum entropy method with 3 different remaining times to maturity. More precisely, we apply our method: 1) before the contract's measurement period; 2) at the beginning of the measurement period; and 3) in the middle of the measurement period. In the first case, accurate meteorological forecasts are not available for the traders. Therefore their decisions are based solely on historical data (or seasonal forecasts). In the second case, meteorological forecasts for few days are available (first 5-10 days of the contract's measurement period). In the last case, meteorological forecasts are available until the maturity day.

Second, we assume three different scenarios in the computation of the physical probabilities described in section 3.1. The first one corresponds to the completely uninformed trader, where all market states have an equal probability to occur. The second set of probabilities corresponds to simple information arriving by the Historical Burn Analysis (HBA) method.

In other words studying the possible HDD outcomes in the previous 30 years we compute the corresponding probabilities for each market state[6]. Finally, the last set of probabilities corresponds to the informed trader. In this case a more advanced model is used. More precisely, a CAR($p$) model proposed by [6] is estimated, and then it is used to forecast the DAT and the HDD index for the corresponding period of interest.

The weather market is characterised by data sparsity and lack of liquidity. We want to examine the number of contracts that are needed in order to extract accurately the implied risk neutral probabilities. For each day and dataset, we consider all the contracts with different strike prices but with the same maturity. We consider 8 data sets for our analysis, since when the number of strike prices and the set of probabilities are considered the total number of scenarios that we analyse grows exponentially. More precisely, as it is presented in the next section, for these 8 datasets we examine 3,048 different scenarios.

In Table 2 the 8 data sets are presented. This includes the day $t$ where our methodology is applied, the maturity date, and the available number of calls and puts are given.

Table 2: Details of the 8 datasets

| Dataset | $t$ | Maturity | Number of Calls | Number of Puts |
|---|---|---|---|---|
| 1 | 20/11/2007 | 31/12/2007 | 3 | 3 |
| 2 | 03/12/2007 | 31/12/2007 | 3 | 3 |
| 3 | 14/12/2007 | 31/12/2007 | 3 | 3 |
| 4 | 21/10/2010 | 30/11/2010 | 4 | 2 |
| 5 | 01/11/2010 | 30/11/2010 | 4 | 2 |
| 6 | 15/11/2010 | 30/11/2010 | 4 | 2 |
| 7 | 06/03/2006 | 03/04/2006 | 4 | 3 |
| 8 | 14/03/2006 | 03/04/2006 | 6 | 3 |

## 4.2 The market states

In this section we will describe the market model for the aforedescribed datasets. More precisely, we will present a description of the underlying market model. As described in Table 1, first we are interested in the number of states, along with the range of values of the HDDs, and the level that characterises each state. Second we need a set of physical probabilities for each state. As explained in the previous section, we will use three different

[6]For an analytical description of the HBA we refer to [3].

sets of probabilities representing a naive, a semi-informed and an informed trader. Third, we need a collection of European options specifying their types, strikes and prices. Due to space limitation we present only two cases chosen because they exhibit typical behaviour: using 3 or 4 options we were able to obtain the risk neutral probabilities that reproduce the market prices of the remaining options in the market not used as input for the proposed maxentropic procedure. All 3,048 cases are available on-line in the supplementary material of this paper.

As it is described in Table 2 the difference between some datasets is the time at which the physical probability is known, and the time at which the data about the price of the option was collected, which influence the price considerably. For example when the remaining time to maturity is just a couple of weeks, future weather is more certain.

In Table 3 we describe the data behind the market model for contract 1, which has 10 states and 6 assets. In Table 3 the level and range of each state is given, as well as the three different sets of probabilities. A closer inspection of Table 3 reveals that the three sets of probabilities are significantly different.

Table 3: Market model for dataset 1

| State | $\omega_1$ | $\omega_2$ | $\omega_3$ | $\omega_4$ | $\omega_5$ |
|---|---|---|---|---|---|
| Level | 347.5 | 715 | 737.5 | 765 | 796 |
| Range | $[0, 695)$ | $[695, 725)$ | $[725, 750)$ | $[750, 780)$ | $[780, 812)$ |
| Phys. Prob. HBA | 0.1282 | 0.0769 | 0.0001 | 0.1026 | 0.1537 |
| Phys. Prob. CAR | 0.1692 | 0.0924 | 0.0910 | 0.1248 | 0.1311 |
| Phys. Prob. EQP | 0.1000 | 0.1000 | 0.1000 | 0.1000 | 0.1000 |
| State | $\omega_6$ | $\omega_7$ | $\omega_8$ | $\omega_9$ | $\omega_{10}$ |
| Level | 824.5 | 846 | 861.5 | 884 | 1100 |
| Range | $[812, 837)$ | $[837, 855)$ | $[855, 868)$ | $[868, 900)$ | $[900, 1300)$ |
| Phys. Prob. HBA | 0.1026 | 0.0769 | 0.0256 | 0.1282 | 0.2051 |
| Phys. Prob. CAR | 0.0957 | 0.0662 | 0.0391 | 0.0796 | 0.1109 |
| Phys. Prob. EQP | 0.1000 | 0.1000 | 0.1000 | 0.1000 | 0.1000 |

To calibrate the risk neutral probability we consider the collection of options of the European type that are available in the market at the day of interest, e.g. 20/11/2007, which is before the measurement period of the contract. The details of the options are presented in Table 4.

In Table 4, the types of the options are listed in the first row, and the strikes and prices are defined in the following two rows, respectively. In the next section we will use

Table 4: Call and Put HDD options for dataset 1

| Type | Put | Put | Put | Call | Call | Call |
|---|---|---|---|---|---|---|
| Strike | 710 | 740 | 760 | 800 | 825 | 850 |
| Price | 18 | 27 | 34 | 47 | 36 | 27 |

a subset of the available options to calibrate the risk neutral probabilities, and then use those probabilities to obtain further option prices of different strike prices, including, but not limited to, those not used during the calibration. To examine on the consistency of the prices we compare the predicted price of the options not used for calibration versus their prices as quoted in the market.

Let us label the options from left to right according to the first row of Table 4, and denote their payoff at the exercise time by $O_i(\hat{S}_j, K_i)$, where $O_i(\hat{S}_j, K_i) = (\hat{S}_j - K_i)^+$ if the option is a call or $O_i(\hat{S}_j, K_i) = (K_i - \hat{S}_j)^+$ if it is a put. Table 5 presents the payoffs of all options depending on the strike price and the state of the market.

Table 5: Option payoffs for dataset 1

| Option | $\omega_1$ | $\omega_2$ | $\omega_3$ | $\omega_4$ | $\omega_5$ | $\omega_6$ | $\omega_7$ | $\omega_8$ | $\omega_9$ | $\omega_{10}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $O_1$ | 362.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $O_2$ | 392.5 | 30 | 2.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $O_3$ | 412.5 | 50 | 22.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $O_4$ | 0 | 0 | 0 | 0 | 0 | 24.5 | 46 | 61.5 | 84 | 300 |
| $O_5$ | 0 | 0 | 0 | 0 | 0 | 0 | 21 | 36.5 | 59 | 275 |
| $O_6$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 11.5 | 34 | 250 |

Note that in this case there are only 6 assets while there are 10 states. Thus the (constrained) linear problem (3.3) is quite undetermined. As it is shown in the next section, we do not even need all the equations to determine the risk neutral probabilities consistently.

Next we consider the 8$^{\text{th}}$ dataset. This refers to the same contract as dataset 7, but with different remaining time to maturity. This contract is very interesting since the final outcome of the HDDs is significantly different than the historical average.

In this case we have nine states and there are nine options available in the market. The dataset for this case is displayed in the next three tables. In Table 6 we specify the underlying market model. The available options traded in the market together with their prices and strikes are presented in Table 7. Finally, in Table 8 the payoffs of the 9 options for the different states are presented.

Table 6: Market model for for dataset 8

| State | $\omega_1$ | $\omega_2$ | $\omega_3$ | $\omega_4$ | $\omega_5$ |
|---|---|---|---|---|---|
| Level | 250 | 560 | 632.5 | 652.5 | 666.25 |
| Range | $[0, 500)$ | $[500, 620)$ | $[620, 645)$ | $[645, 660)$ | $[660, 672.5)$ |
| Phys. Prob. HBA | 0.0001 | 0.0270 | 0.0270 | 0.0001 | 0.0541 |
| Phys. Prob. CAR | 0.0064 | 0.1461 | 0.0809 | 0.0574 | 0.0495 |
| Phys. Prob. EQP | 0.1111 | 0.1111 | 0.1111 | 0.1111 | 0.1111 |
| State | $\omega_6$ | $\omega_7$ | $\omega_8$ | $\omega_9$ | |
| Level | 677.5 | 688.75 | 747.5 | 900 | |
| Range | $[672.5, 682.5)$ | $[682.5, 695)$ | $[695, 800)$ | $[800, 1000)$ | |
| Phys. Prob. HBA | 0.1350 | 0.1080 | 0.2973 | 0.3514 | |
| Phys. Prob. CAR | 0.0452 | 0.0564 | 0.4146 | 0.1435 | |
| Phys. Prob. EQP | 0.1111 | 0.1111 | 0.1111 | 0.1111 | |

Table 7: Call and Put HDD options for dataset 8

| Type | Put | Call | Call | Put | Call | Call | Put | Call | Call |
|---|---|---|---|---|---|---|---|---|---|
| Strike | 640 | 640 | 650 | 650 | 670 | 675 | 690 | 690 | 700 |
| Price | 1 | 61 | 53 | 3 | 36 | 33 | 14 | 24 | 21 |

Table 8: Option payoffs for dataset 8

| Option | $\omega_1$ | $\omega_2$ | $\omega_3$ | $\omega_4$ | $\omega_5$ | $\omega_6$ | $\omega_7$ | $\omega_8$ | $\omega_9$ |
|---|---|---|---|---|---|---|---|---|---|
| $O_1$ | 390 | 80 | 7.5 | 0 | 0 | 0 | 0 | 0 | 0 |
| $O_2$ | 0 | 0 | 0 | 12.5 | 26.25 | 37.5 | 48.75 | 107.5 | 260 |
| $O_3$ | 0 | 0 | 0 | 2.5 | 16.25 | 27.5 | 38.75 | 97.5 | 250 |
| $O_4$ | 400 | 90 | 17.5 | 0 | 0 | 0 | 0 | 0 | 0 |
| $O_5$ | 0 | 0 | 0 | 0 | 0 | 7.5 | 18.75 | 77.5 | 230 |
| $O_6$ | 0 | 0 | 0 | 0 | 0 | 2.5 | 13.75 | 72.5 | 225 |
| $O_7$ | 440 | 130 | 57.5 | 37.5 | 23.75 | 12.5 | 1.25 | 0 | 0 |
| $O_8$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 57.5 | 210 |
| $O_9$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 47.5 | 200 |

## 4.3 Empirical results

As it is described in Section 4.1 we consider 8 datasets. The number of available options varies from 6 in the first dataset to 9 in the last one. For each dataset we consider all possible subsets of the options and their prices. Using the physical state probabilities as reference, we determine the risk neutral probabilities for each subset. Next, using the estimated risk neutral probabilities we compute the option (call and put) prices for a fine grid of strikes.

Finally, we examined the predictive power of the methodology by predicting the price of the options not included in the data (sub)set. We repeat our methodology for the three different sets of physical probabilities.

The total number of sets examined is 3,048. Note that in all cases we were dealing with an undetermined linear system of equations subject to convex constraints. Data will often contain option prices that are inconsistent between them. This induces unboundedness in the dual entropy, but it is not a problem nor does it preclude our method from finding a solution. For example in all $1,107$ sets where the dual entropy was unbounded, we found a bounded solution when a subset of the options used for fitting. This is an expected outcome and it has the added advantage of helping the agent identify options which are not consistently priced between themselves. For our particular dataset, the percentage of converging sets is negatively correlated to the number of options used for fitting (100%, 97%, 87%, 65%, 33%, 7%, 0%, 0% and 0% for $1, \ldots, 9$ respectively), indicating that multiple options were potentially mispriced. Note that mispricing under exchange constraints can happen even if options are priced correctly. This can happen due to the observable granularity of prices in the exchange (for example two options with different strike may have the same observed price; the exchange does not admit decimal granularity in the pricing, and a theoretical price of 1.76 would be observed as the same value as a theoretical price of 1.99). Also, due to the same granularity constraint, a drop of 1 unit in the strike can be accompanied by 1 unit drop in the price for very out-of-the-money (OTM) options, which would seem counter-intuitive. Our method is able to extract a solution by simply sampling from subsets of the observed set.

Out of the $\binom{9}{2} = 36$ sets in dataset 8 containing only 2 options, 35 converged, and only one did not converge. This one was the one using only the 640 and 650 puts (first and fourth options in dataset 8). Convergence is not dependent on the probabilities used, but on the characteristics of the set of options used. For each set of options comprising a given set, if that set did not converge, i.e. the entropy problem didn't have a finite solution, there was a subset of options of that set which would converge. This is achieved by simply sampling within the original set, to capture a smaller subset with consistently-priced options. Additionally note that this problem is most likely to appear with deep

OTM options, where mispricing is more often due to the aforementioned exchange-imposed granularity. Note, for the aforementioned two options, that a choice of $\lambda = (-2.5, 1)$, for example, aligns with the condition in Lemma 3.1 for the unboundedness of the entropy function when using solely these two options for the optimization. We found that often unboundness occurs when OTM options too close to each other (and of the same nature, whether puts or calls) are included in the optimization. These two options, for example, are the ones with the lowest premium in dataset 8.

Second, in the remaining cases, the method provided us with a solution. As it was expected the accuracy of the prediction of the option prices varies depending on the number of options used, the subset and the selected options within the subset. In the cases in which the dual entropy was bounded below, usually 4 options were sufficient to completely characterize the risk neutral measure in a consistent and accurate fashion.

The key step in the maxentropic procedure is the minimization of the dual entropy described by (3.7). A modified Newton algorithm is carried out, [37]. The (fast) algorithm stops when the gradient of the function is less than $10^{-6}$. Finding the minimizing vector $\boldsymbol{\lambda}^*$, the risk neutral density is obtained according to equation (3.6), and consequently the risk neutral probabilities can be determined. In all cases the condition $\sum_i \rho_i p_i = 1$ of constraints (3.4) holds. Each optimization took, on average, around 0.15 seconds in R to derive the risk neutral probabilities.

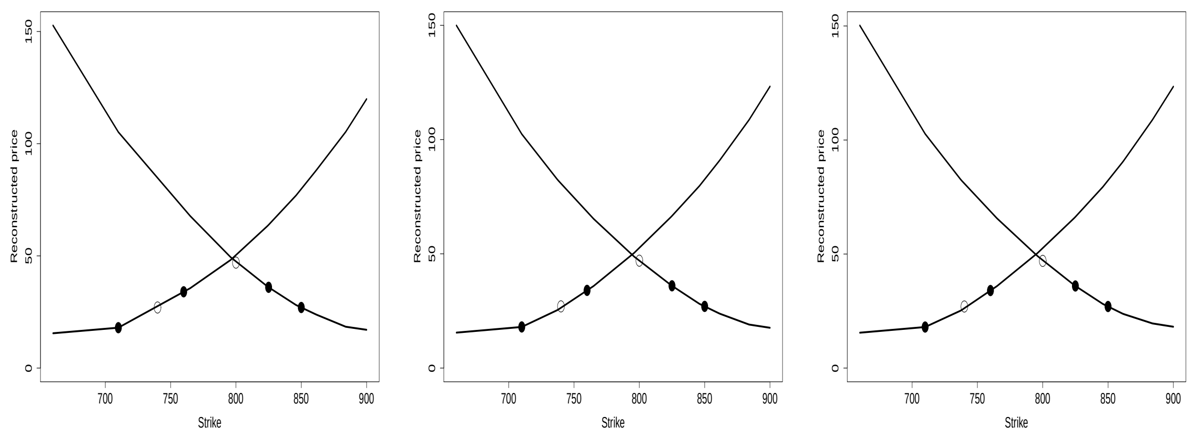
Next, we analyse the results obtained by dataset 1. The data described in Tables 3 – 5 were used. In order to derive the risk neutral probabilties, for this particular set, we consider a subset of two puts and two calls with strike prices of 710, 760, 825 and 850 respectively.

Table 9: Risk neutral probabilities for dataset 1 using 4 option prices

| RNP | $q_1$ | $q_2$ | $q_3$ | $q_4$ | $q_5$ |
|---|---|---|---|---|---|
| HBA | 0.0497 | 0.2703 | 0.0002 | 0.0868 | 0.1301 |
| CAR | 0.0497 | 0.2208 | 0.1101 | 0.0865 | 0.0908 |
| EQP | 0.0497 | 0.2146 | 0.1239 | 0.0790 | 0.0790 |
| **RNP** | $q_6$ | $q_7$ | $q_8$ | $q_9$ | $q_{10}$ |
| HBA | 0.0868 | 0.1012 | 0.0346 | 0.1551 | 0.0853 |
| CAR | 0.0663 | 0.0988 | 0.0652 | 0.1237 | 0.0882 |
| EQP | 0.0790 | 0.0931 | 0.0958 | 0.0953 | 0.0906 |

We can compute the price of any derivative using those 4 options. First, we want to confirm whether the prices in the market are reproduced.

Figure 1 shows the price of the options computed using the risk neutral probabilities shown in Table 9. We estimated three sets of prices, one for each set of physical probabilities, and proceeded to reconstruct the prices of the remaining options (marked by dots), whereas the predicted prices of the options are marked by circles. Figure 1 reveals that the proposed method can reconstruct the option prices and can be used for price forecasting of the remaining options. The results are similar for the three sets of probabilities.




(a) HBA (b) CAR (c) EQP

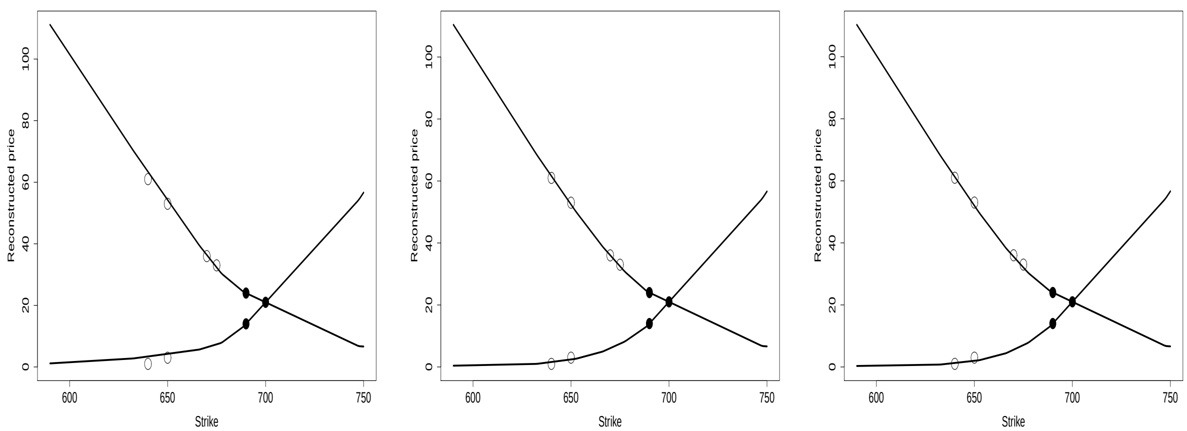
Figure 1: Option prices across strikes for dataset 1, determined from four options under each of the 3 different probabilities. Dots represent the option prices used for fitting, and circles represent the observed option prices not used for fitting.

The second case that we present corresponds to dataset 8. The market data for this case was detailed in Tables 6 – 8. Using the three options with the highest strikes as data, we applied the proposed maxentropic method and obtained the risk neutral probabilities presented in Table 10.

Next, in Figure 2 the curves of the option prices are presented. The description of the panels is as above. From Figure 2 it is clear that using only three options, we can confirm consistency of the valuation of the remaining six options. A closer inspection of Figure 2 reveals that using physical probabilities derived from the CAR model we obtain more accurate valuation.

Table 10: Risk neutral probabilities for dataset 8 determined from 3 option prices

| RNP | $q_1$ | $q_2$ | $q_3$ | $q_4$ | $q_5$ |
|---|---|---|---|---|---|
| HBA | 0.0001 | 0.0379 | 0.0479 | 0.0002 | 0.1071 |
| CAR | 0.0000 | 0.0137 | 0.0683 | 0.0888 | 0.1165 |
| EQP | 0.0000 | 0.0105 | 0.0595 | 0.0958 | 0.1330 |
| RNP | $q_6$ | $q_7$ | $q_8$ | $q_9$ | |
| HBA | 0.2771 | 0.2299 | 0.2557 | 0.0443 | |
| CAR | 0.1496 | 0.2631 | 0.2557 | 0.0443 | |
| EQP | 0.1739 | 0.2274 | 0.2557 | 0.0443 | |




(a) HBA (b) CAR (c) EQP

Figure 2: Option prices across strikes determined for dataset 8, from three options under each of the 3 different probabilities. Dots represent the option prices used for fitting, and circles represent the observed option prices not used for fitting.

In Table 11 we present the out-of-sample mean absolute distance (MAD) between the fitted curve and the options observed but not used for fitting, as a function of the probability and the number of options used for fitting across sets. Our results indicate that using only one or two options for fitting produces large out-of-sample pricing errors. For datasets 1-3, EQP is best for small number of options, but as the number of options used is increasing, the results become similar for all the sets of initial probabilities. On the other hand for datasets 4-8 the CAR and HBA probabilities produce the lower out-of-sample pricing errors.

The above results are confirmed by Table 12. Table 12 groups the datasets by contract. Hence, we can isolate the error in MAD that arises from a specific contract. For the first contract (datasets 1+2+3) the EQP produce the lowest error followed by CAR and HBA.

On the other hand, for the second (datasets 4+5+6) and the third contract (datasets 7+8) CAR and HBA produce lower errors while the MAD using the EQP is highest. It is worth to mention that for the first contract the realised HDDs index is 817.5 and it is very close to the historical average of the last 30 years which is 837.5. On the contrary, the realised HDDs index for the second contract is only 467 while the historical average is 505.9. Finally, in the third contract the realised HDDs index is 672.5 and significantly lower than the historical average of 764.2.

Finally, Table 13 represents the MAD between the fitted curve and the observed prices, not used for fitting, across converging sets of the same dimension, i.e. sets with the same number of observations used for fitting. The grouping was done column-wise by combining datasets 1+4 (first group), datasets 2+5+7 (second group) and datasets 3+6+8 (third group). This grouping will help us isolate the time effect and analyse whether it affects the accuracy of the pricing. In other words, in the first group all contracts are traded before the measurement period when meteorological forecasts are not available. In the second group all contracts are traded at the beginning of the measurement period while the third group contains all the contracts that are traded in the middle and near the end of the measurement period.

A closer inspection of Table 13 reveals that CAR probabilities appear to be more consistent for the reconstruction of the observed option prices not used for fitting, followed by HBA and EQP. Moreover, when using four or more options all methods produce similar out-of-sample MADs. In fact it appears that using 4 or 5 options for fitting is the optimal since more than that, for our dataset, options that are not coherent with each other are included, which increases the likelihood of non-convergence.

Even under the (strong) assumption that prices are coherent between each other, the benchmark MAD under perfect fit would be 0.25, given the integer-based granularity of the premium data in the exchange (any observed price X could be any value between X-0.5 and X+0.5, with an expected absolute error of 0.25 if we assume they are uniformly distributed in that range). For example, the mean absolute distances in the three plots in Figure 2 are 1.40, 0.49 and 0.66 for HBA, CAR and EQP respectively indicating that option prices were, on average, only 1.15, 0.24 and 0.41 away from the benchmark.

Assuming market pricing being coherent, one might expect that as we get closer to the maturity date the accuracy of our pricing will increase. On the contrary, if we expect lack of coherence, it will manifest itself most often during times when there can be disagreement about the inputs and models, probably at the beginning of the measurement period, where there is still a number of potential ways in which the information available is used. Indeed our results in Table 13 confirm this. The first and the third group have smaller errors. For contracts traded in group one traders rely on historical data, while for group 3 traders rely mostly on meteorological forecasts. On the other hand for the second group some meteorological information is available but traders also have to rely on historical data. The mix of information in group two also means higher room for subjective interpretation/valuation by ad hoc approaches. Uncertainty may be higher when the volatility from all information sources is higher. This scenario is expected in the beginning of the measurement period where there is room for subjective approaches by traders and not before, where most of the information arises from historical data, or at the end of the measurement period, where most of the information arises from meteorological forecasts. Hence the second group is the one where most diverse pricing can occur, while there may be more coherence in groups one and three, i.e. more of a "standard" approach taken by traders versus a higher set of inputs that could be used when pricing during periods of group two. Inconsistent sets of market prices are a reality of life, and identifying them opens the gate for arbitrage opportunities.

## 5 Conclusions

We have provided a computationally fast, non parametric, model-free method for inferring risk neutral probabilities, along with their densities, with respect to several possible physical state probabilities. Clearly, when presented with a collection of prices of options, a market analyst does not have a way to decide which ones are more informative than the others. The methodology that we have developed can handle that issue in a simple way: a subset of the data is more informative if it can consistently reproduce the data not used to determine the risk neutral probabilities. When this is the case, we feel more confident for the prices of any other derivative computed using the same neutral risk measures.

Our results indicate that the proposed method can extract accurately the risk neutral

Table 11: Mean absolute distance of out-of-sample points by dataset, probability, and number of options used for fitting across experiments.

| $N$ | Prob | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | HBA | 22.9 | 23.1 | 24.3 | 7.9 | 3.9 | 18.0 | 36.9 | 18.6 |
| | CAR | 16.7 | 16.9 | 22.6 | 7.8 | 4.6 | 18.4 | 20.5 | 18.7 |
| | EQP | 7.7 | 8.1 | 15.2 | 23.1 | 24.5 | 31.0 | 33.8 | 21.5 |
| 2 | HBA | 10.1 | 10.3 | 9.8 | 3.9 | 2.6 | 5.4 | 8.9 | 3.7 |
| | CAR | 7.2 | 7.3 | 7.9 | 3.5 | 2.9 | 5.8 | 5.2 | 5.1 |
| | EQP | 4.8 | 5.0 | 5.9 | 10.9 | 11.1 | 11.4 | 15.4 | 7.6 |
| 3 | HBA | 5.1 | 5.0 | 4.8 | 2.1 | 2.1 | 2.9 | 4.5 | 1.6 |
| | CAR | 4.3 | 4.0 | 4.2 | 1.8 | 2.1 | 2.7 | 4.7 | 1.9 |
| | EQP | 2.4 | 2.3 | 2.8 | 6.0 | 5.1 | 5.4 | 10.8 | 3.6 |
| 4 | HBA | 1.9 | 2.2 | 2.2 | 1.1 | 2.3 | 1.1 | 5.1 | 1.2 |
| | CAR | 1.8 | 1.8 | 2.0 | 0.9 | 2.2 | 0.9 | 5.1 | 1.3 |
| | EQP | 1.2 | 1.1 | 1.3 | 1.6 | 3.1 | 1.1 | 6.6 | 2.2 |
| 5 | HBA | 0.8 | 1.2 | 1.3 | 0.4 | 1.4 | - | 5.4 | 0.8 |
| | CAR | 0.9 | 0.7 | 0.9 | 0.2 | 1.5 | - | 5.4 | 1.1 |
| | EQP | 0.9 | 0.6 | 0.7 | 0.3 | 2.0 | - | 7.6 | 1.4 |
| 6 | HBA | - | - | - | - | - | - | - | 0.7 |
| | CAR | - | - | - | - | - | - | - | 1.0 |
| | EQP | - | - | - | - | - | - | - | 1.1 |

Table 12: The contribution of the contract effect to the accuracy of pricing out-of-sample.

| | Datasets 1+2+3 | | | Datasets 4+5+6 | | | Datasets 7+8 | | |
|---|---|---|---|---|---|---|---|---|---|
| $N$ | HBA | CAR | EQP | HBA | CAR | EQP | HBA | CAR | EQP |
| 1 | 23.4 | 18.7 | 10.3 | 9.9 | 10.3 | 26.2 | 27.8 | 19.6 | 27.7 |
| 2 | 10.1 | 7.5 | 5.2 | 4.0 | 4.1 | 11.1 | 6.3 | 5.1 | 11.5 |
| 3 | 5.0 | 4.2 | 2.5 | 2.4 | 2.2 | 5.5 | 3.1 | 3.3 | 7.2 |
| 4 | 2.1 | 1.9 | 1.2 | 1.5 | 1.3 | 1.9 | 3.1 | 3.2 | 4.4 |
| 5 | 1.1 | 0.8 | 0.7 | 0.9 | 0.8 | 1.1 | 3.1 | 3.2 | 4.5 |
| 6 | - | - | - | - | - | - | 0.7 | 1.0 | 1.1 |

Table 13: The contribution of the time effect to the accuracy of pricing out-of-sample.

| | Datasets 1+4 | | | Datasets 2+5+7 | | | Datasets 3+6+8 | | |
|---|---|---|---|---|---|---|---|---|---|
| $N$ | HBA | CAR | EQP | HBA | CAR | EQP | HBA | CAR | EQP |
| 1 | 15.4 | 12.3 | 15.4 | 21.3 | 14.0 | 22.2 | 20.3 | 19.9 | 22.6 |
| 2 | 7.0 | 5.3 | 7.8 | 7.3 | 5.1 | 10.5 | 6.3 | 6.3 | 8.3 |
| 3 | 3.6 | 3.0 | 4.2 | 3.9 | 3.6 | 6.1 | 3.1 | 2.9 | 3.9 |
| 4 | 1.5 | 1.3 | 1.4 | 3.2 | 3.0 | 3.6 | 1.5 | 1.4 | 1.5 |
| 5 | 0.6 | 0.6 | 0.6 | 2.7 | 2.5 | 3.4 | 1.0 | 1.0 | 1.0 |

probabilities using only few market data points, and it can be used for valuation of other options traded in the market. Hence, the proposed method is ideal to overcome the data sparsity problem that governs the weather market. It is worth to mention that we were able to price accurately options even in the cases where the realised underlying HDDs indices were significantly different from the historical average.

In an extensive analysis we examined whether different sets of physical probabilities, indicating different levels of expertise, can affect the accuracy of out-of-sample valuation. The three sets of probabilities are derived by a naive, a semi-advanced and a state of the art modelling procedure. Our results indicate that the CAR and HBA probabilities can provide a better reconstruction of the original option prices, and also produce a lower MAD in the valuation of the options not used in the fitting procedure. As the number of options used for fitting increases the differences between the three methods become smaller.

Our results indicate that when the available information in the market arrives from historical data or from meteorological forecasts pricing is more coherent. However when there is a mix of information in the market, non-consistent sets of market prices are observed. Testing for consistency using different data subsets should be regarded part of the procedure.

Finally our results show that the proposed method already produces excellent results when as little as 4 options are used for fitting. On the other hand, including OTM options too close to each other usually generates convergence issues, perhaps related to the granularity of the pricing in the exchange, where lower premium options will suffer the most impact as a proportion of the true price, or related to potential incoherent pricing.

## References


[1] Abbas, A. E. (2006). Maximum entropy utility. *Operations Research*, 54(2):277–290.

[2] Alaton, P., Djehince, B., and Stillberg, D. (2002). On modelling and pricing weather derivatives. *Applied Mathematical Finance*, 9:1–20.

[3] Alexandridis, A. and Zapranis, A. (2013). *Weather Derivatives: Modeling and Pricing Weather-Related Risk*. Springer, New York.

[4] Benth, F. E. and Saltyte-Benth, J. (2007). The volatility of temperature and pricing of weather derivatives. *Quantitative Finance*, 7(5):553–561.

[5] Benth, F. E. and Saltyte-Benth, J. (2013). *Modeling and pricing in financial markets for weather derivatives*, volume 17. World Scientific.

[6] Benth, F. E., Saltyte-Benth, J., and Koekebakker, S. (2007). Putting a price on temperature. *Scandinavian Journal of Statistics,*, 34:746–767.

[7] Bertrand, J.-L., Brusset, X., and Fortin, M. (2015). Assessing and hedging the cost of unseasonal weather: Case of the apparel sector. *European Journal of Operational Research*, 244(1):261–276.

[8] Borwein, J. and Lewis, A. (2000). *Convex Analysis and Nonlinear Optimization.* CMS Books in Mathematics. Springer-Verlag, New York.

[9] Borwein, J. M., Lewis, A. S., and Noll, D. (1996). Maximum entropy reconstruction using derivative information, part 1: Fisher information and convex duality. *Mathematics of Operations Research*, 21(2):442–468.

[10] Brockett, P., L., Wang, M., and Yang, C. (2005). Weather derivatives and weather risk management. *Risk Management and Insurance Review*, 8(1):127–140.

[11] Buchheim, L. and Kolaska, T. (2016). Weather and the psychology of purchasing outdoor movie tickets. *Management Science*, forthcoming.

[12] Buchholz, M. and Musshoff, O. (2014). The role of weather derivatives and portfolio effects in agricultural water management. *Agricultural Water Management*, 146:34–44.

[13] Buxey, G. (1988). Production planning under seasonal demand: A case study perspective. *Omega*, 16(5):447–455.

[14] Caliskan Demirag, O. (2013). Performance of weather-conditional rebates under different risk preferences. *Omega*, 41(6):1053–1067.

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

[16] Cao, M. and Wei, J. (2004). Weather derivatives valuation and market price of weather risk. *Journal of Future Markets*, 24(11):1065–1089.

[17] Cao, Q., Ewing, B. T., and Thompson, M. A. (2012). Forecasting wind speed with recurrent neural networks. *European Journal of Operational Research*, 221(1):148–154.

[18] Challis, S. (1999). Bright forecast for profits. *Reactions*, June edition.

[19] Chen, F. Y. and Yano, C. A. (2010). Improving supply chain performance and managing risk under weather-related demand uncertainty. *Management Science*, 56(8):1380–1397.

[20] Cozzolino, J. M. and Zahner, M. J. (1973). The maximum-entropy distribution of the future market price of a stock. *Operations Research*, 21(6):1200–1211.

[21] Davis, M. (2001). Pricing weather derivatives by marginal value. *Quantitative Finance*, 1:1–4.

[22] Elberg, C. and Hagspiel, S. (2015). Spatial dependencies of wind power and interrelations with spot price dynamics. *European Journal of Operational Research*, 241(1):260–272.

[23] Elias, R. S., Wahab, M. I. M., and Fang, L. (2014). A comparison of regime-switching temperature modeling approaches for applications in weather derivatives. *European Journal of Operational Research*, 232(3):549–560.

[24] Fu, L., Trudel, M., and Kim, V. (2009). Optimizing winter road maintenance operations under real-time information. *European Journal of Operational Research*, 196(1):332–341.

[25] Geman, H. and Leonardi, M.-P. (2005). Alternative approaches to weather derivatives pricing. *Managerial Finance*, 31(6):46–72.

[26] Gulko, L. (1999a). The entropic market hypothesis. *International Journal of Theoretical and Applied Finance*, 2(03):293–329.

[27] Gulko, L. (1999b). The entropy theory of stock option pricing. *International Journal of Theoretical and Applied Finance*, 2(03):331–355.

[28] Gulko, L. (2002). The entropy theory of bond option pricing. *International Journal of Theoretical and Applied Finance*, 5(04):355–383.

[29] Gulpinar, N. and Canakoglu, E. (2017). Robust portfolio selection problem under temperature uncertainty. *European Journal of Operational Research*, 256(2):500–523.

[30] Gzyl, H. and Mayoral, S. (2017). Maxentropic solutions to a convex interpolation problem motivated by utility theory. *Entropy*, 19(4):153.

[31] Gzyl, H., ter Horst, E., and Molina, G. ((submitted, under review)). Risk neutral measure determination from price ranges: Single period market models. *Applied Stochastic Models for Business and Industry*.

[32] Hanley, M. (1999). Hedging the force of nature. *Risk Professional*, 1:21–25.

[33] Hardle, K. W., Lopez-Cabrera, B., and Teng, H.-W. (2015). State price densities implied from weather derivatives. *Insurance: Mathematics and Economics*, 64:106–125.

[34] Jaynes, E. T. (1957). Information theory and statistical mechanics. *The Physical Review*, 106:620–630.

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

[36] Judge, G. G. and Mittelhammer, R. C. (2011). *An information theoretic approach to econometrics*. Cambridge University Press.

[37] Lange, K. (2010). *Numerical Analysis for Statisticians*, volume 1. Springer.

[38] Le Cadre, H., Papavasiliou, A., and Smeers, Y. (2015). Wind farm portfolio optimization under network capacity constraints. *European Journal of Operational Research*, 247(2):560–574.

[39] Perez-Gonzalez, F. and Yun, H. (2013). Risk management and firm value: Evidence from weather derivatives. *The Journal of Finance*, 68(5):2143–2176.

[40] Purnanandam, A. and Weagley, D. (2016). Can markets discipline government agencies? evidence from the weather derivatives market. *The Journal of Finance*, 71(1):303–334.

[41] Rajasekera, J. and Yamada, M. (2001). Estimating the firm value distribution function by entropy optimization and geometric programming. *Annals of Operations Research*, 105(1):61–75.

[42] Regnier, E. (2008). Doing something about the weather. *Omega*, 36(1):22–32.

[43] Rouge, R. and El Karoui, N. (2000). Pricing via utility maximization and entropy. *Mathematical Finance*, 10(2):259–276.

[44] Stulec, I. (2017). Effectiveness of weather derivatives as a risk management tool in food retail: The case of croatia. *International Journal of Financial Studies*, 5(1):2.

[45] Sun, B. and van Kooten, G. C. (2015). Financial weather derivatives for corn production in northern china: A comparison of pricing methods. *Journal of Empirical Finance*.

[46] Wilks, D. S. (2002). Realizations of daily weather in forecast seasonal climate. *Journal of Hydrometeorology*, 3(2):195–207.

[47] Wilks, D. S. (2011). *Statistical Methods in the Atmospheric Sciences*, volume 100 of *International Geophysics Series*. Academic Press, Oxford, UK, 3rd edition.

[48] Zapranis, A. and Alexandridis, A. (2008). Modelling temperature time dependent speed of mean reversion in the context of weather derivetive pricing. *Applied Mathematical Finance*, 15(4):355 – 386.

[49] Zhou, R.; Cai, R. T. G. (2013). Applications of entropy in finance: A review. *Entropy*, 15:4909–4931.