I DOTTORATO
A DEGLI STUDI DI MILANO-BICOCCA

Dipartimento di / Department of

Statistics and Quantitative Methods

Dottorato di Ricerca in / PhD program Statistics and Mathematical Finance   Ciclo / Cycle XXXIII

Curriculum in (se presente / if it is) Mathematical Finance

MODELING AND RISK MANAGEMENT WITH APPLICATIONS IN FINANCIAL AND WEATHER DERIVATIVES

Cognome / Surname   Kutrolli   Nome / Name   Gleda

Matricola / Registration number   827262

Tutore / Tutor: Rosanna Grassi

Supervisor: Silvana Stefani
(se presente / if there is one)

Co- Supervisor :   Fred Espen Benth
(se presente / if there is one)

Coordinatore / Coordinator: Emanuela Rosazza Gianin

ANNO ACCADEMICO / ACADEMIC YEAR   2020/2021

Gleda Kutrolli

MODELING AND RISK MANAGEMENT WITH APPLICATIONS IN FINANCIAL AND WEATHER DERIVATIVES

Thesis submitted for the degree of Doctor of Philosophy

Department of Statistics and Quantitative Methods
University of Milano-Bicocca

Tutored by
Prof. Rosanna Grassi
Supervised by
Prof. Silvana Stefani
Prof. Fred Espen Benth

2021

This thesis work is dedicated to my parents Nimete and Petrit Kutrolli, who have always loved me unconditionally and whose good examples have tought me to work hard for the things i aspire to achieve.

Abstract

The purpose of this dissertation is to conduct an in-depth analysis to modeling, pricing and risk management for financial instruments traded in financial and weather market. This work is composed of one scientific report and four working papers that focus specifically on the following topics:

- The scientific report presents and discusses a review of recent literature on financial products used to protect against climate change. Through comparisons of different methodologies proposed by different researchers this study shows the necessity and importance of using weather derivatives and insurance contracts to better protect and develop world wide financial markets.
- In the first paper, we propose temperature-based risk management using hybrid financial instruments built on weather derivatives. The Value-at-Risk technique is exploited, in order to define the level of the critical temperature, to estimate the prices of such derivatives. The results presented in the paper are promising, and show that such contracts can positively cover temperature risk.
- In the second paper, we present a way to hedge temperature risk exploiting weather derivatives contracts by considering 'tail events' and the standard financial approach to tackle them such as Value-at-Risk and Expected Shortfall. We perform the application of risk measures through historical and parametric methods and analyze the effectiveness of Expected Shortfall based approach for hedging meteorological risk. Even, according to risk measure theory, Expected Shortfall captures diversification while Value-at-Risk does not, numerical results show that it is more convenient to enter a single contract that covers more months rather than monthly contracts spanned on the same period.
- In the third paper, we propose forecasting based on a stochastic model of the probability distribution. We suggest to incorporate model uncertainty by considering forecasting using dynamical stochastic evolutions of the probability distribution of the model in question. We have considered and compared the results of two different classes of autoregressive models where the so-called stationary distribution is and is not normal. At the end, we have compared the uncertainty measure based on worst case approach and the proposed approach using a Value-at-Risk based measure giving tighter estimates.
- In the fourth paper, we propose some functional linear models predicting wind speed from temperature data in which both the response and the covariate variables are functions. We have proposed the functional linear model observed in different patterns which are powerful enough to describe the dynamics of wind speed in time.

Acknowledgements

Working and writing a PhD thesis is more than an individual experience, since it take place in a social context and includes several persons, whom I would like to thank sincerly.

First and foremost, I would like to express my wholehearted appreciation and gratitude to my supervisor Prof. Silvana Stefani for her persistent and continuous support at all times throughout my study period. I am very grateful for the opportunity to learn about Italian culture and enjoy all the experience in such environment surrounded by very nice people. I have no words to express how thankful I am for being a student of such a great and kind professor. She has always been encouraging and motivating me, even in difficult situations. In Albania, we have an popular expression that is "teachers and professors are like second parents for their students", and i really felt that in my relationship with her. Her dynamism, vision, sincerity and immense knowledge have deeply inspired me. I could not have imagined having a better advisor and mentor for my PhD study. Thanks a lot, Silvana!

Next, my special appreciation and gratitude goes to my second supervisor Prof. Fred Espen Benth for his unceasing support, advices and guidance during the visiting period i spent in Oslo. He has continued to inspire me with his papers and books since my first contact with mathematical finance theory applied in insurance and energy markets. It was a wonderful experience and for more i am very grateful for the opportunity to know and learn more about Norwegian culture. Thanks, Fred!

I will always be grateful to Prof. Rosanna Grassi, my tutor, for her support and kindness.

I also wish to thank my collaboration team in research in Italy especially Enrico Moretto who have been a great source of support.

I owe many thanks to my friends and colleagues at the University of Milano-Bicocca and University of Oslo, especially my colleagues who started this journey with me, in Finance: Paolo, Maria, Gabriele, and in Statistics: Kenny, Anna, Paolo, and Andrea. Without them this thesis may had needed much longer time to complete. My gratitude goes to all professors and other PhD students for providing a good working and enjoyable environment and for organizing very interesting seminars.

Last but not the least, i would like to thank my family especially my parents Nimete and Petrit who have been like a guiding light that with their love have always shown the right way to follow. Thanks to my brother Maksi always giving me the supportive words, the strength to go through all challanges i faced and my sisters Gentjana, Klodjana and Edjana for their love and trust. Also thanks to my husband Etjon, who has been a constant source of support and ecouragement during the challenges of PhD study and life. I am truly thankful for having you in my life.

I want to finalize this section of acknowledgements directed to my mom, even she is not alive and she has not been physically present throughout this wonderful journey, i wish you would be proud of me!

Thank you to all!
Gleda Kutrolli
Milano, January 2021

Contents

List of Figures

List of Tables

Chapter 1

Introduction

When I started working on this dissertation my original goal was rather simple. As for several years I was working with quantitative financial analysis, I was stunned by how mathematical and statistical methods were applied in finance and were used in order to propose hedging strategies for people who trade and deal with financial derivatives.

After I started my PhD, I read about weather derivatives. The very first questions that came in my mind were about what is a weather derivative? Is there any similarity with financial derivatives? What type of events weather derivatives cover and why should one be interested in weather derivatives? I got my very first responses from the book "Modeling and Pricing in Financial Markets for Weather Derivatives". I was impressed by the role weather derivatives play in many industries today and the way they are used for hedging purposes by managing the weather risk. While reading the book, my curiosity in this area was to see how statistical analysis of weather factors and stochastic processes can be used in modeling the time and space dynamics. At the end, I was fascinated by the way of applying the modern theory of mathematical finance on weather derivatives pricing and hedging. That was the moment I found myself very enthusiastic to combine the background I had in finance with this new area of weather market.

For this reason I started my research with a review of recent literature on financial and insurance products used to protect against climate change. Through comparisons of different methodologies proposed by several researchers, I understood what it was done until then and research needed to be extended in order to use weather derivatives and insurance contracts in a proper manner to better protect and develop world wide financial markets.

Hence, my PhD research began with a keen interest on temperature variable, and seeking answers to the following questions: how we can link weather derivatives with insurance contracts? Since a temperature contract (such as HDD and CDD) depend on a critical temperature or the so called threshold, has it always set to be at $18°$ C for all meteorological stations? Can we use Value-at-Risk or Expected Shortfall risk measures to compute the threshold? Can we use these approaches to measure the risk loadings and how effective is to use the same percentile as the one chosen for the threshold? Once I started my research work on this topic, I realized that the idea of proposing a hybrid instrument was flexible enough to allow considering the actuarial point of view. A hybrid instrument is capable of dealing with a negative risk event such as an insurance contract but priced as a derivative instrument, where risk loadings have been properly considered and charged. Also, there was clear evidence that the standard threshold of $18°$ C was not a proper choice for different regions. A reasonable choice is by applying the Value-at-Risk approach and later extended in Expected Shortfall approach. Since several adverse weather conditions can occur simultaneously, Expected Shortfall can assess consequences of such improbable events more accurately compared to Value-at-Risk.

Another important issue is related to pricing models proposed by researchers in the field. With the dissemination of quantitative methods in risk management and the advent of complex derivative products, mathematical models have played an increasingly important role in financial decision making, especially in the context of pricing and hedging of derivative instruments. But there is some confusion on how sensitive is the value of a given derivative to the choice of the pricing model. Are there some instruments more model sensitive than others? While the use of models led to a better understanding of market risks, it has given rise to a new type of risk, known as "model risk" or "model uncertainty", linked to the uncertainty on the choice of the model itself. Uncertainty on the choice of the pricing model can lead to the mispricing of a derivative products. While model uncertainty is acknowledged by most operators who make use of quantitative models, most of the discussion on this subject has stayed at a qualitative level and a quantitative framework for measuring model uncertainty is still lacking. For this reason, proposing a very novel method that models the probability density functions incorporating the uncertainty, which follows stochastic processes themselves, is important for measuring and managing the risk. The forecast probability is treated as a random variable in some state space of probability density functions, where the notion of model uncertainty is reduced to uncertainty on future volatility. The model risk is handled by a worst case approach which is extended based on Value-at-Risk and Expected Shortfall and is applied in pricing weather derivatives in order to reduce uncertainty.

On the other hand, some attempts of applying functional data analysis in analyzing weather data are done in the literature reviewed. Therefore, the curiosity pushed me to understand how weather data can be seen as functional objects and how we use functional data analysis in forecasting weather variables such as temperature and wind? These questions are followed from other more intriguing questions: can each weather station broadly be categorized in some common geographical zones with similar features by applying functional principal components analysis? In what way can the geographical category characterize the detailed temperature profile and account for the different profiles observed? Since, according to literature review in weather derivatives wind has a similar behavior with temperature, then could a temperature record be used to predict the total annual wind speed by using a linear regression where both the independent and dependent variables are functions? Can the temperature record be used as a predictor of the entire wind speed profile by a functional linear model? It is really wonderful how the functional linear models under scrutiny are able to regularize curves observed over a specific time period and predict curves at unobserved period of time. Moreover, these models are available not just to predict full wind profile and annual wind from temperature climate zones resulted from functional principal components analysis but also predict wind directly from temperature observations.

More concretely, the dissertation consists of one scientific report and four working papers where the scientific report and first three papers (Chapter 2, 3 4 and 6) are connected via modeling and managing the risk in weather derivatives with a specific focus on hybrid financial instruments, managing meteorological risk through VaR and ES and forecasting by applying FDA in weather variables. The last paper (Chapter 5) is an application of modeling and managing the risk in financial derivatives, with a specific focus on dynamic probabilistic forecasting.

- Scientific report attempts to describe methods and models currently in use in the weather derivative market from the studied literature. We start discussing different types of weather-based financial derivatives and focus on worldwide weather indices and explains why it is needed to develop and adopt weather indices. Next, we focus on an important issue in the weather derivatives markets which is the choice of the pricing methodology to use in order to obtain the "fair" value for different contracts. Later we pinpoint innovations in the area of the recent research. At the end, we conclude and discuss about future research.

- In the first paper we propose temperature-based risk management using hybrid financial instruments built on weather derivatives. We model temperature time series and price one-month forward option contracts for hedging adverse outcomes and then show how a "negative" weather performance can be counterbalanced by the "positive" performance of a hedging over-the-counter financial instrument, which can be tailored to meet specific needs. In order to estimate the prices of these derivatives, the conventional Value-at-Risk technique is exploited. This technique is applied to define the level of the critical temperature (called threshold) but also in the risk loadings added for hedging purposes. Numerical findings show that such contracts can positively cover temperature risk.

- In the second paper we hedge meteorological risk concerning temperatures by pricing weather derivatives contracts as efficient risk management instruments, assuming as a key ingredient for this approach some "tail event" triggered by a quantile-based critical threshold. Such thresholds are calculated assuming different quantile levels, applying the Value-at-Risk and Expected Shortfall for the distribution of temperatures as recorded in some historical series. It is shown that both Value-at-Risk and Expected Shortfall are an efficient way for managing the so-called Tail Risk; and since Expected Shortfall is more conservative than Value-at-Risk and due to the subadditivity property, it can be seen that seasonal contracts are generally better off than monthly contracts in reducing global risk. A worst case approach based on Value-at-Risk and Expected Shortfall developed into a Monte Carlo framework has been performed and led to a reduction of uncertainty when calculating weather derivatives prices. Empirical studies in the first and second papers are done under historical daily data for the temperature in Celsius from the weather station Molin Bianco in Arezzo, Tuscany, Italy.

- In the third paper we propose forecasting based on a stochastic model of the probability distribution. Indeed, we suggest to incorporate model uncertainty by considering forecasting using dynamical stochastic evolution of the probability distribution of the model in question. We apply our dynamic probabilistic forecasting to option pricing, where our proposed notion of model uncertainty reduces to uncertainty on future volatility. We have considered and compared the results of two different classes of autoregressive models where the so-called stationary distribution is and is not normal (where the Ornstein-Uhlenbeck process considered is driven by the Brownian motion and Levy process being a compound Poisson process, respectively).

- In the fourth paper we propose some functional linear models predicting wind speed from temperature data in which both the response and the covariate variables are functions. We have proposed the functional linear model observed in different patterns which is powerful enough to describe the dynamics of wind speed in time in different meteorological stations in Lithuania. The functional linear model, firstly, is used to predict wind speed from climate zone using functional principal component analysis. Secondly, annual wind speed is fitted by using temperature as a functional covariate, where the harmonic acceleration roughness in the regression coefficient function is penalized. Thirdly, the full wind speed profile is fitted by the regressing on the full temperature profile, using a level of smoothing by applying generalized cross-validation criterion. The validation procedure based on smoothed functional data of wind speed shows that the proposed models are reliable and can be used for various practical applications.

Part I

The First Part

Chapter 2

Scientific Report: The Financial Instruments for Mitigating and Hedging against the Climate Change

Abstract

Weather often has a significant impact on industrial sectors (such as energy and power industry, agriculture, construction, tourism, insurance and reinsurance companies, etc) making them exposed to weather risk. High exposure to weather risk may lead to financial stress, or even worse bankruptcy. Therefore, weather and its changes have driven a demand for weather risk management. Weather derivatives are contracts whose payoff depends on the temperature, rainfall, snowfall, humidity, sunshine or wind and can be used as part of risk management strategy to hedge risk associated with adverse or unexpected weather conditions. The management of weather risk with derivatives is a recent topic of scientific papers. This report presents and critically discusses a review of recent literature on financial and insurance products used to protect against climate change. Through comparisons of different methodologies proposed by different researchers, this study shows the necessity and importance of using weather derivatives and insurance contracts to better protect and develop world wide financial markets.

Keywords: Weather derivatives, risk management, complete and incomplete market, pricing and modeling

2.1 Introduction

The purpose of weather derivatives is to allow businesses and other organisations to insure themselves against fluctuations in the weather. On the other hand the purpose of insurance contracts and specifically of weather insurance products is to insure against extreme events. The weather derivatives market, in which contracts that provide this kind of insurance are traded, first appeared in the US energy industry in 1996 and 1997. Early trading of weather based instruments among energy companies started as over-the-counter (OTC[1]) trades. OTC trades are still used for weather derivatives for local cities which are not listed in exchanges. In September 1999 the first electronic marketplace for standardized weather derivatives was launched by the Chicago Mercantile Exchange (CME) with the aim of increasing liquidity, market integrity and accessibility. In the

[1] it means that each contract is individually negotiated

beginning, CME only traded two weather products namely Heating Degree Days (HDD) and Cooling Degree Days (CDD) for ten cities in the USA (CME Group, 2011). This market experienced rapid growth and in July 2010 CDD and HDD futures and options for 24 cities in the USA, 11 cities in Europe, and 6 cities in Canada, 3 cities in Japan and 3 in Australia were traded on the CME (Meissner and Burke (2011)). These include New York, Chicago, Philadelfia, London, Paris, Amsterdam, Essen, Stockholm, Barcelona, Rome, Madrid, Oslo, Prague and Berlin as well. But according to CME Group (2018) up to today only weather derivative contracts based on temperatures in London and Amsterdam are available as many other contracts disappeared since 2011; further only Cumulative Average temperature (CAT), HDD and CDD contracts are offered on the CME. An overview of weather derivatives traded on the CME is given by "http://www.cmegroup.com/trading/weather/". Except CME, weather derivatives contracts trade OTC on other exchanges such as Intercontinental Exchange, Swiss Re's ELRiX (electronic risk exchange), and LIFFE (the London International Financial Futures and Options Exchange).

Actually, CME offers HDD contracts from November to June and CAT and CDD from May to September. CDD contracts are only available in the summer months, when temperatures are most likely to be above 18 degrees celsius. In the USA, CME offers CDD contracts in May, June, July, August and September (Benth and Šaltytė Benth (2012)). But CDD written on some European cities (such as Reykjavik in Iceland and Dublin in Ireland) are not available since temperatures during summer are hardly above 18 degrees celsius, so CME offers CAT contracts in summer for these places while in winter HDD contracts are offered as well in USA, Japan and Europe and this kind of contract is available in winter only.

Other types of contracts based on frost days and snowfall are also traded on the CME. The weather derivatives markets are expanding rapidly as diverse industries seek to manage their exposure for weather risks. The notional value of CME weather products in 2004 was $\$2.2B$, and grew ten-fold to $\$22B$ through September 2005 with volume surpassing thousand of contracts traded. In 2006 the value of traded weather instruments rose to $\$45B$. Later, weather derivatives emerged in the financial market as an instrument for hedging various weather related risks in different sectors. According Härdle and Cabrera (2012) the key factor of using weather derivatives as instruments for hedging risk is that it could accelerate the pricing method. Still there is little liquidity in the market and weather is non-tradable, implying that the weather derivatives market is an incomplete market. Weather derivatives comprise of derivatives written on weather variables such as temperature, rainfall, snowfall, humidity, sunshine or wind. In the period of October 14, 1997 to April 15, 2001, the largest proportion of weather derivatives currently traded in the market are based on temperature with 98% of the overall traded volume while rain-related contracts accounted for 0.9%, snow 0.5%, and wind 0.2% (Garman, Blanco and Erickson (2000)). According to (Benth & Šaltytė Benth (2007) and Jewson & Brix (2005)) temperature derivatives contracts are written based on HDD, CDD, CAT and Pacific Rim indices.

A tradable weather derivative contract is defined by the following attributes: the contract type, the contract period, the contract station where the temperature or other index will be measured, the underlying index, the tick size (in CME the tick size is set \$20 which corresponds to a degree day index), the strike level and a payoff function.

The variables defining the payoff function varies according to the type of the contract which can be options, futures, forwards, swaps, collares, straddles, strangles and binaries (Jewson & Brix (2005)).

In this scientific report we attempt to describe all the methods and models currently in use in the weather derivative market. There are many ways of approaching the question of how to price a weather derivative contract; on top of this there are strong financial incentives to invent, new and more accurate methods for such pricing, and there is undoubtedly a lot of progress still to be made.

The rest of the report is organized as follows: section 2 discusses different types of weather based financial derivatives and focuses on world wide weather indices and explains why it is needed to develop and adopt weather indices. It compares methods used on modeling and pricing weather derivatives. In addition, a summary of pricing methods and models is done. Section 3 reviews the literature and the last section concludes the report and discusses about future research.

2.2 Methods

There are numerous methods suggested in the literature that we have chosen to review. These methods mainly compare different types of risk management for different products, including traditional weather insurance, emphasizing why the whole world today should develop and implement weather derivatives to better manage risk. For all the papers considered in this section, we will discuss features, advantages and disadvantages and we will explore different types of financial derivatives capable of managing disaster/fluctuation risk. Many details on how to calculate weather index, how to price and how to create weather derivatives markets in all over the world, are explained.

2.2.1 Weather indices

Weather changes involve investigation of daily meteorological variables (e.g. the future change of precipitation intensity, or the number of frost days, etc). Quantification of these events is based on climate indices derived from daily minimum, average, maximum values or daily sums of meteorological variables (mainly temperature and precipitation). Therefore, their investigation is essential for developing targeted adaptation plans or risk management tools capable of hedging the risk of climate change. Weather indices often describe extreme events that are associated with statistically rare values at the lower or upper tails of the distribution functions (e.g. extremely hot days occur only a few times in a year). However, some indices characterize frequent events, e.g. frost days happen through one-quarter of the year in a specific location. Should we use discrete or continuous distribution for weather variables?

2.2.1.1 Discrete or continuous distributions of weather variables

According to Jewson and Brix (2005) weather variables, such as temperatures can be considered to be a continuous variable, but measurements of this variable are typically rounded to a certain degree. In Europe, the measurements of daily minimum T_{min}

and daily maximum T_{max} temperatures are usually rounded to one decimal place in celsius and consequently T_{avg} (average of T_{min} and T_{max}) has two decimal places with the final digit being either a zero or a five. As a result of this rounding there is only a discrete number of possible outcomes for the measured temperature during a given period, and hence only a discrete number of possible index values can be achieved. This might lead one to conclude that all index distributions should be modelled using discrete distributions. However, the actual number of different possible index values is often very large, and fitting a discrete distribution to such data, and running simulations can be very slow. It is instead, often a reasonable approximation to use a continuous distribution. In general the rule will be: in any case where there are more than one hundred possible values for the index the continuous distribution is used; in all other cases the discrete distribution is the choice.

The most commonly used indices inform about the frequency of a variable of having values that are over/under a given threshold, e.g. an extremely daily maximum temperature is at least 35° celsius; dry periods are defined as consecutive days when daily precipitation is below $1mm$, a frost day when daily minimum temperature is less than 0 degree celsius and so on. Thanks to the fixed thresholds, these indices are easily quantifiable and also interpretable. However, it is not straightforward to use these indices for comparing areas with different climate because the frequency of a variable crossing a given threshold can vastly differ by regions. In order to partially resolve this problem, percentile-based indices were introduced. In this case fixed percentage values (usually the lower or upper deciles) of the distribution function are defined on a reference period to specify the threshold (Stefani et al (2018)), letting the frequency of going over this threshold be quantified for the future.

2.2.1.2 Index-weather formulas in discrete and continuous form

In this section, we will present the most commonly used weather indices used as the underlying values for derivative contracts. The models and the theoretical derivation of prices are most conveniently expressed in a continuous-time framework especially for temperature weather indices while for rainfall the discrete-time framework is mainly recommended. We restate the definitions of the different weather indices in discrete (D) and continuous (C) frameworks accordingly. In the markets, the discrete form for weather indices is always used (Benth and Šaltytė Benth (2012)). For instance, HDD and CDD futures are settled on a daily average temperature, hence discrete representation is suggested. But for theoretical considerations, sometimes it is convenient to formulate the weather indices in continuous form which somehow is an approximation of the market value. Therefore, if it is considered a daily time series model for temperature, then the discrete form of the weather index is suggested to be used. Moreover, if it is considered a continuous-time process, it is tempting to use a continuous representation of weather indices; but the discrete form could be used as well. Therefore, there is no specific criteria which clearly point out when to use a discrete/continuous form of weather indices. Note that more specialized indexes can be constructed to please the buyer of risk protection.

The cumulative average temperature (CAT), measures the sum of average temperature $T(i)$ in a given range $[t_1, t_2]$.

$$(D)\ CAT = \sum_{i=t_1}^{t_2} T(i), \quad (C)\ CAT = \int_{t_1}^{t_2} T(i)di$$

Pacific rim divides CAT index over its duration.

$$PRIM = CAT/(t_2 - t_1)$$

Heating degree days (HDD), and cooling degree days (CDD) measure the sum of deviations of $T(i)$ from a base temperature T_b, usually $65°$ Fahrenheit, or $18°$ Celsius, in a given range $[t_1, t_2]$ (Alexandridis Zapranis (2013)).

$$(D)HDD = \sum_{i=t_1}^{t_2} max(T_b - T(i), 0), \quad (C)HDD = \int_{t_1}^{t_2} max(T_b - T(i), 0)di$$

and

$$(D)CDD = \sum_{i=t_1}^{t_2} max(T(i) - T_b, 0), \quad (C)CDD = \int_{t_1}^{t_2} max(T(i) - T_b, 0)di$$

A cumulative rainfall index measures the sum of daily rainfall, R, on a specific date, i over period, [t_1, t_2].

$$(D)CRI = \sum_{i=t_1}^{t_2} R(i), \quad (C)CRI = \int_{t_1}^{t_2} R(i)di$$

As an alternative, Odening, Musshoff and Xu (2007) suggest a rainfall deficit index defined as:

$$(D)RDI = \sum_{i=t_1}^{t_2} \min(0, \sum_{j=(i-t_1)s+1}^{is} (R(j) - R^{min})),$$

$$(C)RDI = \int_{t_1}^{t_2} \min(0, \int_{(i-t_1)s+1}^{is} (R(j) - R^{min})dj)di$$

where R^{min} is the average s-day precipitation in the respective accumulation period. This index measures the shortfall of the rainfall sum in a s-days period relative to a reference level R^{min}.

The cumulative snowfall index measures the sum of daily snowfall, Y, on a specific date, i over period, [t_1, t_2] (Djordjevic (2018)).

$$(D)CSI = \sum_{i=t_1}^{t_2} Y(i), \quad (C)CSI = \int_{t_1}^{t_2} Y(i)di$$

The cumulative wind speed index measures the sum of daily average wind speeds during a period, where duration is defined as $[t_1, t_2]$, and $W(i)$ is day i average wind speed (Alexandridis & Zapranis (2013)).

$$(D)CWSI = \sum_{i=t_1}^{t_2} W(i), \quad (C)CWSI = \int_{t_1}^{t_2} W(i)di$$

The Nordix wind speed index aggregates the daily deviations from a 20 year mean over a specified period. $w_{20}(i)$ is the 20 year average wind speed on day i (Benth (2018)).

$$(D)NWSI = 100 + \sum_{i=t_1}^{t_2} (W(i) - w_{20}(i))$$

$$(C)NWSI = 100 + \int_{t_1}^{t_2} (W(i) - w_{20}(i))di$$

The historical evolution of the actual German wind power production index $GWPPI(t)$ is obtained as

$$GWPPI(t) = \frac{W(t)}{24C(t)}$$

where $W(t)$ denotes the total wind power production in Germany at day t and $C(t)$ denotes the total installed capacity in Germany at day t (Benth and Pircalabu (2018)).

For the frost index a multivariate approach is suggested by (Rozante, Gutierrez, da Silva Dias, de Almeida Fernandes, Alvim, and Silva (2019)). This index includes in its formulation different processes by which frost occurs such as surface temperature and moisture, sea-level pressure, low-level winds and cloudiness. The frost index (FI) is intended to monitor the chance of occurrence or non-occurrence of frosts, taking into account the contributions of some meteorological variables that directly influence the phenomenon such as temperature (T) and relative humidity (H), both at $2m$, wind speed at $10m$ (V), pressure reduced to mean sea-level (P) and cloudiness (N)—together with their means and standard deviations. For operational purposes, a normalized multivariate weighted contribution is used:

$$\begin{aligned} FI_{(i,j,k)} = w_1 \frac{\overline{T(i,j,k)} - T_m(i,j,k)}{\sigma_T(i,j,k)} &+ w_2 \frac{P_m(i,j,k) - \overline{P(i,j,k)}}{\sigma_P(i,j,k)} \\ &+ w_3 \frac{\overline{V(i,j,k)} - V_m(i,j,k)}{\sigma_V(i,j,k)} \\ + w_4 \frac{\overline{N(i,j,k)} - N_m(i,j,k)}{\sigma_N(i,j,k)} &+ w_5 \frac{\overline{H(i,j,k)} - H_m(i,j,k)}{\sigma_H(i,j,k)} \end{aligned}$$

where $\overline{P}$, $\overline{T}$, $\overline{V}$, $\overline{N}$, $\overline{H}$ are the means calculated in case of frost, σ_P, σ_T, σ_V, σ_N, σ_H are the standard deviations calculated in cases of frost, P_m, T_m, V_m, N_m, H_m are the variables extracted from the numerical model and w_i for $i = 1, ..., 5$ are the weights (constrained to obey $\sum_{i=1}^{5} w_i = 1$) attributed to each meteorological variable.

2.2.2 Weather derivative instruments

Weather contracts are based on the actual observation at one or more points in time on a weather variable or index. The weather derivative instruments are defined based on the underlying index and the measure of weather which governs how pay-out on the contract will occur. The structure of the derivatives are based on a standard derivative structure, which includes futures, forwards, swaps, options (puts and calls) and a combination of put and call options such as collars, straddles, and strangles.

2.2.2.1 Weather Futures

Weather futures are a type of weather derivative that obligate the seller of the derivative to either buy or sell at a time $T > 0$ and at a pre-specified strike price K some commodity. Such contracts enable to protect against losses caused by unexpected shifts in weather conditions (Cummins and Geman (1995)). This is a derivative which is used more and more by the energy companies for hedging against change in demand due to change in temperature. Future derivatives are used only to hedge risk associated with adverse weather. Their underlying variables include: HDDs, CDDs, Average Temperature, Frost, Snowfall, and Hurricanes.

The long futures contract payoff formula is: $payoff = S_T - K$
The short futures contract payoff formula is: $payoff = K - S_T$

where S_T is the price of the asset at the end of the contract, T is the maturity and K is the strike price.

2.2.2.2 Weather forwards

Contrary to spot contract where assets are bought or sold today, forwards contract is a contract in which two parties agree to buy or sell asset at a specified time at a predetermined price today. Forwards are used to hedge risk. Unlike futures, forwards are over-the-counter contracts.

The long forwards contract payoff formula is: $payoff = S_T - K$
The short forwards contract payoff formula is: $payoff = K - S_T$

where S_T is the price of the asset at the end of the contract, T is the maturity and K is the strike price.

2.2.2.3 Weather Options

As said before, there are two types of weather options: calls and puts. The buyer in the HDDs call pay a premium at the beginning of the contract and will benefit of all positive cash-flows the contract might produce untill expiration. If the number of HDDs is greater than predetermined strike level then, the call holder will receive a payout in return. The size of the payout is determined by the strike and the tick size, which is the amount of money that the holder of the call receives for each degree day above the strike level before maturity. In formulating the weather options, the two counterparts need to specify the following parameters:

- the contract type (call or put);
- the underlying index (HDDs or CDDs);
- the contract period;
- the weather station from which the temperature data are obtained;
- the strike level;
- and the tick size.

Normally the option price is stated in the option agreement, and the contract relieves the option holder any obligation to pay the liquidated amount to the writer for refusing to perform (Alaton et al (2002)).

- The call option represents the right to purchase an asset at a predetermined strike price before the option reaches its expiry date and it is purchased in the hope that the underlying asset price will rise. The payoff of a call option can be expressed as follows:
$$P = V(\max(I - K, 0))$$
where V is the tick size which can be determined as 1 EUR per index point, K is the strike level, T is the maturity and I is the index/underlying variable.
- The put option represents the right to sell an asset at a predetermined strike price before expiry date and it is purchased with the hope that the underlying asset price will drop below the strike price. The investor benefit from put options when the market price falls without having to sell short the underlying asset and has limited risk if market goes up against them. The payoff of a put option can be expressed as follows:
$$P = V(\max(K - I, 0))$$
where V is the tick size which can be determined as 1EUR per index point, K is the strike level, T is the maturity and I is the index/underlying variable.

2.2.2.4 Weather Swaps

Swaps are contracts in which two or more parties exchange risks during a predetermined period and the payment are made by both parties: one side paying a fixed price and the other paying a variable price. Swaps are over the counter instrument that can be modified or customized to suit the needs of the parties involved by protecting their business against weather uncertainties. In the case of the standard HDDs, the parties agree on a given strike of the HDDs for the period and the swapped amount. So that the swaps are costless: there is no premium, as the loss from a swap is equated with positive payoff it produces. The parties write a contract to pay each other at some point according to the weather related variable in future. In cases of long swap, the buyer has to pay the seller for low values of the index, thereby hedging against high value index. If the swap is traded without premium, then the strike should be set at a level where the expected payoff is close to zero thus shifting to compensate for the risk that are taking (Cummins and Geman (1995)). Once the swap has been traded, valuation consisting of the calculation of distribution of the possible financial outcome is done.

Swaps can vary depending on the need and the objective. For instance, energy distributors can use the weather swap to hedge against warmer than normal winter. According to Jewson and Brix (2005) the swap payoff is expressed as:

$$P = V(\mu - K)$$

where V is the tick size, K is the strike and μ is the mean of the weather index assumed normally distributed.

According to Broni-Mensah (2012) an uncapped swap can be represented with payoff given as follows:

$$P(I, T) = V(I - K)$$

where I is the value of a specified index at maturity time (T), K is the strike level, and V represents the cash value of one value movement in the sum $(I - K)$. When the maximum payouts for the two counterparties are not equal, the payoff will be as follows:

$$P(I, T) = \begin{cases} V(L_1 - K) & \text{if } I_H(T) < L_1 \\ V(I_H(T) - K) & \text{if } L_2 < I_H(T) < L_2 \\ V(L_2 - K) & \text{if } I_H(T) > L_2 \end{cases}$$

where $I_H(T)$ is the value of HDD index at maturity, V is used to translate the quantity $(I_H(T)K)$ into monetary terms, K is the strike level, and L_1 and L_2 denote the level at which a limit to the payoff is applied. As the swap structure is not symmetrical we seek to determine the value of the strike K that satisfies the equation

$$H(K) = 0$$

with H denoting the expected payoff from a swap contract.

2.2.2.5 Weather Collar

Collar is a strategy that limits the range of loss of earnings on an underlying asset to a specific range. It is a combination of a call and put option enabling one to use a long put or call option with a particular strike by financing it with a short call or put option with a different strike (Cummins and Geman (1995)). This strategy provides the user with the price protection against adverse weather events in the underlying asset thereby, forcing price to move within a defined range. Notably, this is provided at the expense of giving up some of the returns. Importantly the collars contain the premium for one of the parties. The payoff of a collar can be written as follow:

$$P = min(0, min(V(I - K_p), max(0, min(V(I - K_c), 0))))$$

where V is the tick size which can be determined as 1 EUR per index point, K_c is the strike level for a call option, T is the maturity, K_p is the strike level for a put option and I is the index/underlying variable.

2.2.2.6 Weather Strangle

Strangle is an investment strategy involving the purchase or sale of option derivative that allows the holder to profit based on how much the price of the underlying security moves, with relative exposure to the direction of the price movement: the purchase is known as a long strangle while the sale of the derivative is a short strangle. The strangle option is built on a portfolio of put and call option (Jewson and Brix (2005)) thereby, reducing the net debit of the trade and the risk of the loss in the trade although, their strike is different. The owner of a long strangle makes a profit if the underlying price moves far enough away from the current price and it is important that the investor should take a long strangle if the security is volatile. The payoff of a strangle can be expressed as follow:

$$P = V(\max(K_p - I, I - K_c, 0))$$

where V is the tick size which can be determined as 1 EUR per index point, K_c is the strike level for a call option, T is the maturity, K_p is the strike level for a put option and I is the index/underlying variable.

2.2.2.7 Weather Straddle

A straddle is an option strategy involving the purchase of both a put and call option for the same expiration date and strike price on the same underlying resulting to a significant profit in the market direction (Cummins and Geman (1995)). The strategy is profitable only when the stock either rises or falls from the strike price by more than the total premium paid. It gives the trader a greater profit than a butterfly if the final stock price nears the exercise price (Alaton et al (2002)). The payoff of a straddle can be expressed as follow:

$$P = V(\max(K_p - I, I - K_c, 0)), K_c = K_p$$

where V is the tick size which can be determined as 1 EUR per index point, K_c is the strike level for a call option, K_p is the strike level for a put option, T is the maturity and I is the index/underlying variable.

2.2.3 Pricing Approaches

One of the main areas of investigation in the weather derivatives markets is the choice of the pricing methodology to use in order to obtain the “fair” value for different contracts. It is also important that a good pricing approach can be used to increase the transparency, market confidence and further develop weather derivatives into a robust thriving market. But before going for pricing approaches let us discuss about the differences between incomplete and complete markets.

2.2.3.1 Incomplete vs complete market

A market can be incomplete in many different ways (Björk (2009)), and below are mentioned some of them:

- There are more random sources than there are risky underlying assets.

- There are constraints on admissible portfolios, like, for example, a short selling constraint.
- The underlying variable is not traded, like in the case of weather derivative contracts.
- The underlying is traded but the market is not liquid, mention here the case of commodity markets.
- The underlying is traded but the portfolios can not easily and/or without large costs be carried forward in time, like, for example electricity derivatives.

As to financial valuation principles, the Black–Scholes method is the most successful pricing approach in the area of derivatives. This method is based on a strategy in which one creates a portfolio that accurately replicates the payoff of the derivative. The risk associated with the financial derivative is thereby completely eliminated or hedged. Thus, one can argue that the value of a product must be the cost of setting up the hedging portfolio, based on the no-arbitrage principle. The Black–Scholes method has been a landmark in derivative pricing in the complete market. A financial market is complete if all claims are attainable, i.e. if all claims can be replicated by means of a self-financing strategy. If claims, which are not attainable, exist and hence cannot be replicated by means of any self-financing trading strategies, then the market is incomplete. The most distinctive feature of weather derivatives is that, unlike traditional financial derivatives, their prices are linked to a weather event rather than the price of an underlying security or commodity. The commonly used underlying weather indexes, for instance, HDDs and CDDs are non-tradable. On the other hand, there is typically little or no liquidity in weather derivatives. Therefore, weather derivatives market is an incomplete market because the underlying variable in this case, weather is non-tradable and cannot be replicated in any means of self-financing trading activities. In addition, Stojanovic (2011) suggests to study both the complete and incomplete markets, with emphasis on incomplete markets, while complete markets must be viewed as special cases of the incomplete ones. Since incomplete financial market models acknowledge both the hedgeable and unhedgeable risks, then these models are more appropriate for the valuation of weather derivatives. Below we describe some of the more popular pricing models currently being used.

2.2.3.2 Market Price of Risk (MPR)

It is crucial, for the accurate pricing of the various weather derivatives and hedging of the weather risk, the understanding of the MPR. Weather derivatives market is not liquid enough. Furthermore, as just said weather market is not complete, because the underlying is not tradable. This is the reason why the MPR should be considered when valuing weather contracts. In most studies (focused on temperature) so far, the MPR was considered zero (Alaton et al (2002)). However, recently a number of studies examined the MPR and found that it is different than zero. In Møller (2001) one can find numerical choices for the MPR, for which it is stated that the larger the value of MPR the greater the level of compensation awarded. Cao and Wei (2004) apply a generalized Lucas'

(1978) equilibrium pricing model to study the MPR and show that the market price associated with the temperature is significant and represents a considerable portion of the derivative's price. These authors also demonstrate that MPR varies over time, and the assumption of being constant does not have an empirical argumentation. In Xu et al (2008), an indifference pricing approach which is also based on utility maximization is proposed. The most common approach is the one presented in Alaton et al. (2002), and it was followed by Bellini (2005), Benth et al. (2009), and Hardle and Cabrera (2009).

Alaton et al. (2002) suggest that the MPR can be estimated from the market data: by examining what value of MPR gives a price from the theoretical model that fits the observable market price. In Bellini (2005), the implicit MPR is estimated by comparing theoretical future prices, given in previous formulas, to the prices observed in the market under the assumption of a Levy noise process where a time dependence of the MPR is examined. In Hardle and Cabrera (2009), the implied MPR of Berlin was estimated. Their results indicate that the MPR for CAT derivatives is different from zero and shows a seasonal structure that increases as the expiration date of the temperature future increases. Benth et al. (2011) study the MPR in various Asian cities, for which the MPR was estimated by calibrating model prices. Their results indicate that the MPR for Asian temperature derivatives displays a seasonal structure that comes from the seasonal variance of the temperature process.

2.2.3.3 Black-Scholes model

The Black-Scholes model is based on certain assumptions that do not apply realistically to weather derivatives. This derivative pricing model is nonetheless so important in asset pricing that deserves to be presented. One of the main assumptions behind the model is that the underlying of the contract (HDD or CDD) follows a random walk without mean reversion (tends to revert to normal levels within a specific period). Since temperatures are mean reverting, the Black-Scholes model is inadequate for modeling temperature (Garman et al (2000)). For this reason, according to Biganashvili (2013), weather derivatives market needs a standard pricing model so all participants can start communicating in a common language. The large discrepancies between the different models used are preventing the market from developing at an even faster pace. As the Black-Scholes model for financial derivatives was one of the main driving factors of the option markets in the 1980s, the weather markets need a common denominator for today's markets. On the other hand, Meissner and Burke (2011) claim that the Black-Scholes-Merton model despite its limitations is an adequate pricing model for weather derivatives, which is superior to burn analysis approach. In general the Black-Scholes model describes the evolution in a risk-neutral world of the price of weather derivatives through the stochastic differential equation:

$$\frac{dWD(t)}{WD(t)} = rdt + \sigma dW(t) \tag{2.1}$$

where $W(t)$ is a standard Brownian motion, σ is the volatility and r is the interest rate. The solution of equation (6.1) is given as:

$$WD(T) = WD(0) * exp((r - \frac{1}{2}\sigma^2)T + \sigma W(T)) \tag{2.2}$$

$WD(0)$ is the current price of the underlying asset, while the random variable $W(T)$ is normally distributed with mean 0 and variance T, which is also the distribution of $\sqrt{T}Z$ if Z is standard normal random variable. Therefore equation (6.2) can be written as:

$$WD(T) = WD(0) * exp((r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}Z) \tag{2.3}$$

The logarithm of underlying price is thus normally distributed, and the underlying price itself has a lognormal distribution. An application of Black-Scholes model on temperature options is explained in Meissner and Burke (2011). Since HDDs and CDDs are the underlying in temperature options, the payoff equation for options on HDDs and CDDs can be expressed as:

$$\begin{aligned} C(HDD) &= e^{-rt}E[\max(F_0(HDD) - K, 0)] \\ &= e^{-rt}(F_0(HDD)N(d_1) - KN(d_2)) \end{aligned} \tag{2.4}$$

$$\begin{aligned} P(HDD) &= e^{-rt}E[\max(K - F_0(HDD), 0)] \\ &= e^{-rt}(KN(-d_2) - F_0(HDD)N(-d_1)) \end{aligned} \tag{2.5}$$

$$\begin{aligned} C(CDD) &= e^{-rt}E[\max(F_0(CDD) - K, 0)] \\ &= e^{-rt}(F_0(CDD)N(d_1) - KN(d_2)) \end{aligned} \tag{2.6}$$

$$\begin{aligned} P(CDD) &= e^{-rt}E[\max(K - F_0(CDD), 0)] \\ &= e^{-rt}(KN(-d_2) - F_0(CDD)N(-d_1)) \end{aligned} \tag{2.7}$$

where $d_1 = \frac{\ln(\frac{F_0(HDD/CDD)}{K}) + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}$ and $d_2 = d_1 - \sigma\sqrt{T}$.

Equation (4) implies that the colder it is (the more the number of HDDs) the higher is the payoff of the call. Thus, the investor would buy a call option on HDDs if the investor believes it will be colder than the current future price $F_0(HDD)$ implies. Equation (5) instead implies that the warmer it is (the fewer the number of HDDs), the higher is the payoff of the put. Hence, an investor would buy a put on CDD, if the investor believes it will be warmer that the current future price $F_0(HDD)$ implies.

In addition, one can apply sensitivities on Black-Scholes model and as a result obtain what follows:

- Delta: is defined as the rate of change of the price of a HDD with respect to the price of the underlying future price $F_t(HDD)$.

  Call: $\triangle = \frac{\partial C(HDD)}{\partial F_0(HDD)} = e^{-rt}N(d_1)$
  Put: $\triangle = \frac{\partial P(HDD)}{\partial F_0(HDD)} = e^{-rt}(N(d_1) - 1)$

- Gamma: The second derivative of the call/put option with respect to the price of the underlying future price $F_0(HDD)$ is called the Gamma of the option and is given by

  $\Gamma = \frac{\partial^2 C(HDD)}{\partial F_0^2(HDD)} = \frac{\partial^2 P(HDD)}{\partial F_0^2(HDD)} = \frac{e^{-rt}e^{-\frac{d_1}{2}}}{F_0(HDD)\sigma\sqrt{t}\sqrt{2\pi}}$

- Theta: Theta is the rate of change of the price of a HDD with respect to time to maturity with all else remaining the same.

  Call: $\Theta = \frac{\partial C(HDD)}{\partial \tau}$

  $\Theta = \frac{1}{T}(-(\frac{F_0(HDD)\sigma e^{-rt}e^{-\frac{d_1^2}{2}}}{2\sqrt{t}\sqrt{2\pi}})) - rKe^{-rt}N(d_2) + rF_0(HDD)e^{-rt}N(d_1)$

  Put: $\Theta = \frac{\partial P(HDD)}{\partial \tau}$

  $\Theta = \frac{1}{T}(-(\frac{F_0(HDD)\sigma e^{-rt}e^{-\frac{d_1^2}{2}}}{2\sqrt{t}\sqrt{2\pi}})) + rKe^{-rt}N(-d_2) - rF_0(HDD)e^{-rt}N(-d_1)$

- Vega: The vega of a weather derivative, is defined as the rate of change of its price with respect to the volatility of the underlying asset.

  $\Lambda = \frac{\partial C(HDD)}{\partial \sigma} = \frac{\partial P(HDD)}{\partial \sigma} = \frac{F_0(HDD)e^{-rt}\sqrt{t}e^{-\frac{d_1^2}{2}}}{100\sqrt{2\pi}}$

- Rho: The rho of a weather derivative is defined as the rate of change of its price with respect to the interest rate.

  Call: $\rho = \frac{\partial C(HDD)}{\partial r} = \frac{1}{100}Kte^{-rt}N(d_2)$

  Put: $\rho = \frac{\partial P(HDD)}{\partial r} = -\frac{1}{100}Kte^{-rt}N(-d_2)$

2.2.3.4 Actuarial Methods

These methods use conditional expectation of the weather derivatives future payoffs to calculate the price. This, fundamentally is the same method used by insurance companies. When pricing risks different probabilities and statistical analysis are required for different events to be insured. Based on the historical probabilities an insurance premium is calculated accordingly. These methods are less applicable for weather derivatives as the underlying variable such as temperature, rainfall, snowfall, wind, etc tend to follow a recurrent, predictable pattern (Stefani et al (2018)). If the contract was to insure for extreme conditions such as extreme heat or cold periods, then the actuarial method would be useful. It is argued that this is the only appropriate method for extreme weather conditions (Cao, Li and Wei (2003)). Moreover, the estimated expected payoff is in the real world, meaning that the actuarial approach is correct only when the expected payoff from the derivative is the same in both the real and the risk-neutral world (Hull 2003, 2005). An actuarial method with normal distribution was used in the paper of Beyazit and Koc (2010) to price put/call European options:

$$Call = Ve^{-r(T-t)}E[max(H_T - K, 0)],$$

$$Put = Ve^{-r(T-t)}E[max(K - H_T, 0)]$$

where V is the tick size, r the risk free rate, T the expiration time, K the strike price and H_T the cumulative HDD/CDD index at maturity. If the index follows a normal distribution, the put/call price functions will be expressed as:

$$Call = Ve^{-r(T-t)}((\mu_n - K)(\phi(-\alpha_n) - \phi(-\frac{\mu_n}{\sigma_n})) + \frac{\sigma_n}{2\pi}(e^{-\frac{\alpha_n^2}{2}} - e^{-\frac{1}{2}(\frac{\mu_n}{\sigma_n})^2}))$$

$$Put = Ve^{-r(T-t)}((K-\mu_n)(\phi(\alpha_n)-\phi(-\frac{\mu_n}{\sigma_n}))+\frac{\sigma_n}{2\pi}(e^{-\frac{\alpha_n^2}{2}}-e^{-\frac{1}{2}(\frac{\mu_n}{\sigma_n})^2}))$$

where μ_n denote the mean, σ_n standard deviation of average annual temperature, $\phi(.)$ the cumulative distribution function for standard normal distribution, and $\alpha_n = \frac{K-\mu_n}{\sigma_n}$.

2.2.3.5 Burn Analysis

This approach is commonly used in the insurance industry and essentially uses a simulation using historical information to estimate uncertain weather related payments. The procedure of this method starts by collecting the historical weather data. After this being done, the method determines what would have been paid out from the option for every year in the past and concludes by making an average of these amounts. This approach is easy to implement and understand and in the valuation of complex transactions involving correlated weather indices, the correlation is embedded into historical data. However, if an extreme event is included in the data, it can distort the results of the analysis as it tends to omit low frequency extreme events (Spillet (2001)). The main advantage of burn analysis in comparison with other methods is that it does not include any form of weather (temperature) forecasting. From what we stated above, it is quite common that market participants use this method in order to get the first notion about the fair price of an option. According to Schiller, Seidler, and Wimmer (2012), burn analysis approach is the simplest method for evaluating weather derivatives (despite all its simplifications, it is used by many market traders). The main idea of burn analysis is to calculate the future payoff of a derivative by considering the payoffs the same derivative yielded in the past. If, for example, a derivative for measurement period $[\tau_1, \tau_2]$ should be priced for the year $n+1$, we would calculate the fictive indices the same derivative had in the years $n, n-1, n-2$, etc. This yields a series $Y_1, Y_2, ..., Y_n$ of n indices for the past n years. Using the linear model

$$Y_t = \beta_0 + \beta_1 * t + \epsilon_t, t = 1, ..., n$$

then, the constant (intercept) parameter β_0 and the trend (slope) parameter β_1 can be estimated as:

$$\widehat{\beta_1} = \frac{\sum_{t=1}^{n}(t-\frac{n}{2})(Y_t-\overline{Y})}{\sum_{t=1}^{n}(t-\frac{n}{2})}$$

$$\widehat{\beta_0} = \overline{Y} - \frac{n}{2}\widehat{\beta_1}$$

where $\overline{Y} = \frac{1}{n}\sum_{t=1}^{n} Y_t$ is the mean of the calculated indices over the past n years. Three assumptions are established:

- The expected error$E(\epsilon_t) = 0$ for all years $t = 1, ..., n+1$.
- The variance of the errors $Var(\epsilon_t) = \sigma^2$ is constant for all years $t = 1, ..., n+1$.
- The covariance of the errors $Cov(\epsilon_t, \epsilon_s) = 0$ for all years $t \neq s$.

Under these assumptions, by the Gauss–Markov theorem, the estimator $\widehat{Y}_t = \widehat{\beta}_0 + \widehat{\beta}_1 t$ is the best linear unbiased estimator for Y_t. Hence, the index Y_{n+1} of the next year $n+1$ can be predicted as

$$\widehat{Y}_{n+1} = \widehat{\beta}_0 + \widehat{\beta}_1 * (n+1)$$

In order to derive a measure of the certainty of the prediction $\widehat{Y}_{n+1}$, a fourth assumption need to be established:

- The errors $\epsilon_t, t = 1, ..., n+1$, are independent identically normally distributed.

2.2.3.6 Index Modeling Method

This is a pricing method for weather derivatives based on a distribution statistically modeling the claim size. Once claim sizes have been observed, we select and fit a distribution. Then the mean of this distribution can be computed to find the expected value of the claim. 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, quantiles of the claim outside the range of observed data can be estimated. According to Schiller, Seidler, & Wimmer (2012) the approach extends the burn analysis method by estimating the distribution of the weather index. If the distribution can be estimated relatively well, the index modeling approach yields a more stable price estimation than the Burn Analysis. In addition, the fourth assumption added to burn analysis approach extends the burn analysis to an index modeling approach, since $\epsilon_i \sim N(0, \sigma^2)$ implies $Y_i \sim N(\beta_0 + \beta_1 i, \sigma^2)$. With this assumption one can use the well-known theory of linear models (Rencher 2008) to estimate the variance of the error of the prediction $\widehat{Y}_{n+1}$:

$$Var(\widehat{Y}_{n+1} - Y_{n+1}) = \frac{(n+2)(n+1)(n-2)}{n(n-1)(n-4)} s^2$$

$$s^2 = \frac{1}{n-2} \sum_{i=1}^{n} (Y_i - \widehat{Y}_i)^2$$

is the unbiased estimate for the variance σ^2 of the errors.

2.2.3.7 Monte Carlo Simulations

This method consists on a computer-based generation of random numbers which can be used to statistically construct weather scenarios. Generally the payoffs of an instrument are determined by simulating numerous weather scenarios based on HDDs or CDDs. As a result, the fair price is the average of all possible payoffs approximately discounted to account for the time value of money. Let turn back to the application of Black-Scholes formula in weather derivatives. From equation 6.3 to draw samples of the terminal weather derivative price $WD(T)$ it suffices to have a mechanism for drawing samples from standard normal distribution. First let produce a sequence of $Z_1, Z_2, ...$ of independent standard normal random variables. Given a mechanism for generating the Z_i, one can estimate $E(e^{-rt}max(WD(T) - K, 0))$ using the following algorithm:

$$for\ i = 1, ..., n$$
$$generate:\ Z_i$$
$$calculate:\ WD_i(T) = WD(0) + exp((r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}Z_i)$$
$$calculate:\ C_i = e^{-rt}max(WD(T) - K, 0)$$

$$calculate:\ \widehat{C}_n = \frac{C_1 + ... + C_n}{n}$$

for any $n \geq 0$, the estimator $\widehat{C}_n$ is unbiased, in the sense that its expectation is the target quantity:

$$E[\widehat{C}_n] = C \equiv E[e^{-rT}max(WD(T) - K, 0)]$$

The estimator is strongly consistent, meaning that when $n \to \infty$

$$\widehat{C}_n \to C$$

with probability 1. For finite but at least moderately large n, the point estimate $\widehat{C}_n$ is supplemented with a confidence interval. Let

$$s_C = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(C_i - \widehat{C}_n)^2}$$

denote the sample standard deviation of $C_1, ..., C_n$ and let z_δ denote the $1 - \delta$ quantile of the standard normal distribution ($\Phi(z_\delta) = 1 - \delta$). Then

$$\widehat{C}_n \pm z_{\delta/2}\frac{s_C}{\sqrt{n}}$$

is an asymptotically, as $n \to \infty$, valid $1 - \delta$ confidence interval for C. This is a simple application of Monte Carlo simulation approach using Black -Scholes model on weather derivatives according to Glasserman (2013).

2.2.3.8 Indifference Pricing Method

This approach is a utility based approach which has been presented by Brockett et al (2006) and Xu et al (2008) and it is different from other pricing methods because it is based on the basic principle of equivalent utility and according to Alexandridis and Zapranis (2013) it makes use of investors risk preferences and a corresponding utility function. The method uses the expected utility to produce indifference prices. One utility function that can be used for this approach is an exponential utility function:

$$U(X) = -e^{-\lambda X}$$

where $X \in \mathbb{R}$ and $\lambda > 0$. But the utility function can be a mean-variance utility function as well such as:

$$U(X) = E(X) - \lambda\sigma^2(X)$$

with the risk aversion parameter $\lambda > 0$. This method uses two market participants a seller (e.g. insurance company) and a buyer as agent seeking protection. At the beginning both actors optimize their investment portfolios in order to maximize wealth at maturity. Differently from other approaches this method is less ambitious since it does not attempt to predict a transacted market price. Instead, this approach calculates price boundaries for seller's and buyer's and states if transactions are likely to occur or not. To illustrate the indifference pricing approach, let consider a dynamic market setting that consists of two assets, a weather derivative with a price process $WD = (WD_t)_{0 \le t \le T}$, where T is a fixed time horizon, and a nontraded asset Y on which a European-type claim is written. The payoff of this European derivative is denoted by $g(Y_t)$, payable at time T. Moreover, no trading of the derivative is allowed after its inscription/purchase. The individual risk preferences are modeled via a utility function u. In this model, the investor seeks to maximize the expected utility of terminal wealth with and without talking into account the European claim. The initial wealth of the investor is denoted by w. The optimisation problem without considering this claim is a classical Merton (1976) model of optimal investment, namely,

$$V(w) = sup_\nu Eu[w + \int_0^T \nu_t dWD_t]$$

where $V(w)$ is the attainable maximum expected utility of terminal wealth with initial wealth w. Considering the possibility of buying/selling δ units of this claim, the buyer and seller's optimisation problems are defined respectively by

$$V^b(w) = sup_\nu Eu[w + \int_0^T \nu_t dWD_t - \delta\pi^b + \delta g(Y_T)]$$

and

$$V^s(w) = sup_\nu Eu[w + \int_0^T \nu_t dWD_t + \delta\pi^s - \delta g(Y_T)]$$

where π^b and π^s denote the prices of buying and selling one unit of the claim, respectively.

The indifference seller's (buyer's) price of the European claim $g(Y_T)$ is defined as $\pi^s(\pi^b)$, such that the investor is indifferent to the following two scenarios: optimizes the expected utility without using the derivative and optimizes his or her expected utility taking into account, on the one hand, the liability (payoff) $g(Y_T)$ at expiration T, and on the other hand, the compensation π^s (cost π^b). Therefore, the indifference prices $\pi^s(\pi^b)$ must satisfy

$$sup_\nu Eu(w + \int_0^T \nu_t dWD_t) = sup_\nu Eu[w + \int_0^T \nu_t dWD_t + \delta\pi^s - \delta g(Y_T)])$$

and

$$sup_{\nu}Eu(w + \int_0^T \nu_t dWD_t) = sup_{\nu}Eu[w + \int_0^T \nu_t dWD_t - \delta\pi^b + \delta g(Y_T)])$$

2.2.3.9 McIntyre Pricing Method

A simple analytical model for pricing weather derivatives is presented by McIntyre and Doherty in 1999. These authors assume that the cumulative weather variable (which can be temperature, precipitation, etc) is normally distributed with mean m and standard deviation σ. The McIntyre model for pricing weather call and put options derivatives is expressed:

$$Call = (m - K)N((m - K)/\sigma) + \sigma^2 f(K)$$

$$Put = (K - m)N((K - m)/\sigma) + \sigma^2 f(K)$$

where m is the mean weather variable (for instance, temperature), K is the strike price, $N(.)$ is the cumulative standard normal distribution, σ is the standard deviation, and $f(.)$ is the probability density function for a standard normal random variable defined as

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp^{-\frac{(x-m)^2}{2\sigma^2}}$$

The larger the volatility the larger the movement and the greater risk and as a consequence a higher price for the option. The mean of the underlying weather derivative indicates the price maker's expectation of future observations and should take into account recent trends, forecasts and positions. The implied volatility and implied mean together represent the risk and the purchase price of the option.

2.2.3.10 Stochastic Process

In this approach a stochastic differential equation is chosen to represent the diffusion of the weather index. The process is calibrated to either historical data sets or market quotes for weather derivatives, should they exist. The equation is then solved using the boundary conditions provided by the payment terms of the derivative transaction. Common features of the processes chosen would be mean reverting or auto-regressive processes. The one main advantage of this method is that the risk statistics are easily expressed. In Tables [1-7] a list of pricing models for the dynamics of temperature, rainfall, maize yield using basket (rainfall and temperature) weather derivatives, snowfall and wind, respectively, are displayed.

Note: All the notations in models presented in Tables [1-7] are as in the reviewed papers.

2.3 Literature Review

Weather affects worldwide economy. Global warming has led to more extreme weather phenomena and wider weather fluctuations than expected.

Table 2.1: Evolution of pricing models for the dynamics of temperature variations. (Part 1)

Authors	Year	Model
Alaton et al	2002	$dT(t) = dS(t) + \beta(T(t) - S(t))dt + \sigma(t)dB(t), \beta \in \mathbb{R}$ $S(t) = a + bt + c\sin(2t\pi/365 + \nu)$ $\sigma^2(t) = \frac{1}{N_\mu}\sum_{t=0}^{N_\mu - 1}(T(t+1) - T(t))^2$ N_μ is the number of days corresponding to months $\mu = 1, ..., 12$
Brody et al	2002	$dT(t) = \beta(t)(\theta(t) - T(t))dt + \sigma(t)dB^H(t)$ B^H-fractional Brownian motion [2] $T(t) = \frac{1}{2}(T_{min}(t) + T_{max}(t))$ $\theta(t)$ - the expected value at which the temperature reverts at a rate $\beta(t)$ on day t . $\sigma^2(t) = \frac{1}{N}\sum_{t=1}^{N}\overline{T}^2(t)$
Campbell & Diebold	2002	$T_t = m_t + s_t + \sum_{l=1}^{L}\rho_{t-l}T_{t-l} + \tilde{\epsilon}_t, t = 1, 2, ..., \tilde{\epsilon}_t = \sigma_t\epsilon_t$ $\epsilon_t \sim iid(0, 1)$ $m_t = \sum_{m=0}^{M}\beta_t t^m$ $s_t = \sum_{p=1}^{P}(\sigma_{c,p}cos(2\pi p\frac{d(t)}{365}) + \sigma_{s,p}sin(2\pi p\frac{d(t)}{365}))$ $\sigma_t^2 = \sum_{p=1}^{Q}(\gamma_{c,q}cos(2\pi q\frac{d(t)}{365}) + \gamma_{c,q}sin(2\pi q\frac{d(t)}{365}))$ $+\sum_{r=1}^{R}\alpha_r * (\sigma_{t-r}\varepsilon_{t-r})^2 + \sum_{s=1}^{S}\beta_s\sigma_{t-s}^2$ $d(t)-$ repeating step function that cycles through $1, ..., 365$ $L = 25, M = 1, P = 3, Q = 3, R = 1, S = 1$
Benth & Šaltytė Benth	2005	$dT(t) = dS(t) + \beta(T(t) - S(t))dt + \sigma(t)dL(t)$ L(t) - Levy process [3] $S(t) = a + b\cos(\frac{2\pi}{365}(t - t_0))$ $\tilde{\varepsilon}_t = \sigma_t\varepsilon_t$, $\varepsilon_t \sim$ iid $\sigma_t^2 = \sum_{q=1}^{Q}[\gamma_{c,q}cos(\frac{2\pi qd(t)}{365}) + \gamma_{s,q}sin(\frac{2\pi qd(t)}{365})] + \sum_{r=1}^{R}\alpha_r\tilde{\varepsilon}_{t-r}^2$
Benth & Šaltytė Benth	2007	$dT(t) = dS(t) - \beta(T(t) - S(t))dt + \sigma(t)dB(t)$ $S(t) = a + bt + \sum_{i=1}^{I_1}a_i\sin(2i\pi(t - f_i)/365)$ $+\sum_{j=1}^{J_1}b_j\sin(2j\pi(t - g_j)/365)$ $\sigma^2(t) = c + \sum_{i=1}^{I_2}c_i\sin(2i\pi t/365) + \sum_{j=1}^{J_2}d_j\sin(2j\pi t/365)$
Benth & Šaltytė Benth	2011	$T(t) = \Lambda(t) + Y(t)$ $\Lambda(t) = a + bt + c\sin((2\pi(t - d))/365),$ $Y(t) = e_1^{'}X(t) \sim CAR(p)$ $dX(t) = AX(t)dt + \sigma e_p dB(t)$ $\sigma^2(t) \sim V(t), dV(t) = -\lambda V(t) + dL(t)$

Table 2.2: Evolution of pricing models for the dynamics of temperature variations. (Part 2)

Authors	Year	Model
Benth & Šaltytė Benth	2012	$T(t) = \mu(t) + \epsilon(t)$ $\mu(t) = S(t) + \sum_{i=1}^{p} \alpha_i (T(t-i) - S(t-i), \epsilon(t) = \sigma(t)\delta(t)$ $S(t) = a_0 + a_1 t + \sum_{i=1}^{J} b_{1j} \cos(2\pi j(t - b_{2j})/365)$
Taib & Benth	2012	$T(t) = S(t) + Y(t)$ $Y(t) = \sum_{i=1}^{p} \alpha_i Y(t-i) + \epsilon(t), \epsilon(t) \sim N(0, \sigma^2)$ $S(t) = a_0 + a_1 t + a_2 \sin(\frac{2\pi(t-a_3)}{365})$
Wang, Li, Li, Huang & Liu	2015	$dT_t = dT_t^m + \alpha(T_t^m - T_t) + \sigma_t dW_t$ W_t - Wiener process $T_t^m = A + Bt + Csin(\frac{2\pi t}{365} + \phi)$ $\sigma_t^2 = c + \sum_{i=1}^{I} c_i \sin(\frac{2i\pi t}{365}) + \sum_{j=1}^{J} c_j \cos(\frac{2j\pi t}{365})$ $\alpha = -\log(\frac{\sum_{i=1}^{n}((T_{i-1} - T_{i-1}^m)/\sigma_{i-1}^2)(T_i - T_i^m)}{\sum_{i=1}^{n}((T_{i-1} - T_{i-1}^m)/\sigma_{i-1}^2)(T_{i-1} - T_{i-1}^m)})$
Szabȯ & Szėpszȯ	2016	$X_T(t) = \triangle T(t) + \overline{T} + \epsilon(t)$ $\overline{T}$ – temperature polynomial averages for the reference period $\triangle T(t)$ – the departure of polynomial values from this reference with t time $\epsilon(t)$ – the residuals of raw model results from the polynomial fits
Kabaivanov & Markovska	2017	$dT(t) = \theta(\mu - T(t))dt + \sigma dW(t)$ Mean reversion speed θ and mean μ are constant values
Kabaivanov & Markovska	2017	$T(t) = f(t) + S(t)$ $f(t) = p_1 \sin(2\pi t) + p_2 \cos(2\pi t) + p_3 \sin(4\pi t) + p_4 \cos(4\pi t) + p_5$ $dS(t) = \theta(\mu - S(t))dt + \sigma dW(t) + J(\mu_j, \sigma_j) d\Pi(\alpha)$ $S(t)$ – the mean-reverting process with jumps that are driven by a Poisson process with density α

Table 2.3: Evolution of pricing models for the dynamics of temperature variations. (Part 3)

Authors	Year	Model
Prabakaran & Singh	2017	$dT_t = (\frac{dT_t^m}{dt} + b(T_t^m - T_t))dt + dL_t$ $T_t^m = A + Bt + Csin(wt + \phi)$ $dL_t = \sigma_t dW_t + dY_t + dZ_t, L_t \sim ARCH(1)$ $dY_t = -\alpha Y_t dt + dQ_t, dZ_t = -\beta Z_t dt + dR_t$ $Q_t = \sum_{i=1}^{N_i^Y} U_i, U_i \sim N(\mu_Y, \delta_Y^2),$ $R_t = \sum_{i=1}^{N_i^Z} V_i, V_i \sim N(\mu_Z, \delta_Z^2)$ dY_t - fast mean-reverting OU process driven by compound Poisson processes dZ_t - slow mean-reverting OU process driven by compound Poisson processes
Evarest, Berntsson, Sigull & Yang	2018	$T_d(t) = \begin{cases} T_{t,1} : dT_{t,1} = (\mu_1 - \beta T_{t,1})dt + \sigma_1 T_{t,1} dW_t, & \text{with prob } q_1 \\ T_{t,2} : dT_{t,2} = \mu_2 dt + \sigma_2 dW_t, & \text{with prob } q_2 \end{cases}$ β – the mean reversion speed $\frac{\mu_1}{\beta}$ – the long-term mean σ_1 – the volatility of the mean-reverting heteroskedastic process in base regime μ_2 and σ_2 the mean and volatility of the shifted regime process $S_d(t) = A_1 sin(\frac{2\pi}{365}(t - A_2)) + A_3 t + A_4$ – seasonality process
Gyamerah, Ngare & Ikpe	2019	$dT_i(t) = dS_i(t) + \beta_i(t)(T_i(t) - S_i(t))dt + \sigma(t)T_i(t)dB_i(t)$ $S(t) = a + bt + c\sin(2t\pi/365 + \nu)$ $T_i(t)$ – represents the daily average temperature for location i $S_i(t)$ – the deterministic seasonal component for location i $\beta_i(t)$ – the time-varying speed of mean-reversion for location i $\sigma(t)T_i(t)$ – the daily average temperature volatility through time for location i

Table 2.4: Evolution of pricing models for rainfall.

Authors	Year	Model
Wilks	1998	$Y_{t,k} = r_{t,k} * X_{t,k}$ $p^{01}_{t,k} = P(X_{t,k} = 1 \mid X_{t-1,k} = 0)$ $p^{11}_{t,k} = P(X_{t,k} = 1 \mid X_{t-1,k} = 1)$ $X_{t,k} = \begin{cases} 1, & \text{if } \Phi[\epsilon_{t,k}] < p^{01/11}_{t,k} \\ 0, & \text{otherwise} \end{cases}$ $\mathrm{r}_{t,k} = r_{min} - \delta_{t,k} \ln(\Phi(z_{t,k})), \epsilon_{t,k} \sim N(0,1)$ $\delta_{t,k} = \begin{cases} \beta_{t,k}, & \text{if } u_t <= \alpha_{t,k} \\ \gamma_{t,k}, & \text{if } u_t > \alpha_{t,k} \end{cases}$ $r_{min} = 0.01mm, u_t \sim U[0,1], z_{t,k} \sim N(0,1)$ $\beta_{t,k} >= \gamma_{t,k} > 0, 0 < \alpha_{t,k} < 1$
Ritter, Mubhoff & Odening	2014	$Y_{t,k} = r_{t,k} * X_{t,k}$ $p^{01}_{t,k} = P(X_{t,k} = 1 \mid X_{t-1,k} = 0)$ $p^{11}_{t,k} = P(X_{t,k} = 1 \mid X_{t-1,k} = 1)$ $X_{t,k} = \begin{cases} 1, & \text{if } \Phi[\epsilon_{t,k}] < p^{01/11}_{t,k} \\ 0, & \text{otherwise} \end{cases}$ $r_{t,k} = r_{min} - \delta_{t,k} \ln(\Phi(z_{t,k})), \epsilon_{t,k} \sim N_{K+1}(0, \Sigma)$ $\Sigma : \sigma(k,l) = corr[\epsilon_{t,k}, \epsilon_{t,l}]$ $\delta_{t,k} = \begin{cases} \beta_{t,k}, & \text{if } \frac{\Phi[\epsilon_{t,k}]}{p^{01/11}_{t,k}} <= \alpha_{t,k} \\ \gamma_{t,k}, & \text{if } \frac{\Phi[\epsilon_{t,k}]}{p^{01/11}_{t,k}} > \alpha_{t,k} \end{cases}$ $r_{min} = 0.01mm, z_{t,k} \sim N_{K+1}(0, Z)$ $Z : \eta(k,l) = corr[z_{t,k}, z_{t,l}]$
Szabò & Szèpszò	2016	$X_P(t) = (\frac{\triangle P(t)+100}{100})\overline{P}(\frac{\epsilon(t)+100}{100})$ $\overline{P}$ – precipitation polynomial averages for the reference period $\triangle P(t)$ – the departure of polynomial values from this reference with t time $\epsilon(t)$ – the residuals of raw model results from the polynomial fits

Table 2.5: Evolution of pricing models for maize yield using basket (rainfall and temperature) weather derivatives.

Authors	Year	Model
Dzupire, Ngare & Odongo	2019	$Y = I + \epsilon, y$ – maize yield, $\epsilon \sim N(0, \sigma^2)$ $I = \alpha_0 + \alpha_1 R_{cd} + \alpha_2 CDD + \alpha_3 R_{cd}^2 + \alpha_4 CDD^2 + \alpha_5 R_{cd} CDD$ $R_{cd} = \sum_{i=1}^{n} \min(0, \sum_{j=(i-1)s+1}^{is} r_j - r^{min})$ $CDD = \sum_{i=1}^{n} 1_{T_i > 23}, r^{min}$ – minimum rainfall amount

Table 2.6: Evolution of pricing models for snowfall.

Authors	Year	Model
Dischel et al	1999	$S_n = a\theta_n + bS_{n-1} + \gamma\epsilon_n$ $\epsilon_n \sim N(\mu, \sigma^2)$ – iid $a + b = 1, \gamma = 1$ θ_n – the snowfall of the n-th day in an average year S_{n-1} – is the one day lagged snowfall
Luo et al	2010	$dS(t) = [\alpha\theta(t) + \beta S(t)]dt + \gamma dm_1 + \delta dm_2$ $S(t)$ – is the daily snowfall level for day t $\alpha\theta(t) + \beta S(t)$ – is the expected value of daily snowfall dm_1 and dm_2 – the actual distributions which have no assumption about the shape but are bootstrapped from the actual historical snowfalls

Weather derivatives have a payoff that depends on a weather index which represent the weather conditions against which protection is being sought. The effect of hedging using weather derivatives can also be achieved using an insurance contract that has a payoff based on a weather index (Jewson & Brix (2005)). Weather derivatives are economical in comparison to insurance, require no proof of damage or loss and provide protection from the uncertainty in normal weather (Geyser & Van der Venter (2001)). There are some differences between weather derivatives and index-based weather insurance contracts that may mean that one is preferable to the other in certain circumstances. For instance, some companies may not be happy with the idea of trading derivatives but comfortable with buying insurance. According to Jewson & Brix (2005) and Stefani et al (2018) other ways in which insurance and financial derivatives differ include the following:

- cover different kind of risks: weather insurance protect from low frequency, high impact extreme weather events such as tornadoes, floods and weather related fires while financial derivatives are more suited to protect against higher frequency, lower impact events (minor fluctuations of the underlying variable; for example,

Table 2.7: Evolution of pricing models for wind.

Authors	Year	Model
Šaltytė Benth & Benth	2010	$X_t^k = S_t^k + \sum_{i=1}^{p^k} \phi_i^k X_{t-i}^k + \sum_{j=1}^{q^k} \theta_j^k \epsilon_{t-j}^k + \varepsilon_t^k$ $t \in [0,T]$, $k = d, 3h$, $l_d = 1$, $l_{3h} = 8$ ϕ_i^k and θ_j^k – parameters from $ARMA(p_k, q_k)$ $\varepsilon_t^k = \sigma_{t,k}\delta_t^k$ $\sigma_{t,k} = c_0 + \sum_{d=1}^{3} c_d cos(\frac{2k\pi t}{365})$ δ_t^k – iid process $S_t^k = a_0^k + \sum_{i=0}^{l} a_{2i+1}^k \cos(\frac{(2i+2)\pi t}{365 * l_k}) + \sum_{j=0}^{l} a_{2j+2}^k \sin(\frac{(2j+2)\pi t}{365 * l_k})$
Šaltytė Benth & Šaltytė	2011	$Z(s;t) = \mu(s;t) + \varepsilon(s;t)$ $S(s;t) = a_1^s + \sum_{l=1}^{L}[a_{2l+1}^s cos(\frac{(2l+2)\pi t}{365}) + a_{2l+2}^s sin(\frac{(2l+2)\pi t}{365})]$ $\mu(s;t) = S(s;t) + \sum_{i=1}^{p_s} \phi_i(s)(Z(s;t-i) - S(s;t-i))$ $\varepsilon(s;t) = \sigma(s;t)\epsilon(s;t)$, $\epsilon(s;t) \sim N(0,1)$ $\sigma(s;t) = b_0^s + \sum_{j=1}^{J}[b_{2j}^s cos(\frac{2\pi jt}{365}) + b_{2j+1}^s sin(\frac{(2j+1)\pi t}{365})]$
Alexandridis & Zapranis	2013	$dW_t^{(l)} = S(t) + k(t)(W_{t-1}^{(l)} - S(t-1))dt + I_p\sigma(t)dB_t$ $S(t) = a_0 + b_0 t + \sum_{i=1}^{I_1} a_i sin(2\pi i(t - f_i)/365)$ $+\sum_{j=1}^{J_1} b_j sin(2\pi j(t - g_j)/365)$ $\sigma^2(t) = c_0 + \sum_{i=1}^{I_2} c_i \sin(\frac{2i\pi t}{365}) + \sum_{j=1}^{J_2} d_j \cos(\frac{2j\pi t}{365})$
Benth & Pircalabu	2018	$P(t) = \Lambda(t)\exp(-X(t))$ $dX(t) = \alpha(\mu - X(t))dt + dL(t)$ $\alpha > 0$, $\mu > 0$, $L(t) = \sum_{k=1}^{N_t} J_k$, $N_t \sim Po(\lambda)$, $J_k \sim Exp(1/\lambda)$ $\Lambda(t) = a_1 + a_2 \sin(\frac{2\pi t}{365}) + a_3 \cos(\frac{2\pi t}{365})$

a temperature drop/increase by a few degrees, a higher/lower amount of snow or rainfall or a higher/lower wind speed)

- with weather derivatives, the payout is designed to be in proportion to the magnitude of the phenomena while weather insurance pays a once-off lump sum that may or may not be proportional to the magnitude of the phenomena and as such lacks flexibility

- it may be necessary to perform a frequent (daily, weekly or monthly) revaluation of derivative positions, known as mark to market or mark to model, but this is usually not necessary for insurance

- weather insurance normally pays out if there has been proof of damage or loss while weather derivatives require only that a predetermined index value has behaved in prespecified way

- a tax payment could be required: most commonly, insurance reimbursements might incur a tax deduction while derivative ones do not
- the accounting treatment may be different
- contractual details are different

All of these listed differences vary in a certain degree from country to country. As a consequence weather derivatives will not totally replace insurance contracts since there are a number of significant differences between them.

Traditional insurance can protect revenue losses or damages from weather-related events and can be expensive and requires no proof of damage or loss. But such protection has several drawbacks. First of all, weather caused damages, to some extent, are social risks and many insured subjects may suffer simultaneously. As a result the insurer who issued such policies may face financial problems to reimburse such claims. Therefore, insurers may withdraw from the markets when find out that cumulative losses are too large (Chen and Hamwi (2012)). Secondly, insurance companies might face losses regardless of the behaviour of the financial markets. Thirdly, insurance usually only covers extreme weather and catastrophic losses thus insured are not fully protected (Ender and Zhang (2015)). Weather derivatives insurance has been attractive to developing countries with small size of farms. In China large size farmers prefer buying traditional weather insurance because it can cover their actual loss. Also, large size farmers pay lower insurance premiums per unit coverage than small size farmers, since the insurer's underwriting cost will be lower due to the economies of scale and also because of less adverse selection and moral hazard problems (Boyd et al (2011)). In addition, large size farmers receive huge government subsidies for weather insurance and premiums (Mahul and Stutley (2010)). In this report we recall and emphasize that weather derivatives can be used as a risk management tool for all kind of industries affected by weather. A review on the recent scientific papers follows divide in three categories: weather derivatives, insurance and climate change.

2.3.1 Weather Derivatives

Weather derivatives are financial instruments that can be used in order to reduce risk associated with adverse or unexpected weather conditions. The payoff of a weather derivative depends on underlying weather variables such as temperature, rainfall, wind or snowfall.

2.3.1.1 Temperature-based models and risks

Since the majority of the traded weather derivatives are written on temperature indices then temperature is the most common used variable in weather derivatives. More recent studies propose dynamic models which directly simulate the future behavior of temperature using different pricing approaches. Lee and Oren (2009) propose an equilibrium pricing model for weather derivatives in a multi-commodity setting. The optimal payoff in a multi-commodity setting is given as

$$x^*_{i,1}(P_i) = \mu_{I_i} - E[I_i|P_i] - \alpha_{i,1} W_0 B_1 (1 + \frac{1}{W_0 B_1}(E[W_1|P_i] - \mu_{W_1}))$$

where

$$\alpha_{i,1} = \frac{\frac{1}{v_i}(\mu_{W_1} - B_1 W_0) - (\mu_{I_i}\mu_{W_1} - E[E[I_i|P_i]W_1]) - \sigma_{W_1 I_i}}{\mu^2_{W_1} + \sigma^2_{W_1} - E[E[W_1|P_i]W_1]},$$

$$W_0 = \frac{\sum_{i=1}^2 \frac{\Lambda_i}{\Gamma_i} + \sum_{j=1}^2 \frac{\Lambda_j}{v_j \sigma^2_{W_1}} \frac{\mu_{W_1}}{v_m \sigma^2_{W_1}}}{B_1(\frac{1}{v_m \sigma_{W_1}} + \sum_{i=1}^2 \frac{1}{v_i \Gamma_i} + \sum_{j=1}^2 \frac{1}{v_m \sigma_{W_1}})}$$

where $\triangle_i = \frac{\mu_{W_1}}{v_i} - (\mu_{I_i}\mu_{W_1} - E[E[I_i|P_i]W_1] - \sigma_{W_1 I_i})$, $\Gamma_i = \mu^2_{W_1} + \sigma^2_{W_1} - E[E[W_1|P_i]W_1]$, $\Lambda_j = \mu_{W_1} - v_j \sigma_{I_j W_1}$, $B_1 = (1 + r)B_0$ is the riskless bond price at time 1, r is the interest rate and $B_0 = 1$, W_1 is the weather derivative price, and P_i is the unit spot price of type i at terminal time.

In the multi-commodity economy the weather derivative has two effects; the risk hedging effect and the risk sharing effect while in a single-commodity economy there is only a risk hedging effect since there is no counter-party to share risk. The risk hedging effect can be measured by the certainty equivalent difference of the maximized utility with and without weather derivative. The risk sharing effect reflects possible diversification of weather risk across industries with different weather dependence (e.g. some industries may benefit from high temperature while others may be adversely affected). Such risk sharing effect can be measured by the certainty equivalent difference of the maximized utility between the multi-commodity and a single-commodity economy. Authors were able to derive closed form expression for equilibrium prices and the measurement of the risk hedging and sharing effects. An important result was that weather derivatives improves hedging and risk diversification capability, especially in situations where commodity derivatives are not available. Numerical analysis is performed by employing Monte-Carlo simulations.

Benth and Šaltytė Benth (2011) propose a continuous-time autoregressive model for the temperature dynamics with volatility being the product of a seasonal function and a stochastic process. Authors use the Barndorff-Nielsen and Shephard model for the stochastic volatility. The authors find that the proposed temperature dynamics is flexible enough to model temperature data accurately. Also, future prices for cooling and heating degree days and cumulative average temperatures are calculated, as well as option prices written on them. This model is applied on data from Stockholm, Sweden. Next, some issues related to the mean reversion of the model are discussed, where in particular the so-called half-life of the temperature model is derived. To understand how fast the temperature dynamics is reverting back to its long-term average, the notion of half-life (which is defined to be the expected time it takes before the process is returned half way back to its mean from any position) of the stochastic process is discussed.

Meissner and Burke (2011) suggest that the Black-Scholes model (BSM) despite its limitations is an adequate pricing model, which is superior to Burn analysis [4] which is used in pricing weather futures and options contracts. These authors show that the BSM model can be applied to pricing temperature based options. But the BSM model requires completeness in the market which means that it can be used to perform delta hedging. The authors find that there is an active market for temperature futures, which can be used

[4] for more see section 3.3.4

to delta hedge. The BSM model assumes a log normal distribution of the underlying at option maturity and by extensive empirical testing authors show that the log normal distribution provides a reasonably well fit for HDD and CDD data. Therefore, this support the application of the BSM model to price and risk management temperature options in practice.

Benth and Šaltytė Benth (2012) present a stochastic model for daily average temperature which contains seasonality, a low-order autoregresive component and a variance describing the heteroskedastic residuals. Also this paper addresses the issue of using historical burn analysis on derivatives with aggregated values as the underlying variable. A discussion of continuous time models, discrete time model (Campbell and Diebold (2005)) and weather derivatives pricing is done. In Benth and Benth (2005) the dynamics of the deseasonalized temperature is assumed to follow an Ornstein-Uhlenbeck process. Alaton et al. (2002) consider the same dynamics except that the volatility is assumed to be constant for each month. The dynamics are generalized in so-called continuous time AR (CAR) process in Benth et al (2007) which is applied to Stockholm data. Later Härdle and Cabrera (2009) studied this class of processes for German temperature data and Benth et al (2011) for Asian temperatures, all validating the relevance of this class of models. When estimating these models discretization of the stochastic process is applied leading back to the time series models. Temperatures are naturally evolving continuously over time, so it is very appealing to use continuous time stochastic processes to model the dynamics although the data may be on a daily scale and the weather derivatives market settles contracts based on indices of daily average temperature. There is another fundamental aspect related to the nature of the temperature futures markets. Temperature futures can be traded continuously in the opening hours of the exchange. Thus a model for the forward price dynamics is naturally formulated as a continuous time stochastic process. In Benth et al (2008) the dynamics of temperature future prices is to use the risk adjusted to predict index value developing a continuous time model of Brownian motion type by using a Girsanov transform which effectively shifts the seasonal function by some constant called the MPR.

Taib and Benth (2012) study pricing of weather insurance contracts based on temperature indices. Authors take data from Malaysia and analyze three pricing methods such as burn approach, index modeling and temperature modeling. The first two give similar results and a high degree of uncertainty in their premium estimates while premium estimated from the temperature modeling approach is prone of Monte Carlo error. But the temperature modeling approach has the advantage that it can account for current information of the weather situation. Pricing of the insurance contract using a seasonal autoregressive time series model for daily temperature variations (which gives a very precise information on the dynamics of temperature and very detailed probabilistic information on the index and therefore assess correctly the premium and probabilities of loss and profits) is proposed and the profit/loss distribution from the contract in the perspective of the insured and the insurer is investigated. Moreover, chances of receiving a profit form a weather index insurance contract are analyzed. The authors find that chances are not very large using burn analysis approach but far better in the temperature modeling case.

Alexandridis and Zapranis (2012) shows that almost 70% of US companies are affected by weather in some way. Burn analysis is considered the benchmark approach for pricing temperature derivatives. It is impractical to apply no arbitrage pricing models to weather derivatives as it is not possible to construct a risk free portfolio consisting of weather index and the derivative. In addition the no arbitrage condition does not result in a unique price as many martingale measures exist.

Wang, Li, Li, Huang and Liu (2015) present a feasible model for the daily average temperature on the area of Zhengzhou, China and apply it to weather derivatives pricing. Authors apply the mean-reverting Ornstein-Uhlenbeck process to describe the evolution of the temperature and use Monte Carlo simulations to price heating degree day (HDD) call option for this city. Many other mathematical tools, such as wavelet functions, B-spline functions, and polynomial function, can be used for a more accurate modeling of temperature data. In addition, these authors suggest that it would be interesting to treat carefully the jumps of temperatures due to extreme phenomena by integrating a term which describes their behavior within the stochastic model. At the end it is emphasized that is important to construct accurate weather models for other weather factors, such as rain, snow, or fog, since the risks implied by these factors also deeply affect economic development of China.

Yuan, Göncü and Ökten (2015) display the analysis of the estimation of the sensitivities of weather derivatives in a stochastic model of temperatures. These authors use path-wise derivative and kernel methods to derive Monte Carlo estimators for the sensitivity (Greeks) of temperature-based weather derivatives. These sensitivities can be used by investors for choosing the most suitable weather contracts for partial hedging or speculation. Temperature data from New York, Atlanta and Chicago are used in the discussion of numerical results. The sensitivities with respect to the long-term mean temperature parameters are almost identical for all three cities. Therefore, from this aspect there is no distinction between these contracts. However, there is more of a distinction between contracts with respect to their delta, gamma, vega and mean reversion parameter. The delta of the New York HDD option is smallest in absolute terms, suggesting a smaller sensitivity with respect to possible measurement or forecast errors in the initial temperature. In case of partial hedging with a correlated asset, a lower gamma means less re-balancing of the partial-hedging portfolio and possibly lower transaction costs. Results show that gamma of all three options is low with New York HDD call/put options having the lowest value. As a future work, authors suggest, the sensitivities derived in this study can be extended to various dynamic temperature models with more general stochastic processes.

Erhardt (2016) explores a method to model the financial risks of holding portfolios of long-term temperature derivatives for any subset of the 30 North American cities whose derivatives are actively traded on the CME. The author incorporated spatial dependence among 30 cities using a multivariate normal distribution. A normal distribution for payments is obtained through affine transformations of cCDD [5] and cHDD [6]. Moreover, it is demonstrated how the normal distribution could be used to estimate some common risk measures. The value at risk (VaR) is defined as:

$$VaR(L, \alpha) = Q_\alpha = min(Q : P(L \leq Q_\alpha) \geq \alpha)$$

[5] cumulative CDD

[6] cumulative HDD

The VaR is easily obtained since the loss L is a continuous random variable for larger values of α and therefore the distribution function $F_L(l) = P(L \leq l)$ is stricktly increasing and has inverse distribution function F_L^{-1} then $Q_\alpha = F_L^{-1}$. While the conditional tail expectation is defined as:

$$CTE_\alpha(L) = E(L|L > Q_\alpha),$$

which is the expected loss conditional upon exceeding the VaR. In addition, the most notable limitation of the method is its reliance on sufficiently long time periods to allow for normality in the degree day sums. This method is not recommended for modeling cumulative degree days over short time periods to allow the normal approximation to hold.

Kabaivanov and Markovska (2017) focus on modeling environment changes in a way that allows to price weather derivatives in a flexible and efficient way. These authors show that option based approach toward resource management can offer very special insights on rare events and allow to reuse derivative pricing methods to improve natural resources management. To show this the authors use stochastic modeling with mean reverting processes and Monte Carlo simulated temperatures to evaluate Asian weather options. Two different models one with standard Orstein Uhlenbeck process and a second one including jump diffusion and forecast different temperature paths over a 90 days horizon are calibrated. Both models are able to cope with temperature forecasting and yield meaningful results for Asian weather options, but the one which includes jumps and accounts for seasonal effects is more accurate. Authors conclude that the advantages of using environment modeling go beyond the pure valuation of derivative instruments because it provides a common framework that can fit together stochastic models, management decisions, financial impact and the effects on individual behavior. That allows to assess not only derivative contracts but also to forecast and measure the result of regulations and environment policies.

Prabakaran and Singh (2017) construct the temperature model under Ornstein Uhlenbeck process which is driven by a Levy process rather than a standard Brownian motion. These authors extend their approach to model and price weather derivatives. Also it is discussed on how weather forecasting and seasonal forecasting could potentially improve their valuation of weather derivative contracts. The simplest case of using weather forecasts in weather derivative pricing is considered, which is the calculation of the fair price of a linear swap contract on a separable [7] and linear index such as CAT. After that the calculation of the fair price of a linear swap contract on a separable non-linear index such as HDD is considered and the general case which includes the calculation of the fair price for all other contracts (non-linear swaps and options) and the calculation of the distribution of outcomes for all contracts which is the most difficult case.

Alexandridis, Gzyl, Ter Horst, and Molina (2017) propose the use of the maximum entropy method to extract the risk neutral probabilities directly from the weather market

[7] For a separable index, the variance is the sum of the terms in the covariance matrix of daily index values during the contract period. In case of CAT separable index is the sum of the terms in the covariance matrix of the daily temperatures.

prices. The maximum entropy method for a call/put European option is defined as:

$$\pi_m = \sum_{j=1}^{K} \rho_j O_m(\widehat{S}_j, K_m) p_j, m = 1, ..., M$$

where ρ_j is the density of the risk neutral probability q with respect to the probability p and $\sum_{i=1}^{K} \rho_i p_i = 1$, K_m are the strike prices, M is the number of option prices used, $O(\widehat{S}_j, K_m)$ is the payoff of the mth option (which will be either a call or a put of European type), and π_m is its observed price. This method is computationally fast, model free, non-parametric and can overcome the data sparsity problem that governs the weather market. These authors infer consistent risk neutral probabilities along with their densities from the market price of temperature options, and price accurately options even in the cases where the realised underlying HDDs indices were significantly different from the historical average. The empirical results indicate that CAR and historical burn analysis probabilities can provide better reconstruction of the original option prices and produce a lower mean absolute distance in the valuation of the options not used in the fitting procedure. It is concluded that when the available information in the market arrives from historical data or from meteorological forecasts pricing is more coherent.

Evarest, Berntsson, Sigull and Yang (2018) discuss pricing of weather derivatives whose underlying is temperature and where the temperature follows a two state regime switching model with a heteroskedastic mean reverting process as base regime and a shifted one defined by Brownian motion with mean different from 0. The model allows the volatility of the underlying process in the base regime to vary with changes in temperature process. This model is applied for pricing futures contract on HDD, CDD and CAT and corresponding call option contracts on these futures. The authors suggest that for realistic contract payoff, it is important to estimate MPR based on the available market prices and after that make comparison between market prices and expected payoff from the model. The Monte Carlo simulation approach for the underlying temperature dynamics model is described and then this approach is used to price the call option contracts.

Gyamerah, Ngare and Ikpe (2019) propose to use a machine learning ensemble technique to determine the relationship between maize yield and weather variables. This approach aim to eliminate the product design basis risk which makes most smallholder farmers and agricultural stakeholders unwilling to pay for the price of weather derivatives. These authors develop a mean-reverting model with a time-varying speed of mean reversion, seasonal mean, and local volatility that depended on the local average temperature. The model is extended to a multi-dimensional model for different but correlated locations. The application of these models is done on futures, options on futures and basket futures for cumulative average temperature and growing degree days (GDD). Since there is not a real weather derivative market in Africa from which the prices of indices under scrutiny can be obtained, it is assumed a constant MPR in the pricing models. With these efficient and reliable pricing models, basis risk will be mitigated. As a result there will be an increase in the willingness to pay for the contracts on the farmers side and trading activities in the weather derivatives market will also increase and it will be cost efficient to buy contracts for different but correlated farming locations rather than a single farming location. The agricultural sector of Ghana, which is vulnerable to climate shocks, is considered as an illustration.

2.3.2 Rainfall-based models and risks

Even the most traded weather derivatives are based on temperature indices, several economic sectors are exposed to rainfall risk as well. Farmers and financial investors are affected by indirect losses caused by scarce or abundant rainfall. With rainfall derivatives firms can transfer rainfall risk to the capital market and give the buyer the opportunity to reduce rainfall risk exposure in order to profit from weather uncertainty. The process of pricing rainfall derivatives it is not as smooth as in temperature case and the literature on pricing rainfall derivatives is thin. Ritter, Mubhoff and Odening (2014) estimate a daily multi-site rainfall model from which optimal portfolio weights derive. This method, which is calibrated to the historical data and then simulates future rainfall, allows to reduce geographical basis risk more efficiently than simpler approaches such as inverse distance weighting. The reduction of geographical basis risk is done by combining weather derivatives with different reference stations in Germany. Including a new weather station requires partially a re-estimation of the multi-site rainfall model.

Härdle and Osipenko (2017) emphasized that weather derivatives are contingent claims with payoffs determined by future events as temperature, snowfall and rainfall. Authors develop a utility-based model for pricing baskets of weather derivatives under default risk on the issuer side in OTC markets. The terminal wealth of buyer j at time T is expressed as:

$$\Pi_{j,T} = I_j + \alpha_{j,T}^T W_T + \beta_{j,T} B_T = I_j + V_{j,T}$$

with I_j being a random income that depends on some weather indices entering the final payoff W_T. $\beta_{j,T} B_T$ and $\alpha^{j,T} W_T$ are the payoffs of the risk free asset and of the basket of the weather derivatives; together they constitute $V_{j,T}$, the terminal portfolio value of buyer j. If $\alpha_{j,t+1:T}$ denote trading strategies of agent j from $t+1$ to T then the portfolio choice problem of buyer j in each $t = 0, 1, \ldots, T1$ is expressed as:

$$\max_{\alpha_{j,t+1:T} \in \mathbb{R}^{Sx(T-t)}} E_t(U_j(\Pi_{j,T}))$$

$$\alpha_{j,t+1}^T W_t + \beta_{j,t+1} B_t - V_{j,t} = 0$$

While the terminal wealth of investor m at T is defined as:

$$\Pi_{m,T} = -\alpha_{m,T}^T W_T + \beta_{m,T} B_T = V_{m,T}$$

with $\alpha^{m,T} W_T$ and $\beta_{m,T} B_T$ being the payoffs of the weather derivatives portfolio and the risk free asset respectively. Investor's portfolio choice problem in each $t = 0, 1, \ldots, T1$ is expressed as:

$$\max_{\alpha_{m,t+1:T} \in \mathbb{R}^{Sx(T-t)}} E_t(U_m(\Pi_{m,T}))$$

$$\alpha_{m,t+1}^T W_t - \beta_{m,t+1} B_t + V_{m,t} = 0$$

In time $t < T$ investor m maximises expected utility of the terminal wealth with respect to all future trading strategies, subject to a self-financing portfolio. In this model agents maximise the expected utility of their terminal wealth, while dinamically rebalance their weather portfolios over a finite investment horizon. Using dynamic programming approach to portfolio optimisation over a finite investment horizon semi-closed forms for the equilibrium prices of weather derivatives and for the optimal trading strategies are obtained. The authors find an adverse effect of increasing counterparty default risk and capital costs on the demand for weather derivatives and on their prices. At the end, the proposed model is applied to price rainfall options using historical data of agricultural provinces Changde and Enshi in China.

Undli and Schatvet (2018) examine the effect on a Norwegian hydropower producer's operating income by hedging volumetric risk with the use of weather derivatives. Precipitation evolves more irregularly and unevenly than temperature changes, furthermore it does not have the same geographical correlation structure found for temperature. Also due to many zero values in precipitation data a logarithmic transformation is not the right approach and therefore it is conducted a quadratic and linear regression. After that the indifference pricing method [8] and McIntyre pricing method is a simple analytical model presented by McIntyre in 1999 for pricing weather derivatives, which assumes that data follows a normal distribution are applied, resulting that McIntyre performs better. In this case derivatives allow hydropower producers to transfer weather risk to a third party. It is shown that for periods characterized by low precipitation and high standard deviation the production of hydropower can effectively be hedged by using monthly options on precipitations resulting in increased operating income and reduced volatility. The pricing model developed can be used in agriculture as well where a farmer is interested in hedging weather risks due to rainfall and temperature simultaneously and economically. It can also be used to price weather derivatives in other weather related industries affected by rainfall, temperature or both. To further develop application of weather derivative in hydropower case one have to alternate indifference and McIntyre pricing models applied to a larger dataset and different geographical locations.

2.3.3 Basket of temperature and rainfall

Other important results come from scientific work on rainfall concluding with a paper that mixes rainfall and temperature effect together. Dzupire, Ngare and Odongo (2019) present an incomplete market pricing approach to analyze the evaluation of weather derivatives and the viability of a weather derivatives market in terms of hedging. For the specification of indifference prices for the seller and buyer of a basket of weather derivatives written on rainfall and temperature is developed a utility indifference method. The agent's risk preference is described by an exponential utility function and the prices are derived by dynamic programming principles and corresponding Hamilton Jacobi-Bellman equations from the stochastic optimal control problems. At the end it is shown how the basket weather derivatives, whose underlying indices are rainfall and temperature processes, contribute to maize yield variability.

[8] is a utility based approach which has been presented by both Brockett et al. (2006) and Xu et al (2008)

2.3.4 Wind-related models and risks

Another group of research work in weather derivatives market is focused on wind. In this report, the various aspects of weather derivative have been presented. So far, the focus has been on modeling and pricing temperature and rainfall derivatives. In this section, the focus is on wind derivatives. Šaltytė Benth and Benth (2010) propose an ARMA time-series model for the wind speed at a single spatial location, and estimate the model with data based on three different wind farm regions in New York. Large discrepancies in the behaviour of daily average and three-hourly wind speed (DAWS and WS) records are demonstrated. In order to compare the power prediction of DAWS and WS the mean square prediction error (MSPE) is calculated. The authors find that MSPEs for DAWS data are much smaller than those for the three-hourly WS, indicating that more accurate predictions are obtained from modeling aggregated data directly rather than at the finer time scale. Moreover, some Box-Cox and back transformations are applied in the data.

Šaltytė Benth and Šaltytė (2011) propose a spatial–temporal model for the wind speed (WS) applied on daily WS records from 18 meteorological stations in Lithuania. The model contains seasonality, a higher-order autoregressive component, a variance describing the remaining heteroskedesticity in residuals and is estimated at the single spatial meteorological station independently on spatial correlations. The spatial dependencies are modeled by a Gaussian random field. Referring to a validation procedure based on out of sample observations these authors conclude that the model predicts well the WS and can be observed in many other practical applications.

Alexandridis and Zapranis (2013) model the dynamics of the wind generating process using a non-parametric non-linear wavelet network which is validated in New York. The proposed methodology is compared against alternative methods, proposed in prior studies. Their results indicate that wavelet networks can model the wind process very well and consequently constitute an accurate and efficient tool for wind derivatives pricing. These authors provide pricing equations for wind futures written on two indices, the cumulative average wind speed index and the Nordix wind speed index. The characteristics of the wind speed process are very similar to the process of daily average temperatures. It is indicated a slight downward trend and seasonality in the mean and variance. In addition the seasonal variance is higher in the winter while it reaches its lower values during the summer period.

Benth and Pircalabu (2018) propose a non-Gaussian Ornstein–Uhlenbeck model for the wind power production index. The model allows for an analytical formula for pricing wind power futures. These authors provide an empirical study, where the model is calibrated to 37 years of German wind power production index that is synthetically generated assuming a constant level of installed capacity. Also, the MPR based on one year of observed prices for wind power futures with different delivery periods, is studied. Generally, the authors find a negative risk premium whose magnitude decreases as the length of the delivery period increases. The result suggests that wind power producers are willing to accept a lower price when selling wind power futures. Moreover, the MPR is more volatile for shorter delivery periods and it is argued that this behaviour might be related to liquidity aspects and the information contained in short-term weather forecasts, which the proposed model does not incorporate.

2.3.5 Snowfall-based models and risks

The last group of weather derivatives we consider in this report is based on snowfall. Although rainfall and snowfall are two different weather variables and contracts are traded separately, according to Alexandridis and Zapranis (2013), they share a lot of common characteristics. Luo et al (2010) uses statistical modeling methods for index and daily snowfall modeling and compare them. The index model which performs better is the generalized Edgeworth expansion with adjusted exponential distribution since it gives the lowest chi-squared statistic. For the daily snowfall modeling, the GARCH model is the one who explains and performs better the daily snowfall pattern.

Tang and Jang (2011) perform an empirical analysis that examines geographical diversification and financial hedging as two strategies against snowfall risk. Risk management is split into two terms, namely operational hedging and financial hedging. Operational hedging aims to reduce the exposure to risk factors by changing the operations, whereas financial hedging is designed to transfer risks to the insurance company or the market by purchasing insurance or trading financial instruments (such as forwards, futures, or options). Also, financial hedging enables the firm to invest in more attractive investment opportunities and the companies do not have to change their operations only for risk management purposes. The empirical analysis results suggest that financial hedging might be a more effective strategy for ski resorts. A contract that covers only a quarter or even a month instead of the entire season is more effective. To achieve better weather risk management, outcomes are provided based on simulating (using Monte Carlo simulations) the interactions between geographical diversification and financial hedging. As a result, business diversification could improve the outcomes of financial hedging.

Tang and Jang (2012) construct a snowfall forward to hedge the snowfall risk for Winter Sports, a public traded, single-property resort. Demonstrating hedging effectiveness is optimal in this case, as the firm is not diversified geographically nor business-wise. Establishing a hedge is deciding the strike, which is a predetermined level of snowfall index used to decide profit or loss at the expiration of the contract and is chosen to be the historical mean. To achieve a “costless” hedge, whose expected value is zero for both sides on entering into the contract, the strike has to be set at a level that is neutral so that neither side has a built-in profit on entering the position. The objective of hedging is to minimize the volatility regardless of the direction of price/quantity change, not to maximize the level of cash flows. To find optimal hedge the authors regress unhedged, operating cash flow against quarterly and annual snowfall to find an effective hedge. Monte-Carlo simulation is used as well, generating estimates of cash flow alternatives in order to test the statistical significance of the regression. Furthermore, these authors present the results of this hedging strategy in the period 1991-2003 by displaying cashflow with and without a derivative present. Cash flow volatility was reduced by 25.8% at most, when actual snowfall was high. At the end it is concluded that it is more efficient to use snowfall forwards as a risk management tool in months with high snowfall levels.

Some recent work as master thesis is done by including an analysis on how temperature affects on beer sales during the year (Hoornaert (2018)). This is done by regression analysis, while a second one (Haugan and Rasmussen (2019)) is about snowfall in which four ways of pricing models on snow are considered: historical densities, Edgeworth densities, burn analysis and ADS model that is a model proposed by Alaton, Djehiche and Stillberger (2002). All these models are compared and at the end it is concluded that Edgeworth densities model performs better when one consider snowfall as weather index.

2.3.5.1 Edgeworth historical density

The option prices for CDD and HDD have also been calculated using Edgeworth adjusted historical densities. There might be some situations requiring the changes in the prices due to the distributional characteristics of the data, particularly temperature data. Due to non-normality the pricing needs to be modified by taking into consideration of moments of distribution higher than second order, which is known as the so-called generalized Edgeworth series expansion and has been applied to option pricing by Rubinstein (2000). In the model $a(x)$ is the density of normal distribution function which is extracted from historical distribution of temperature data by using first two moments. Then, by using skewness η and kurtosis κ measures of the historical data the densities can be modified and adjusted according to the following formula (Stuart and Ord (1987)):

$$f(x) = [1+\frac{1}{6}\eta*(x^3-3x)+\frac{1}{24}(\kappa-3)(x^4-6x^2+3)+\frac{1}{72}\eta^*2(x^6-15x^4+45x^2-15)]a(x)$$

where x is standard normal variable and $f(x)$ denotes the Edgeworth density of $a(x)$. Accordingly, the adjusted Edgeworth densities can be calculated as weights of the put option payoffs during the selected period for the specific strike levels. The skewness and kurtosis adjusted call and put prices can be calculated according to the following formula:

$$Call(t) = e^{-r(T-t)}\frac{1}{\sum_{i=1}^{N} f_i(x)}\sum_{j=1}^{N} f_j(x)max(\sum_{k=1}^{D} X(k) - K, 0)_j$$

$$Cut(t) = e^{-r(T-t)}\frac{1}{\sum_{i=1}^{N} f_i(x)}\sum_{j=1}^{N} f_j(x)max(K - \sum_{k=1}^{D} X(k), 0)_j$$

In the above formula K is the strike price, N is the number of observations, D is the number of days in a particular period and $X(k) = max(T(k) - 18, 0)$ for CDD and $X(k) = max(18 - T(k), 0)$ for HDD.

2.3.6 Insurance-related market

In the past, insurance contracts and catastrophe bonds were widely used by companies in weather-sensitive industry sectors. Insurance contracts are used to protect the buyer of the contract against adverse weather conditions, and written on rare weather events such as extreme cold or heat and hurricanes or floods. These events are highly liked to create great catastrophes with huge impact on the revenues of the company.

Bobojonov, Aw-Hassan and Sommer (2014) examine the potential of three index insurance schemes (1) a statistical index, (2) an index based on agro-meteorological approach and (3) a remote sensing-based index for minimizing risk. It also discusses how index-based insurance markets contribute to rural development in scenarios of increasing climate risks in Syria. The study identifies that all three insurance schemes have a very high potential to cope with increasing climate risk. Insurance schemes designed according to these indexes performed very well in terms of covering revenue losses in most of the extreme drought years observed in the country. Farmers purchasing an insurance contract may have better access to credit and find it easier to invest in agricultural production and improve productivity. Because such alternative index-based insurance programmes are low cost, such methods are more affordable for poor farmers and thus can potentially make an excellent contribution to economic growth in rural areas.

Porrini and Schwarze (2014) analyse the performance of different insurance models in relation to information imperfections (such as adverse selection and moral hazard) and market imperfections (such as charity hazard [9] and transaction costs). In addition, the different models are examined in terms of the extent to which stimulate mechanisms that facilitate the mitigation of greenhouse gas emissions, adaptation to the inevitable impacts of climate change and the development of climate risk finance management. Some concluding remarks are offered regarding the possible future development of a European insurance model as a means of developing an economically effective response to natural hazards caused by climate change.

Clarke et al (2016) argue that weather-indexed insurance products currently being sold to farmers are derivatives, not indemnity insurance products. The model the author proposes is one of rational demand, where the consumer is assumed to be a price taking, risk averse, expected utility maximizer with, for some results, decreasing absolute risk aversion (DARA). One critical aspect of the model is the nature of the joint probability structure of the index insurance product and the consumer's loss. The net transfer from index insurance is assumed to be imperfectly correlated with the consumer's net loss, and so index insurance purchase both worsens the worst possible outcome and improves the best possible outcome; a consumer might incur a loss but receive no net income from the index insurance product, or incur no loss but receive a positive net income. Some upper bounds derive, for rational purchase of hedging instruments. For the case of indemnity insurance, that is insurance without basis risk[10], risk aversion and DARA alone cannot bound the purchase of indemnity insurance below full insurance; an infinitely risk-averse individual would rationally purchase full insurance. However, tighter bounds may be derived for actuarially fair or unfair hedging products with basis risk, both under the restriction of risk aversion alone, and that of risk aversion and DARA. Loosely speaking the bound for DARA arises because an individual who cares enough about the risk of wanting to purchase a sizable hedge must, under the assumption of DARA, care enough about the downside basis risk and, for the case of an actuarially unfair price, the dead weight cost of hedging to limit the size of the hedge. Authors present a ratio that may be useful for understanding the level of basis risk in a product from a consumer's perspective, and then apply the ratio to the numerical example of

[9] which refers to the reduced incentive to insure oneself against disaster damage in anticipation of governmental and/ or private assistance

[10] an increase in insurance purchase transfers wealth from high- to low-wealth states, subject to some dead weight cost

weather index insurance across one state in India. With a belief constructed from the empirical joint distribution function of yields and weather indexed claim payments, it is shown that optimal demand from any risk averse expected utility maximizer is zero if the price for index insurance is more than 1.56 times the expected claim income.

Doms (2017) analyzed how weather index insurance based put and call options as well as strangles reduce performance risk of 20 German crop farms (which are not exposed by extreme climatic conditions) and conclude with a comparison between them based on hedging efficiency which is defined as the percentage change of the volatility of farm specific total gross margins with and without weather index insurance. A main point of interest in the weather index insurance research is the minimization of the basis risk. One could find that customized contracts are better suited to reduce performance risk than standardized contracts and hedging efficiency varies considerably from farm to farm and depends highly on the contract type, the analyzed customized call-options and strangles clearly outperform the customized put-options. Moreover the results indicate that weather index insurances are highly farm specific. For reasons of comparability, specific indexes are selected from numerous possible indexes based on assumptions regarding possible sources of yield related performance risk. These indices were applied to each of the farms under scrutiny equally and as a consequence there is need for a farm specific risk analysis before a farmer decides for a specific insurance type.

Additional work should investigate whether the designed insurances reduce the performance risk of farms located in regions with extreme farming conditions. It is also of interest whether a change of the parameters of the standardized contracts might improve their risk reducing capacity. One also might analyze contracts based on other indexes such as mixed indexes.

Surminski and Hudson (2017) determine how the "risk reduction linkages of insurance" [11] can be assessed and developed further in a multisector partnership (MSP) setting. Traditionally, efforts to evaluate disaster insurance are focused on affordability, availability, commercial viability and financial sustainability. These authors add another feature of "risk reduction" as an indicator of the impacts that insurance can have on the underlying risk levels. Four methodologies (such as stress testing, estimation of effectiveness of policyholder-level mitigation measures, analysis through a risk reduction framework and investigation of the design principles of insurance) are considered and it is explored for several European examples of insurance development. While very different in scope and history, all examples share one common feature: they can all be considered as MSPs designed to foster greater collaboration between different stakeholders. Those examples are used to act as testing grounds for the assessment of the risk reduction implications of insurance. Author's findings show that the potential for risk reduction of new or reformed schemes could be strengthened through multisectoral partnerships.

Keucheyan et al (2018) examine the ongoing financialization of climate risk insurance. This paper describes the structure of modern natural risk insurance and reinsurance, a structure that is currently undergoing profound changes due to the combined impact of two processes, financialization and the emergence of new risks,

[11] the lack of linkages between insurance underwriting and risk reduction has been identified as a key barrier for insurability and affordability

including climate change. One response is discussed to the challenges of climate change experienced by the insurance industry since the 1990s: so-called "catastrophe bonds", a financial mechanism meant to insure against possible natural disasters. At the end the paper shows that financialization and new risks alter the role of insurer "of last resort" that the state [12] has played since the 19-*th* century. In the context of the "fiscal crisis of the state", part of the state's insuring capacity has been transformed into financial mechanisms such as sovereign catastrophe bonds and microinsurance. The last two sections reflect on the growing role of big data and financial modeling in climate risk insurance and, presenting a critique of the process of financialization on both economic and political grounds.

Stefani et al (2018) rely on the studies of Benth and Benth (2005), Benth and Benth (2011) and Benth and Benth (2012) to create a model for temperature forecasting. In this paper weather derivatives are considered as a hybrid instrument which encompasses properties of insurance contracts. To estimate the prices of these derivatives, the conventional VaR technique is exploited. This technique is applied to define the level of the critical temperature (called threshold) but also in the risk loadings added for hedging purposes.

Fusco, Miglietta and Porrini (2018) focus on the relation between insurance variables and agroclimatic variables, such as the different levels of precipitation and temperature, focusing in particular on the drought phenomenon. With a decrease in precipitation and an increase in temperature the need to cover risks with adequate insurance instruments increases. These authors collect agroclimatic and insurance data for each Italian province for the period 2004-2011, and measure the influence of climatic agroclimatic variables on insurance variables, such as Total Premiums, Insured Value and Certificates. The results of the analysis show the significance of the precipitation variable and its negative effect with each insurance dependent variable. The same result can be observed focusing on the effect of minimum temperature on two insurance variables, i.e. Total Premiums and Certificates. Models tested explain a range between 44% and 51% of the variation in their insurance dependent variables. The analysis confirms that climate factors represent an incentive for the adoption of insurance instruments highlighting the necessity to increase farmers' information and to support insurance instruments through public subsidies. The paper concludes that adverse climatic events should not be considered as exceptional events, but as one of the negative externalities with which the agricultural enterprise must live. This major concern, especially in the agricultural sector particularly vulnerable to adverse climatic events, shows the importance of providing suitable financial hedging instruments for farmers.

2.3.7 Related works in climate change

Until now we have reviewed numerous models proposed by different researchers and applied in different areas. But weather models consist on uncertainties, which stem from the followings:

- Natural variability is the inherent part of the climate system causing its continuous

[12] plays the role of legal regulator of the insurance market, including of risk pooling practices by insurers

change without any external forcing. For instance, two consecutive years can be extremely dry or wet over a region due to this instance.

- Different climate models use various numerical approximations and parametrization schemes to describe physical processes resulting in dissimilar results, as well. The largest diversity occurs in the description of cloud and precipitation processes.

- Climate change is highly influenced by the anthropogenic activity. Its global future path is not known yet and therefore different (optimistic, pessimistic) scenarios are constructed in order to achieve a proper estimation. These hypothetical scenarios are taken into account within models through various greenhouse gas concentration pathways (Szabȯ and Szėpszȯ, 2016).

According to this, one cannot make any reasonable statements based on results of a single weather model run but only through quantifying the projection uncertainties. This could be achieved by applying the ensemble method, when more models and scenarios are considered together and future changes are expressed as probabilistic information.

For example at the Hungarian Meteorological Service a detailed assessment is made on the uncertainties of temperature and precipitation projections (Szabȯ and Szėpszȯ, 2016), based on the modified method of Hawkins and Sutton (2009 and 2011). The study is based on how the total uncertainty could be reduced via model improvements and scenario developments. These investigations may enable to increase the reliability of the climate change information with finding a representative model. To conclude this section a brief list of research works in assessing climate change are presented. Hwang, Tol and Hofkes (2016) investigate the role of emissions control in welfare maximization under fat-tailed risk about climate change. A classification of fat-tails is done and it is discussed about the effects of fat-tailed risk on climate policy. One of the most important findings is that emission control may prevent the "strong" tail-effect from arising, at least under some conditions such as bounded temperature increases, low risk aversion, low damage costs and bounded utility function. Moreover the fat-tailed risk with respect to a climate parameter does not necessarily lead to an unbounded carbon tax. To better understand what a fat-tail is: a probability density function has a fat tail when its moment generating function is infinite, that is, the tail probability approaches 0 more slowly than exponentially. The method used in this paper is Gauss-Hermite quadrature [13] which is used for numerical integration while for the simulations the authors use a one-box temperature response model. Some implications of applying this kind of technique are: climate policy greatly reduces the effect of fat-tailed uncertainty on damage costs and consumption and that the effect is sensitive to the unit cost of emissions control. So, optimal carbon tax does not necessarily accelerate as implied by Weitzman's Dismal Theorem (dismal theorem argues that "fat tails" in the distribution of warming may pose problems for cost–benefit analysis as it may imply that society might be willing to exchange today's consumption for future consumption at an infinite rate (Weitzman (2009))) if the option for emissions control is present. It should be

[13] Gauss-Hermite quadrature is a deterministic integration method used to calculate the expectation operator this method uses predetermined integration nodes and weights.

emphasized that the implications of this paper are only meaningful under the following conditions: climate change is so uncertain that social welfare (carbon tax function) is unbounded in the absence of emissions control and it is possible to control greenhouse gas emissions and the level of emissions control is chosen optimally. To this end the magnitude of uncertainty needs to be measured in detail for the benefits and the costs of emissions control to be correctly estimated.

Veliz, Kaufmann, Cleveland and Stoner (2017) presented the first empirical estimates for the effect of climate change on electricity prices in Massachusetts USA. This paper finds that climate change alters the load duration curve [14] which raises prices. The size of this price will depend on the degree to which policy makers can create an environment that prompts generators, the distribution system and electricity consumers to adapt. Adaptation can be enhanced by policies aimed at electricity supply and consumption. On the supply side higher prices can be damped if policy creates a more certain environment for investment in new peaking capacity. On the demand size higher prices can be damped if policy favors energy conservation measures that reduce and/ or reschedule the electricity used for cooling. Authors use statistical models to translate the monthly changes in temperature that are forecast by climate models into monthly changes in electricity consumption and translate these monthly changes into hourly rates of electricity consumption using Monte Carlo techniques. After that the translation of hourly rates of electricity consumption into hourly prices, using statistical model that quantifies the relationship between hourly prices for and consumption of electricity, is done. At the end, the hourly forecasts for price and consumption is used to compute the effect of climate driven changes in temperature on electricity expenditures in Massachusetts.

2.4 Results and Future Research

In the latest Global Risk Report 2019 [15], published by the World Economic Forum, extreme weather events were ranked number 1 global risk with the highest likelihood and number 3 risk in terms of impact. Extreme weather events are perceived to bear higher risks than cyber security, weapons of mass destruction, data fraud and involuntary immigration.

As a result, it is important to include weather risk management tools such as natural weather hedging, insurance and weather derivatives into the general risk management strategy. From the literature review it is approved that weather derivatives can not be used like a substitute of weather insurance, but we can use them for different purposes or mix them together. Stefani et al (2018) proposed a hybrid contract that deals with a negative risk event, such as an insurance contract, but that is priced as a derivative instrument and this is done in case of temperature. In the future it would be interesting to study other hybrid contracts applied in other weather variables such as rainfall, snowfall, wind, etc. The weather derivative market is a classical incomplete market since the weather indices are not tradable assets, thus traditional no arbitrage pricing methods

[14] A load duration curve (LDC) is used in electric power generation to illustrate the relationship between generating capacity requirements and capacity utilization.

[15] http://www3.weforum.org/docs/WEF_Global_Risks_Report_2019.pdf

such as Black-Scholes as suggested in Meissner and Burke (2011) are not applicable in pricing weather derivatives. Further research needs to be executed on the valuation methods of weather derivatives. Today, there is no standard pricing model like the Black-Scholes model in conventional derivatives; having a standard pricing model could remove the discrepancies between the actual applied different models. But why should we be interested in weather derivatives? According to McDonald (2013) the first motive is speculation; using a derivative to construct a bet that is highly levered and tailored to a specific view. Initial costs of placing bets can be relatively small to the potential gains or losses from the bets. The second motive states that in some cases, derivatives provide a less costly financial outcome compared to combining underlying assets and will lead to reduced transaction costs. As a third motive, the author mentions the possibility to circumvent regulatory restrictions, taxes and accounting rules by trading derivatives, defined as regulatory arbitrage. The fourth and last motive, which it is considered as most vital, is risk management. For this report purpose, we are more interested in weather derivatives as a risk management tool, hedging weather risk/revenue affected by the weather.

Moreover, there is a need for analysing the different weather risk management tools empirically. Are weather derivatives in the long-run superior to insurance? Weather derivative settlement is objective and efficient, but can it add more value to a company than insurance can?

The relatively low correlation (even negative correlation) between weather derivatives and conventional financial assets suggests that weather derivatives can be excellent for diversification purposes. The use of derivatives on precipitation can have positive cash flow effect on industries with direct weather exposure, especially for businesses that have seasonal cash flows dependent on weather conditions. Also the presence of both production and price risk implies that options became a useful hedging tool. It is concluded as well that when the underlying uncertainty is non-linear in nature, the asymmetric payoff profile of options are more suitable for hedging purposes. We can say that there are some limitations to these contracts. Because the pay-off of a weather derivative depends on a weather index, not on the actual amount of money lost due to the weather, it is unlikely that the pay-off will compensate exactly for the money lost. The potential for such a difference is known as basis risk. In general, the basis risk is smallest when the financial loss is highly correlated with the weather and when contracts of optimal size and structure, based on the optimum location, are used for hedging. So, for a company to decide how to hedge its risk there is often a trade-off between basis risk and the price of the derivative.

Although it has been shown to use weather derivatives as a tool of risk management, there still exists imperfections and weaknesses that are important to address. Basis risk, as explained previously, is an imperfection that should always be taken into consideration.

Another challenge related to weather derivatives is collecting data, more specifically gathered data from multiple weather stations, none of them 100% representative the one we are interested in. Gathering data from different stations not located exactly at the place under study contributes to an increase in the geographical spatial/basis risk. In addition, some further research could be done on the costs of setting-up a large network of reliable weather stations.

For further research and improvement, it is recommended looking at the possibility of using multivariate, and/or non-linear regression when calculating the tick size used in derivative pricing. This could improve the adjusted coefficient of determination, which in turn should decrease basis risk. Using not only simple linear regression, would most likely further improve weather derivatives as a risk management tool.

Although it is shown that the level of critical temperature in CDD and HDD indices must be defined in an appropriate value according to the location of the station, we have to see and compare for more methods or techniques on how to obtain the proper value of the threshold.

All models can be refined for example in case of the indifference pricing approach by choosing a different utility function as power utility which means a different risk preference by the investor of by changing the basket of weather derivatives. Since weather derivatives are traded on and over the counter, one may compare prices as developed from different methods with those by actuarial approaches. In addition a sensitivity analysis can be applied on the models used to price weather derivatives. Moreover, an important issue to investigate is the analysis of the adoption of mixed-based weather derivatives, which in turn use composite weather indexes, since they have the potential to reduce basis risk.

Until now researcher suggest to use different models in contracts with payoff depends on an underlying index but still needed to clarify what model and in which case we can use a specific model which are the advantages and disadvantages in using the models used for example for rainfall, snowfall, etc. Some research work is done on modeling including temperature and rainfall variables but still it is necessary to do research on the possibility and the scope of building a composite index of rainfall and temperature. This would be useful for weather derivatives for hedging weather risk in some crops, whose growth is highly dependent on both rainfall and temperature. Except that it is interesting to study the effect of using artificial intelligence systems in forecasting weather variables, which can improve the accuracy from the learning process. Weather models developed until now from different researchers could be extended to include more complex dynamics of the behavior of temperature. Moreover, a comparison of different numerical techniques would be an interesting area to consider in the future. The least square optimization methodology may have applications for valuing previously intractable basket options in weather derivatives.

Another area where the research can be done is related to the appeal, the demand and the willingness to pay for weather derivatives in different businesses. This would help in structuring weather derivative products. After that it is interesting to study the effect of meteorological forecasts on the prices of weather derivative products. Of course, this could be done after a market for these products is in place. More research on portfolio effects of weather derivatives is needed. An important development would be to model more precisely the correlation between temperature and various commodity contracts. This could be extended to model the known lag between movements in load versus movements in prices.

We can conclude that weather derivatives are an essential tool in weather risk management. These instruments provide efficient and objective cash settlement and furthermore reduce the volatility of the financial performance triggered by weather variables. In some case studies it is indicated that the significant payoffs in some years are vital to cover the premium payments of other years in which the payoffs equal zero. As a consequence, only a long-term weather hedging horizon will prove to mitigate the weather exposure.

Bibliography

[1] Alaton, P., Djehiche, B., & Stillberger, D. (2002). On modelling and pricing weather derivatives. *Applied mathematical finance*, **9**(1), 1–20.

[2] Alexandridis, A., & Zapranis, A. D. (2012). *Weather derivatives: modeling and pricing weather-related risk*. Springer Science & Business Media.

[3] Alexandridis, A., & Zapranis, A. (2013). Wind derivatives: Modeling and pricing. *Computational Economics*, **41**(3), 299–326.

[4] Alexandridis, A., Gzyl, H., Ter Horst, E., & Molina, G. (2017). Extracting Risk Neutral Densities For Weather Derivatives Pricing Using The Maximum Entropy Method. *11th International Conference on Computational and Financial Econometrics*.

[5] Bellini, F. (2005). The weather derivatives market: Modelling and pricing temperature. Faculty of Economics (Doctoral dissertation, Ph. D. University of Lugano, Lugano, Switzerland).

[6] Benth, F. E. & Saltyte-Benth, J. (2005). Stochastic modeling of temperature variations with a view towards weather derivatives. *Applied Mathematical Finance*, **12**(1), 53–85

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

[8] Benth, F. E., Saltyte-Benth, J. & Koekebakker, S. (2008). Stochastic modelling of electricity and related markets. *World Scientific*.

[9] Benth, F. E., Härdle, W. K., & Cabrera, B. L. (2011). Pricing of Asian temperature risk. *In Statistical Tools for Finance and Insurance* (pp. 163–199). Springer, Berlin, Heidelberg.

[10] Benth, J. Š., & Benth, F. E. (2010). Analysis and modelling of wind speed in New York. *Journal of Applied Statistics*, **37**(6), 893–909.

[11] Benth, F. E., & Šaltytė Benth, J. (2011). Weather derivatives and stochastic modelling of temperature. *International Journal of Stochastic Analysis*.

[12] Benth, J. S. & Benth, F. E. (2012). A critical view on temperature modeling for application in weather derivatives markets. *Energy Economics*, **34**(2), 592–602.

[13] Benth, F. E., & Saltyte-Benth, J. (2012). *Modeling and pricing in financial markets for weather derivatives (Vol. 17)*. World Scientific.

[14] Benth, F. E., (2018). Weather Derivatives: Modeling and Pricing. *MAT4770 Energy stochastics*. https://www.uio.no/studier/emner/matnat/math/MAT4770/v18/timeplan/mat4770-carmatemp-v18.pdf [Accessed 30 January 2020].

[15] Benth, F. E., & Pircalabu, A. (2018). A non-Gaussian Ornstein–Uhlenbeck model for pricing wind power futures. *Applied Mathematical Finance*, **25**(1), 36–65.

[16] Beyazit, M. F., & Koc, E. (2010). An analysis of snow options for ski resort establishments. *Tourism Management*, **31**(5), 676–683.

[17] Biganashvili, T. (2013). Weather Derivatives as a Risk Management Tool (Doctoral dissertation).

[18] Björk, T. (2009). *Arbitrage Theory in Continuous Time*. Oxford university press.

[19] Bobojonov, I., Aw-Hassan, A., & Sommer, R. (2014). Index-based insurance for climate risk management and rural development in Syria. *Climate and Development*, **6**(2), 166–178.

[20] Boyd, M., Pai, J., Qiao, Z., & Ke, W. (2011). Crop insurance principles and risk implications for China. *Human and ecological risk assessment*, **17**(3), 554–565.

[21] Brockett, P. L., Wang, M., Yang, C., & Zou, H. (2006). Portfolio effects and valuation of weather derivatives. *Financial Review*, **41**(1), 55–76.

[22] Brody, D. C., Syroka, J., & Zervos, M. (2002). Dynamical pricing of weather derivatives. *Quantitative finance*, **2**(3), 189–198.

[23] Broni-Mensah, E. (2012). *Numerical solutions of weather derivatives and other incomplete market problems (Doctoral dissertation*, The University of Manchester (United Kingdom)).

[24] Campbell, S. D & Diebold, F. X. (2005) Weather forecasting for weather derivatives. *J. Am. Stat. Assoc*. **100**(469), 6–16.

[25] Cao, M., Li, A., & Wei, J. Z. (2003). *Weather derivatives: A new class of financial instruments*. Available at SSRN 1016123.

[26] Cao, M., & Wei, J. (2004). Weather derivatives valuation and market price of weather risk. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, **24**(11), 1065–1089.

[27] Chen, Y., & Hamwi, I. (2012). Why some disaster insurance does not exit? *Asia Pacific Journal of Risk and Insurance*, **6**(1), 1–14.

[28] Clarke, D. J. (2016). A theory of rational demand for index insurance. *American Economic Journal: Microeconomics*, **8**(1), 283–306.

[29] Dischel, B. (1998). At last: A model for weather risk. *Energy and Power Risk Management*, **11**(3), 20–21.

[30] Đorđevič, B. S. (2018). Hedging By Using Weather Derivates In Winter Ski Tourism. *Economics of Agriculture*, **65**(1), 125–142.

[31] Doms, J. (2017). Put, call or strangle? About the challenges in designing weather index insurances to hedge perfomance risk in agriculture. *German Association of Agricultural Economists (GEWISOLA)*.

[32] Dzupire, N. C., Ngare, P. & Odongo, L. (2019). Pricing basket weather derivatives on rainfall and temperature Processes. *International Journal of Financial Studies*. **7**(3), 35.

[33] Ender, M., & Zhang, R. (2015). Efficiency of weather derivatives for Chinese agriculture industry. *China Agricultural Economic Review*, **7**(1), 102–121. https://doi.org/10.1108/CAER-06-2013-0089

[34] Erhardt, R. (2016). Incorporating spatial dependence and climate change trends for measuring long-term temperature derivative risk. *Variance*, **9**(2), 213–226.

[35] Evarest, E., Berntsson, F., Singull, M., & Yang, X. (2018). Weather derivatives pricing using regime switching model. *Monte Carlo Methods and Applications*, **24**(1), 13–27.

[36] Fusco, G., Miglietta, P. P., & Porrini, D. (2018). How drought affects agricultural insurance policies: The case of Italy.

[37] Garman, M., Blanco, C., & Erickson, R. (2000). Seeking a standard pricing model. *Environmental Finance*, (3), 1–8.

[38] Cummins, J. D., & Geman, H. (1995). Pricing catastrophe insurance futures and call spreads: An arbitrage approach. *The Journal of Fixed Income*, **4**(4), 46–57.

[39] Geyser, J. M., & Van de Venter, T. W. G. (2001). Hedging maize yield with weather derivatives (No. 1737-2016-140384).

[40] Glasserman, P. (2013). *Monte Carlo methods in financial engineering (Vol. 53)*. Springer Science Business Media.

[41] Gyamerah, S. A., Ngare, P., & Ikpe, D. (2019). Hedging Crop Yields Against Weather Uncertainties—A Weather Derivative Perspective. *Mathematical and Computational Applications*, **24**(3), 71.

[42] Haugan, A. M., & Rasmussen, J. B. (2019). Pricing weather derivatives, and constructing portfolios that reduce the weather risk for ski resorts (Master's thesis, NTNU).

[43] Härdle, W. K., & López Cabrera, B. (2009). Infering the market price of weather risk. *In Discussion paper, SFB 649. Humboldt Universität zu Berlin*.

[44] Härdle, W. K. & Cabrera, B. L. (2012). The implied market price of weather risk. Applied Mathematical Finance, **19**(1), 59–95.

[45] Härdle, W. K. & Osipenko, M. (2017). A dynamic programming approach for pricing weather derivatives under issuer default risk. *International Journal of Financial Studies*, **5**(4), 23. https://doi.org/10.3390/ijfs5040023

[46] Hoornaert, A. (2018). Hedging Weather Risk Using Weather Derivatives (Doctoral dissertation, Ghent University).

[47] Hull, J. C. (2003). *Options, futures and other derivatives. 5a edicion*. Ed. Prentice Hall.

[48] Hull, J. C. (2005). *Fundamentals of Futures and Options Markets. 5th Edition*. Ed. Prentice Hall.

[49] Hwang, I. C., Tol, R. S. J. & Hofkes, M. W. (2016). Fat-tailed risk about climate change and climate policy. *Energy Policy*. 89, 25–35.

[50] Jewson. S & Brix. A (2005). *Weather derivative valuation: the meteorological, statistical, financial and mathematical foundations*. Cambridge University Press.

[51] Kabaivanov, S., & Markovska, V. (2017). Modelling environment changes for pricing weather derivatives. *Scientific Annals of Economics and Business*, **64**(4), 423–430.

[52] Kelly, D. L., Kolstad, C. D., & Mitchell, G. T. (2005). Adjustment costs from environmental change. *Journal of Environmental Economics and Management*, **50**(3), 468–495.

[53] Kendall, M. G., Stuart, A., & Ord, J. K. (1987). Kendall's advanced theory of statistics, vol 3: design and analysis, and time series.

[54] Keucheyan, R. (2018). Insuring climate change: new risks and the financialization of nature. *Development and Change*, **49**(2), 484–501.

[55] Le Den, X., Persson, M., Benoist, A., Hudson, P., de Ruiter, M., de Ruig, L., & Kuik, O. (2017). Insurance of weather and climate-related disaster risk: Inventory and analysis of mechanisms to support damage prevention in the EU.

[56] Lee, Y. & Oren, S. S. (2009). An equilibrium pricing model for weather derivatives in a multi-commodity setting. Energy Economics, **31**(5), 702–713.

[57] Luo, L. (2010). Snowfall derivative pricing: Index and daily modeling for the snowfall futures (Doctoral dissertation, Concordia University).

[58] Mahul, O., & Stutley, C. J. (2010). Government Support to Agricultural Insurance: Challenges and Options for Developing Countries. *The World Bank*.

[59] McDonald, R. L., (2013). Derivatives Markets. *3rd Edition ed. Boston: Pearson*.

[60] McIntyre, R., & Doherty, S. (1999). Weather risk-an example from the uk. *Energy and Power Risk Management June*.

[61] Missner, G. & Burke, J. (2011). Can we use the Black-Scholes model to value temperature options? *International Journal of Financial Markets and Derivatives* **2**(4). 298–313.

[62] Møller, T. (2001). On transformations of actuarial valuation principles. *Insurance: Mathematics and Economics*, **28**(3), 281–303.

[63] Porrini, D., & Schwarze, R. (2014). Insurance models and European climate change policies: an assessment. *European Journal of Law and Economics*, **38**(1), 7–28.

[64] Prabakaran, S., & Singh, J. P. (2017). Modeling and Pricing of Weather Derivative Market. *Global Journal of Pure and Applied Mathematics*, **13**(12), 8103–8126.

[65] Rencher, A. C., & Schaalje, G. B. (2008). *Linear models in statistics*. John Wiley & Sons.

[66] Ritter, M., Musshoff, O., & Odening, M. (2014). Minimizing geographical basis risk of weather derivatives using a multi-site rainfall model. *Computational Economics*, **44**(1), 67–86.

[67] Rozante, J. R., Gutierrez, E. R., da Silva Dias, P. L., de Almeida Fernandes, A., Alvim, D. S., & Silva, V. M. (2019) Development of an index for frost prediction: Technique and validation. *Meteorological Applications*.

[68] Rubinstein, M. (2000). On the relation between binomial and trinomial option pricing models. *The Journal of Derivatives*, **8**(2), 47–50.

[69] Šaltytė Benth, J., & Šaltytė, L. (2011). Spatial–temporal model for wind speed in Lithuania. *Journal of Applied Statistics*, **38**(6), 1151–1168.

[70] Schiller, F., Seidler, G., & Wimmer, M. (2012). Temperature models for pricing weather derivatives. *Quantitative Finance*, **12**(3), 489–500.

[71] Spillet, R (2001): Don't Blame the Weather, Futures and Options World.

[72] Stefani, S., Moretto, E., Parravicini, M., Cambiaghi, S., Sonubi, A., Kutrolli, G., & Tulli, V. (2018). Managing adverse temperature conditions through hybrid financial instruments. *Journal of Energy Markets*, **11**(3), 25–41.

[73] Stojanovic, S. (2011). *Neutral and Indifference Portfolio Pricing, Hedging and Investing: With Applications in Equity and FX*. Springer Science Business Media.

[74] Surminski, S., & Hudson, P. (2017). Investigating the risk reduction potential of disaster insurance across Europe. *The Geneva Papers on Risk and Insurance-Issues and Practice*, **42**(2), 247–274.

[75] Szabó, P., & Szépszó, G., (2016). Quantifying Sources of Uncertainty in Temperature and Precipitation Projections over Different Parts of Europe. In: Mathematical Problems in Meteorological Modelling. Mathematics in Industry (eds.: Bátkai, A., Csomós, P., Faragó, I., Horányi, A., Szépszó, G.), *Springer International Publishing*, 207–237.

[76] Taib, C. M. I. C. & Benth, F. E. (2012). Pricing of temperature index insurance. Review of development finance, **2**(1), 22–31.

[77] Tang, C. H. & Jang, S., (2011). Weather risk management in ski resorts: Financial hedging and geographical diversification. *International Journal of Hospitality Management*, **30**(2), 301–311.

[78] Tang, C. H. & Jang, S. (2012). Hedging Weather Risk in Nature-Based Tourism Business: An Example of Ski Resorts. *Journal of Hospitality Tourism Research*, **36**(2). 143–163.

[79] Undli, H. & Schatvet, M. (2018). Using weather derivatives to hedge precipitation exposure for a Norwegian hydropower producer. *MSTHesis, NTNU*.

[80] Véliz, K. D., Kaufmann, R. K., Cleveland, C. J., & Stoner, A. M. (2017). The effect of climate change on electricity expenditures in Massachusetts. *Energy Policy*, 106, 1–11.

[81] Wang, Z., Li, P., Li, L., Huang, C., & Liu, M. (2015). Modeling and forecasting average temperature for weather derivative pricing. *Advances in Meteorology*.

[82] Weitzman, M. L. (2009). On modeling and interpreting the economics of catastrophic climate change. *The Review of Economics and Statistics*, **91**(1), 1–19.

[83] Wilks, D. S. (1998). Multisite generalization of a daily stochastic precipitation generation model. *Journal of Hydrology*, **210**(1-4), 178–191.

[84] Xu, W., Odening, M., & Musshoff, O. (2008). Indifference pricing of weather derivatives. American Journal of it Agricultural Economics, **90**(4), 979–993.

[85] Yuan, W., Göncü, A., & Ökten, G. (2015). Estimating sensitivities of temperature-based weather derivatives. *Applied Economics*, **47**(19), 1942–1955.

[86] CME Group. (2011). The Weather Derivatives Markets at CME Group: A brief history.

[87] CME Group. (2018). Weather Futures and Options Codes. Retrieved from Chicago Mercantile Exchange: http://www.cmegroup.com/trading/weather/files/updated-weather-product-slate.pdf

Chapter 3

Managing adverse temperature conditions through hybrid financial instruments

Joint work with Silvana Stefani, Enrico Moretto, Matteo Parravicini, Simone Cambiaghi, Adeyemi Sonubi, and Vanda Tulli
(Published in the *Journal of Energy Markets*, 2018, Volume 11, pp. 25-41)

Abstract

Recent international policy initiatives focus on reducing carbon emissions to limit warming. It is almost universally recognized that risks connected to climatic changes are unpredictable in their consequences. Moreover, attempts (for instance the 2016 Paris conference) to manage climatic changes at a global level have been counterbalanced by a not clear-cut US policy. Surprisingly, the financial world does not seem to care much about this problem. Yet, it is estimated that 80% of world industries (i.e. agriculture, construction sector and hospitality activities) are affected (totally or in part) by climate. Rain or low temperatures disrupt tourism; heavy rain or high temperatures devastate crops and damage farmers. This work contributes to existing literature by proposing a temperature-based risk management using hybrid financial instruments based on weather derivatives. Based on well-established literature we firstly model temperature time series; we then price one-month forward option contracts for hedging adverse outcomes. Our results exploit daily temperature data-set (1951-2016) collected in Arezzo, Italy. We then show how a "negative" weather performance can be counterbalanced by the "positive" performance of the hedging Over-The-Counter financial instrument that can be tailored to meet specific needs.

Keywords: Climate change, weather derivatives, temperature, risk hedging

3.1 Introduction

As it is very well known, climate change is one of the most important and crucial environmental, political as well as economic topics of the 21st century. It already affects and will affect in the future the entire spectrum of life on planet Earth.

To tackle this issue, huge efforts have been done and are still undergoing on a worldwide basis. Mentioning only the most recent acts, the so called 'Paris Agreement', that has become effective in November 2016 and has been signed by almost 200 countries, has stated that all parties should "*hold the increase in the global average temperature to well below* $2°C$ *above pre-industrial levels and pursue efforts to limit the* *temperature increase to* $1.5°C$ *above pre-industrial levels, recognizing that this would significantly reduce the risks and impacts of climate change*" [**Paris**].

Two straight effects of climate change are sequences of longer and more intense heat waves, more frequent damaging storms, and a likely average raise in temperatures above 2°C. This, of course, forces farms, companies as well as public authorities to consider to hedge against climate changes to mitigate the resulting risk. To acknowledge the important of 'adverse weather' risk it is useful to recall that, since 1997, the Chicago Mercantile Exchange (CME [1]) allows traders to negotiate weather derivatives that go under the names of Heating Degree Day (HDD), Cooling Degree Day (CDD), and Cumulative Average Temperature (CAT) (CME Group, 2017). Such contracts rely on the fact that the most relevant issue in managing weather risk is some uncommon behavior of temperatures. Further, hedging has been done traditionally also through insurance contracts whose effectiveness, however, is limited to clearly identified extreme events.

Loosely speaking, insurance contracts can, somehow, be intended to be similar to financial derivatives as one party is willing to sell some risk and the other accepts to buy it.

In financial literature, a number of methodologies (namely Burn analysis, Index modeling and Daily simulation; see, for instance, Jewson, Brix, & Ziehmann (2005), Alaton, Djehiche, & Stillberger (2002), and Schiller, Seidler, & Wimmer (2012)) has been extensively applied. Burn analysis is based on historical data and determine the price of a weather derivative adding to the obtained value some risk loadings as it is common practice in the actuarial field. Despite the triviality of this technique, Burn analysis is a widely used pricing methodology, especially if the data set is large enough.

It is important to point out that weather derivatives are usually more suited to protect against minor fluctuations from the normal status of the underlying variable, being most effective in a sort of low risk scenario. To achieve a proper hedging against high risk instances, instead, insurance contracts are to be obviously preferred.

If, on one hand, weather derivatives deal with a small fraction of sources of weather risk, such contracts benefit of a liquid, regulated market like for instance CME; their prices are, therefore, closer to the correct value of the traded risk. Insurance contracts, instead, are very illiquid but they can cover, for instance, farmers from peculiar damages such the one due to hailstorms and can be very flexibly shaped to risks localized in a very small area. On the other hand, damages recognition and validation involve bureaucratic paperwork and are subject to the final decision of a claim handler. This leads to a delay or a reduction in damage payments, as well as uncertainty in the final payment.

In both cases, i.e. buying a derivative contract or an insurance contract, a crucial issue is the goodness and source of collected historical data; this is an important motivation in this article, since there is a gap in dealing with a meteorological risk by means of a weather derivative rather than an *ad-hoc* insurance contract. In the first case, more marketable contracts are penalized by the fact that they consider only an 'average' risk and might run short of a reliable coverage. Just to give an example, temperature in a US city which is the underlying for a future in CME, can be very different than temperature in South Italy, thereby making it useless for a farmer buying such contract

[1] http://www.cmegroup.com/trading/weather/

for a coverage. In the latter case the cost of an insurance agreement, as it recalls a 'tailor-made' *over-the-counter* derivative whose price is not a market price, generally ends up being larger. However, its specificity allows, in most cases, for a more effective risk reduction.

To the best of our knowledge, so far no effort has been done by insurance companies as well as banks or other financial institutions to offer a hybrid instrument that could link derivatives to insurance contracts. The aim of this contribution is, then, to suggest a methodology that, exploiting local data and heavily relying on results by F.E. Benth (Benth & Benth (2013), Benth & Benth (2011), Benth & Benth (2005)), allows to find the price of an insurance or derivative contract applying well established derivative pricing methodologies.

The numerical section of this paper exploits and benefits from a detailed and very ample data-set that spans from 1951 to 2016, for a total of 66 years available in SIR [2]. In particular, daily max and min temperature data for the town of Arezzo (Tuscany, Italy) have here been considered.

The main result of this paper is to bridge the two above mentioned approaches and to lead to the determination of prices for temperature derivatives dealing with meteorological risk in the Arezzo area. The approach is flexible enough to allow considering the actuarial point of view, where risk loadings have to be properly considered and charged.

This article is structured as follows: Section 3.2 recalls the main points in pricing weather derivatives, Section 3.3 deals with their stochastic modeling, Section 3.4 proposes the valuation of one-month forward options for two different months and shows their effectiveness in hedging meteorological risk. Section 3.5 concludes.

3.2 Basic Concepts on Temperature Derivatives

Weather derivatives are usually structured as futures, forwards, options and swaps based on different underlying weather indexes. In this paper, in line with derivatives quoted in the contracts, we will work with OTC instruments, in particular, forward option contracts. Our analysis is focused on derivative products whose underlying are the HDD, CDD and the level of daily accumulated temperatures over a given period Benth & Benth (2013). The average temperature for a given day t is calculated as the mean of the recorded maximum and minimum temperature. Given a weather station, let $T_{min}(t)$ and $T_{max}(t)$ denote the minimum and maximum temperatures measured in one day t. The average temperature in day t is defined by

$$T(t) = \frac{T_{max}(t) + T_{min}(t)}{2} \tag{3.1}$$

For a given site, the degree days are the difference of the daily average temperature from the base temperature (in general 65 degrees Fahrenheit or 18 degrees Celsius). An HDD is the number of degrees by which the day's average temperature is below the base temperature, while a CDD is the number of degrees by which the day's average temperature is above the base temperature. Cooling degree days and heating degree

[2] http://www.sir.toscana.it/

days are never negative. Thus, if the daily average temperature is less than 18^oC, HDD will accumulate for the period, and if the daily average temperature is greater than 18^oC, CDD will accumulate. Consequently, HDD and CDD are calculated as follows:

$$HDD(t) = \max(18 - T(t), 0) \tag{3.2}$$
$$CDD(t) = \max(T(t) - 18, 0) \tag{3.3}$$

where $T(t)$ is the current temperature. The HDD and CDD indexes are the aggregated indexes over an agreed period of time that is called measurement period. CAT index is the cumulative average temperature, and this index is used to substitute CDD index (since in many cites the average daily temperature is hardly above 18^oC) in summer and in winter HDD for the same reasons.

$$HDD(t_1, t_2) = \sum_{t=t_1}^{t_2} HDD(t) \tag{3.4}$$
$$CDD(t_1, t_2) = \sum_{t=t_1}^{t_2} CDD(t) \tag{3.5}$$
$$CAT(t_1, t_2) = \sum_{t=t_1}^{t_2} T(t) \tag{3.6}$$

These indexes are accumulations of daily HDDs and CDDs, over a month or an entire season. In this study, one-month forward option contracts for which the underlying is an index linked to specific weather events and the forward prices, are evaluated to build hedging strategies over one month. Therefore, we have a contract which "delivers" HDD over a specified monthly duration valid for a given month, in return for an agreed forward price. The HDD is calculated from temperatures in the given month. There are four basic elements in a contract: (i) the underlying variable, HDD or CDD; (ii) the contract period; (iii) the meteorological station from which the temperature data are recorded; (iv) the tick size that is the value in EUR of one HDD or CDD unit.

As mentioned before, temperature derivatives are the most commonly weather contracts traded on the market. Our basic pricing framework is shown as follows:

1. Collect historical daily average temperature data;
2. Construct a technique for modeling temperature time series;
3. Building scenarios through Monte Carlo Method;
4. Calculate the price of the contract, adding the risk loading factor.

In this paper, we evaluated a one month forward HDD based on temperature daily data in the city of Arezzo (Tuscany) in the timespan 1951-2016. The payoff of HDD for temperature is

$$X(\tau_1, \tau_2) = k \times \sum_{s=\tau_1}^{\tau_2} \max(c - T(s), 0) \tag{3.7}$$

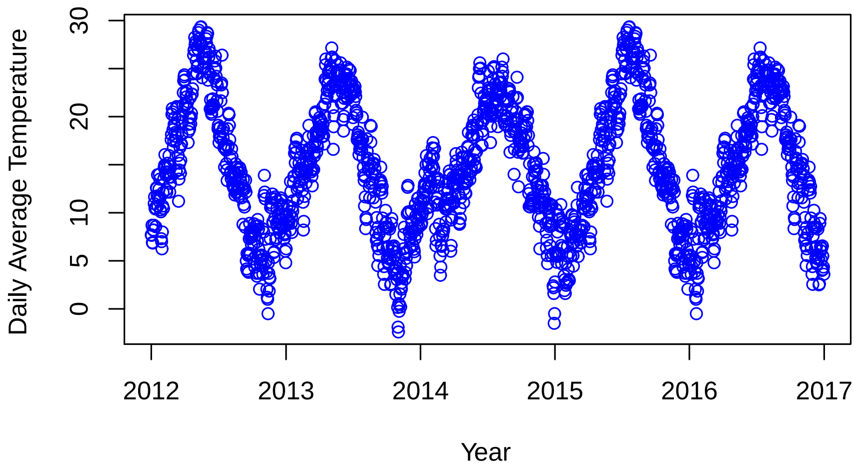


Figure 3.1: Time series of daily average temperatures of Arezzo, snapshot of the 2012-2016 period

where k is the multiplier factor/tick size and it is set to 20 EUR, and c is the threshold. While in general c is set as 18^oC, it is necessary to find out the most appropriate threshold for the calculation of the HDD index. As threshold we use the tenth percentile (or first decile) of the daily average empirical temperature. As will be illustrated later the threshold used for January, 2017 is calculated to be 2.43^oC. Figure 3.1 shows the daily average temperature in Arezzo for the last 5 years (2012-2016), where it is possible to observe a clear seasonality of the data.

3.3 Stochastic Modeling of Weather Derivatives

3.3.1 Data Collection

The dataset consists of the daily maximum and minimum temperatures in Celsius (°C) from the weather station Molin Bianco in Arezzo, Tuscany (WMO ID: 16172, latitude: 43°27'34"34.81"N, longitude: 11°50'44.5"E, and elevation: 248 meters above sea level) over the timespan 1951-2016.

The daily average temperature (DAT) is then calculated using a simple arithmetic mean of max and min temperatures. From a physical-meteorological point of view, such calculation for the daily average temperature would be inaccurate, since one should consider the temperatures of a day at different times. For the purposes of this paper (i.e. pricing of derivatives having as underlying the temperature index) however, it can be efficiently proceeded with the calculated daily average temperature. The sample, consisting of 24,107 observations, shows an absolute average maximum of 34°C recorded on June 18, 1990, when the maximum temperature was 39°C and the minimum 29°C and an absolute average minimum of -8.10°C registered on October 30, 1993, when the maximum temperature was -1.2°C and the minimum temperature was -15°C. The average temperature of the whole sample is 13.9°C.

The empirical distribution is described with a histogram in Figure 3.2.

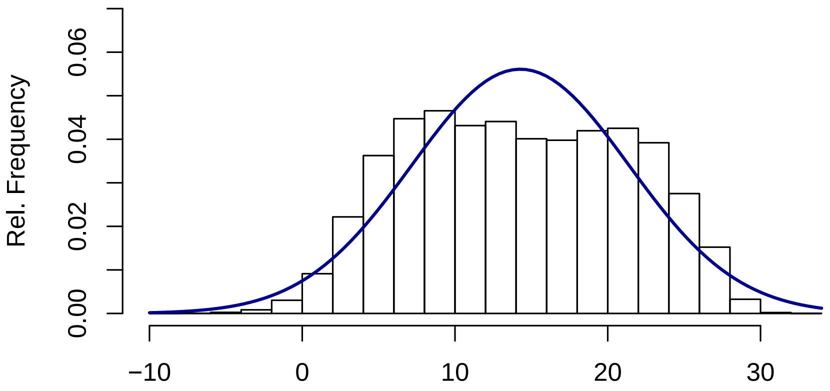


Figure 3.2: Histogram of the daily average temperature of Arezzo (Tuscany) (1951-2016).

Table 3.1: Descriptive Statistics of the daily average temperature in Arezzo (Tuscany) 1951-2016.

	Mean	Var	Std Dev	Min	Max	Skew.	Kurt.
Temperature(°C)	13.869	50.455	7.103	-8	34	0.019	2.105

Table 3.1 presents some descriptive statistics for the daily average temperatures. It can be noted that the distribution of the temperature is not normal since there exist no proper peak in the distribution.

The data set is used to construct a technique for modeling temperature. We start from a simple linear regression for the long run trend. Then, we track and insert in the model the seasonal component. Subsequently, to make sure that the model is able to incorporate the effect of past data, we will proceed with an autoregressive model. At the end, a residual analysis will be implemented and a verification of the goodness of the model will be made by evaluating what the model is able or not to capture.

3.3.2 Temporal Modeling for temperature

To model the temperature dynamics we use a time series decomposition approach. The temperature $T(t)$ is decomposed into a mean component $\mu(t)$, which models the trend, and a residual component $\epsilon(t)$ which models the fluctuations around a trend over time. The temporal decomposition of $T(t)$ can be written as:

$$T(t) = \mu(t) + \epsilon(t) \tag{3.8}$$

where $\mu(t)$ is a deterministic function over time that is defined as follows:

$$\mu(t) = \theta(t) + \sum_{i=1}^{p} \alpha_i (T(t-i) - \theta(t-i)) \tag{3.9}$$

where $\theta(t)$ represents the (linear) trend and seasonality in terms of the sine/cosine functions in $T(t)$, and α_i are the parameters of an AR(p) process. The variables $\epsilon(t)$ represent the residuals at time t of the following form:

$$\epsilon(t) = \sigma(t)\gamma(t) \tag{3.10}$$

Where $\sigma(t)$ is a (possibly) time–dependent volatility function, and $\gamma(t)$ is a zero-mean temporally independent Gaussian random process with standard deviation equal to one.

3.3.3 Estimation and Validation of Temperature model

For modeling the temperature dynamics in Arezzo (Tuscany), we apply a time series decomposition approach. Therefore, the temperature time series is decomposed into components, such as trend, seasonality, AR process and residual term. All these components appear in the data simultaneously as you can see in Equations 3.8 and 3.9. By estimating and eliminating the different components of the time series step by step and examining the residuals at each step, we can obtain a good fit of data to the proposed model.

3.3.3.1 Trend

We choose to model the trend as a linear function of time ensuring stationarity in our temperature time series. Even though the linear trend might be a simplification when a long time series of temperature is considered, a constant trend from year to year seems to be validated in our case. At first, we run a linear regression. The obtained slope is significantly different from zero, meaning that the temperature has increased in Arezzo by about 0.5°C in fifty years (Figure 3.3).

The positive trend corresponds to the increase in the global mean temperature. Furthermore, meteorological phenomena are very much influenced by the geographical location, and this is in fact one of the main causes of the weaknesses in going to perform the pricing of weather derivatives.

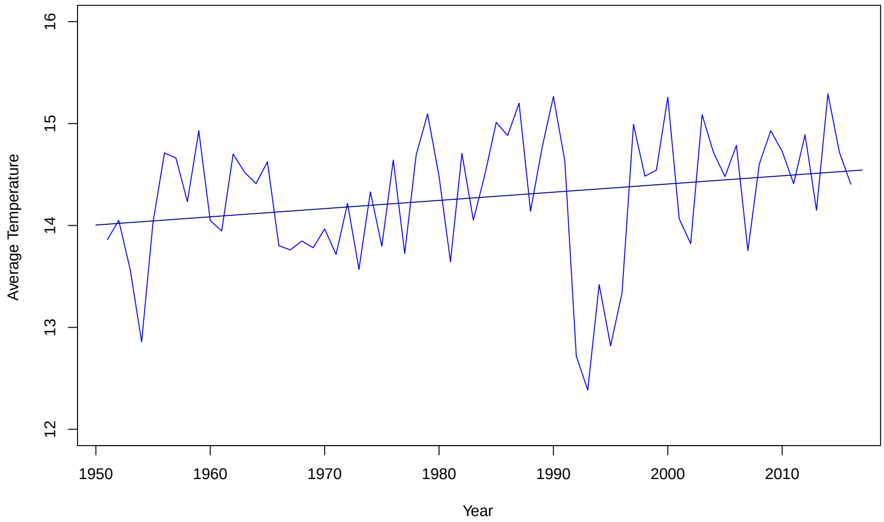


Figure 3.3: The linear trend in temperatures in Arezzo, Tuscany (1951-2016).

3.3.3.2 Seasonal Component

The deterministic function $\theta(t)$ is modeling the trend and seasonality of temperature. A truncated Fourier series is sufficiently flexible to describe temperature:

$$\theta(t) = a_0 + a_1 t + a_2 \cos\left(\frac{2\pi t}{365}\right) + a_3 \sin\left(\frac{2\pi t}{365}\right) \tag{3.11}$$

Such a sum of trigonometric functions explains the seasonal variations in the temperature, such as low temperatures in the winter and high in the summer. According to the standard statistical significance tests, we conclude that a rather low order truncated Fourier series explains well the seasonal variations in the temperature. The estimates of the fitted function are reported in Table 3.2 for Arezzo. As all confidence intervals do not contain

Table 3.2: Estimated coefficients and 95% confidence interval under OLS for formula (3.11)

	a_0	a_1	a_2	a_3
estimated parameter	14.0185	2.1814×10^{-5}	-8.016	-1.0917
95% conf.int lower bound	13.9119	1.4157×10^{-5}	-8.0914	-1.1672
95% conf.int upper bound	14.1251	2.95×10^{-5}	-7.9407	-1.0164

0, parameters in Table 3.2 are significantly different from zero.

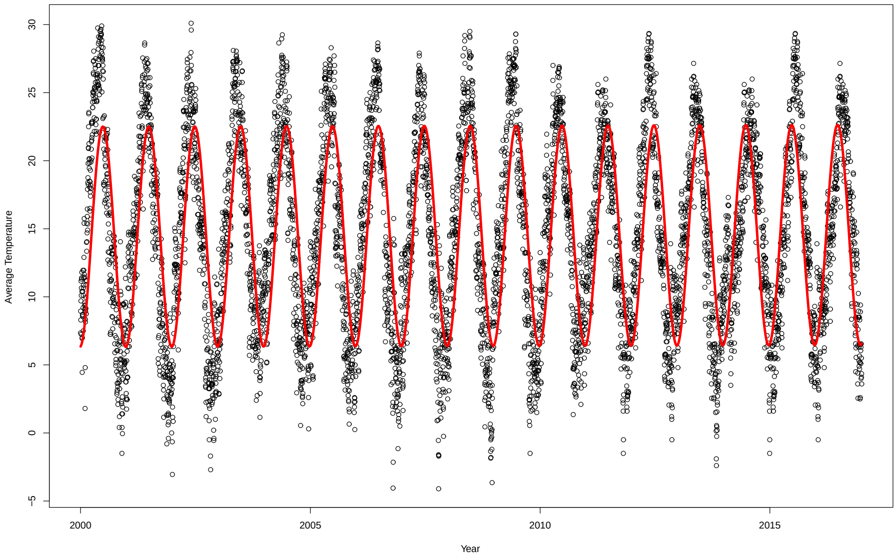


Figure 3.4: Fitted trend (red), seasonality and AR(3) process.

3.3.3.3 AR Process

In this section we eliminate the estimated trend and seasonal effects from the temperature data and find the best AR(p) model fitting the deseasonalized temperature data. Referring to the temperature model in Equation (3.9) we consider an AR(3) under the hypothesis that the temperatures of day t can be correlated with those of the days $t-1, t-2, t-3$. We also estimated other AR(p) processes on deseasonalized data, and found that AR(3) fitted the data better. In Table 3.3 the estimated values of those parameters are reported.

Estimates α_1, α_2 and α_3 are significant at the 1% level. As can be seen from Figure 3.4, adding the self-regulating econometric process in Equation (3.9) and (3.11), the fitness is improving.

Table 3.3: Estimated coefficients from AR(3).

	α_1	α_2	α_3
Estimated values	0.809776834	0.00933351	0.079971453
p-value	0	0	$1.684667 \cdot 10^{-35}$

3.3.3.4 Estimation of Volatility

One way to model the variance $\sigma^2(t)$ is to fit a truncated Fourier series

$$\sigma^2(t) = b_1 + \sum_{l=1}^{4} b_{2l} \cos\left(\frac{2\pi\ell t}{365}\right) + b_{2l+1} \sin\left(\frac{2\pi\ell t}{365}\right) \tag{3.12}$$

to the empirical daily variance of the residuals. The alternative is to use the daily empirical variance of residuals to approach the properties of a white noise. First, we have calculated the daily variance of empirical residuals, remembering that for each day of the year in our case we have 66 values, equal to the number of samples in years. After that, we calculate the estimated OLS residuals. On the left hand side of Figure 3.5, the empirical variance $\sigma^2(t)$ together with the estimated one is presented. It is noted that the variation in temperatures in the cold season is considerably more marked than during the hot season. The variability also increases in the period between late spring and early summer, as well as early autumn, in short in the transition between the cold and warm seasons and vice versa.

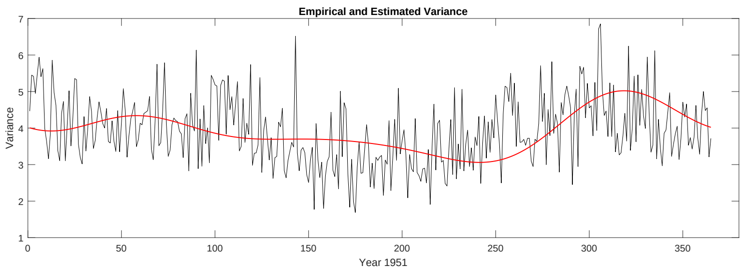


Figure 3.5: Empirical (black line) and estimated (red curve) variance of the residuals of the temperature about Arezzo in 1951

At the end it is possible to deduce the residual process by dividing the components (Figure 3.5):

$$\epsilon(t)/\sigma(t) \tag{3.13}$$

From table 3.4 it is discovered that the behaviour of residuals is that of a normal standard. Moreover Figure 3.6 shows the histogram of residuals.

Table 3.4: Normality Tests of Residuals.

	JB *	Pearson (χ^2)	Kurt.	Skew.	Lilliefors (KS *)
statistic value:	7.982	12.150	1.800	-5.32e-12	0.060
p - value:	0.029	0.434	0.0125	1	0.282

JB denotes Jaque-Bera. KS denotes Kolmogorov-Smirnov.

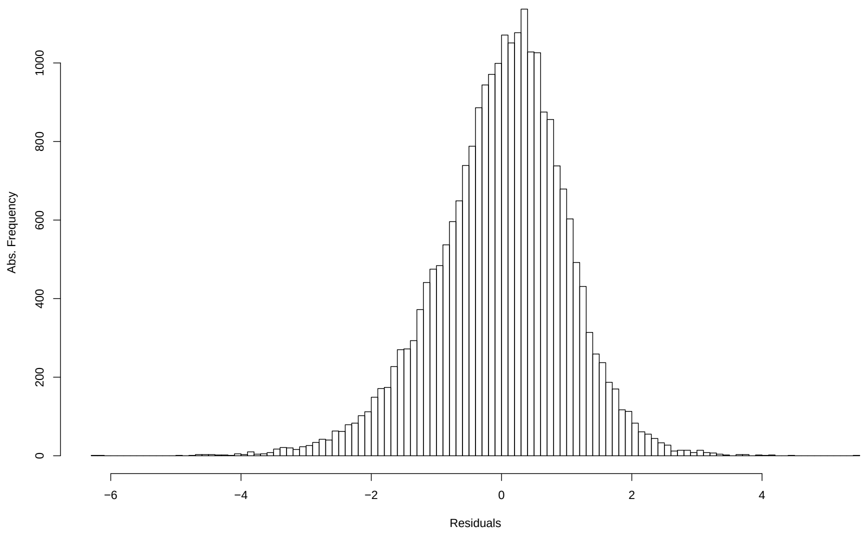


Figure 3.6: Histogram of overall residuals

3.3.3.5 Validation of the Model

To detect potential remaining dependencies in the residuals we analyze the ACF of the squared residuals. From the ACF of the squared residuals Figure 3.7, a clear seasonal pattern is observed indicating a time – dependency in the variance of residuals. After having eliminated the seasonal dependency by dividing the residuals by the square root of the fitted variance, the resulting ACF is shown in Figure 3.8. The graph basically shows that we are left with zero-mean uncorrelated in time series which is close to a normal distribution.

After removing seasonality from time series we validate the model by comparing empirical values with estimated values from the model.

3.4 Pricing of Temperature

As for HDD forward option contract the price is defined as:

$$P(t,\tau_1,\tau_2) = \exp\big(-r(\tau_2 - t)\big)E\big[X(\tau_1,\tau_2)|F_t\big] \tag{3.14}$$

where $X(\tau_1,\tau_2)$ (according to expressions in Equation (3.7)) represents the average value of considered payoffs for temperature. We use the Monte Carlo method to simulate the stochastic component of the model given to Equation (3.10) for temperature. For this reason we have chosen to use a fixed number of paths that means that for each case we are going to simulate $1\ 000$ values. For each of these paths we calculate the

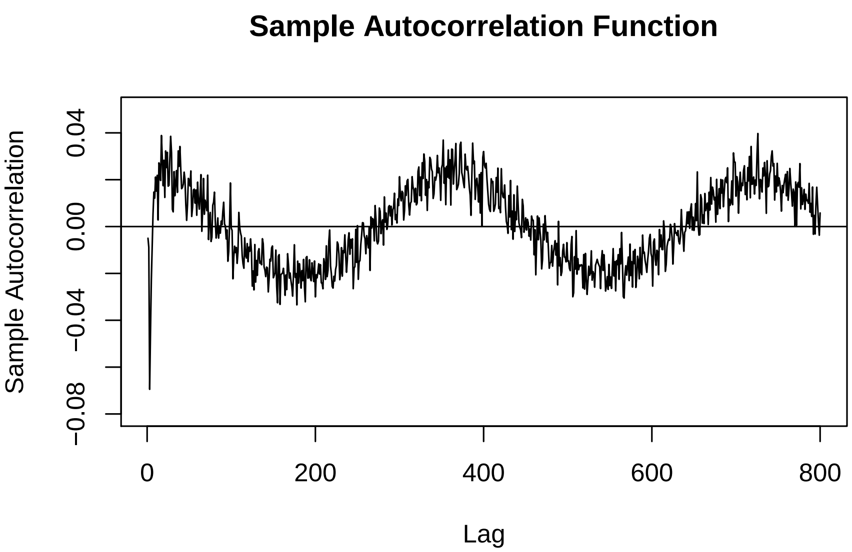


Figure 3.7: Autocorrelation function of empirical residuals

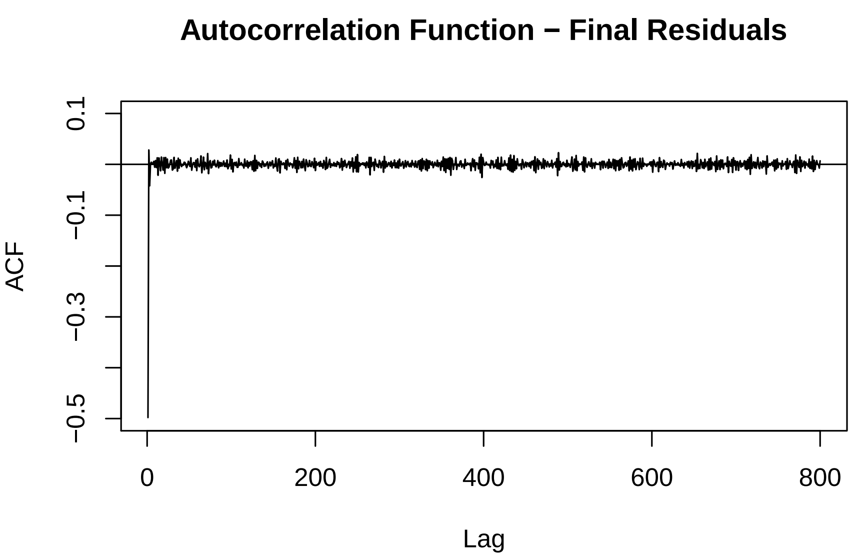


Figure 3.8: Autocorrelation function of final residuals

payoff. Then, the average of the payoffs from all the generated paths will represent the expected price of the derivative $X(\tau_1, \tau_2)$. Concerning temperature this value has resulted **31.99** EUR. After that, we proceed to calculate the price of the derivative through the discounting formula. Referring to the current discount rates in Italy, r will be set equal to 0 and, consequently, the current value will correspond to the simple average value of X. Therefore, the price of the contract for temperature is **139.58** EUR (for more details see Table 3.5).

But, to this price the component called risk loading must be added, which has the purpose of guaranteeing to the insurance institution/ bank a profit margin for the risk assumed. The value of the component is obtained, by calculating 5% of the 95% percentile of the payoff distribution, which, in this case, is equal to **34.74** EUR. As a result, the final price will be **174.32** EUR.

Table 3.5: Details on calculation of the final price for 67 years - Month of January (Table continue on next page).

Year	MC-Monthly-HDD	Monthly-Payoff	Final Payoff
1951	0	0	
1952	4.4	88	
1953	7.8	156	
1954	37.2	744	
1955	8.81	176.2	
1956	4.84	96.8	
1957	3.87	77.4	
1958	8.92	178.4	
1959	4.04	80.8	
1960	6.11	122.2	
1961	80.02	1600.4	
1962	6.58	131.6	
1963	0	0	
1964	25.15	503	
1965	24.31	486.2	
1966	39.94	798.8	
1967	1.63	32.6	
1968	6.22	124.4	
1969	0	0	
1970	0	0	
1971	0.43	8.6	
1972	0	0	245.7
1973	8.97	179.4	
1974	3.36	67.2	
1975	0.93	18.6	
1976	26.44	528.8	
1977	14.69	293.8	
1978	27.37	547.4	
1979	0	0	
1980	2.43	48.6	
1981	3.79	75.8	
1982	59.52	1190.4	
1983	2.18	43.6	
1984	20.48	409.6	

Table 3.5: Continued.

Year	MC-Monthly-HDD	Monthly-Payoff	Final Payoff
1985	0	0	
1986	0.61	12.2	
1987	3.79	75.8	
1988	2.36	47.2	
1989	0	0	
1990	1.43	28.6	
1991	0	0	
1992	0	0	
1993	3.89	77.8	
1994	0	0	205.68
1995	0.73	14.6	
1996	0	0	
1997	0	0	
1998	0	0	
1999	0	0	
2000	0	0	
2001	0.68	13.6	
2002	3.39	67.8	
2003	0	0	
2004	0	0	
2005	0	0	
2006	0	0	
2007	0.78	15.6	
2008	0	0	
2009	0	0	
2010	0	0	
2011	0	0	
2012	0	0	
2013	0	0	
2014	0	0	
2015	1.36	27.2	
2016	6.55	131	141.21
2017	2	31.99	139.58

3.4.1 Application of an hedging strategy analysis for temperature

Let consider a farmer in Arezzo who has signed a one month temperature contract with a bank in January 1, 2017 to protect himself from excessively cold temperatures which could lead to frost with consequent damages on his crops. The stipulated contract is as it is described above, with a price of **174.32** EUR. At the end of contract period the historical temperatures as of January, 2017 (obtainable from the SIR (Figure 3.1) are recorded and the payoffs are calculated. From 6 to 16 of January there was a particularly icy period that brought the daily average temperatures below the "minimum" threshold (which we remember had been set at 2.43řC) thus "triggering" the protection provided by the derivative, which in fact in those days will have earned **451.60** EUR. On the other hand, a similar one month contract, signed for February 2017, gave a different payoff. Favorable weather conditions did not allow the HDD to trigger and a loss is faced by the farmer. But of course this is counterbalanced by meteorological conditions favorable for the crops. In this way, considering the collection obtained net of the price paid, for January the farmer will have obtained a profit of **277.28** EUR. The farmer then has covered itself efficiently against adverse weather conditions. A summary of the obtained results are shown in Table 3.6.

Table 3.6: Hedging strategy analysis for January 2017.

January - HDD with threshold 2.43řC	
Price	139.58 EUR
Risk Loading	34.74 EUR
Final Price	174.32 EUR
Specific Case: January 2017	
Price	174.32 EUR
Payoff	451.60 EUR
Profit	277.28 EUR

3.5 Conclusions

The aim of this paper is to propose an hybrid contract capable of dealing with some 'negative event' risk, as an insurance contract, but, at the same time, being priced as a derivative instrument. By numerical techniques, a contract capable of hedging against temperature risk in a limited geographical area (i.e. the town of Arezzo in Tuscany) has been priced. Results presented in Section 3.4 are promising as numerical findings show that such contracts can positively cover temperature risk. This is not only a theoretical achievement as financial institutions (for instance rural banks) can offer effective hedging instruments.

Results presented in this paper can be applied, *mutatis mutandis*, to other kinds of weather risk sources, such as excessive rainfall.

Declaration of Interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Acknowledgment

The authors of this paper gratefully thank the staff at the *Servizio Idrologico Regionale della Regione Toscana* (SIR) for its excellent management of data source.

Bibliography

[1] Authentic texts of the Paris Agreement

[2] Benth, F.E., Benth, J.S. (2013). *Modeling and Pricing in Financial Markets for Weather Derivatives*. Advanced Series on Statistical Science & Applied Probability. Volume 17 World Scientific

[3] Alaton, P.; Djehiche, B.; Stillberger, D. (2002). On modelling and pricing weather derivatives, *Applied Mathematical Finance*, **9**(1), 1–20.

[4] Benth, F.E., Benth, J.S. (2011). *Weather derivatives and stochastic modelling of temperature*, International Journal of Stochastic Analysis.

[5] Benth, F.E., Benth, J.S. (2005). Stochastic modelling of temperature variations with a view towards weather derivatives, *Applied Mathematical Finance*, **12**(1), 53–85.

[6] Schiller, F., Seidler, G., Wimmer M. (2012). Temperature models for pricing weather derivatives. *Quantitative Finance*, **12**(3), 489–500.

[7] Jewson, S., Brix, A., & Ziehmann, C. (2005). Weather derivative valuation. *Cambridge University Press*.

Chapter 4

Managing Meteorological Risk through Expected Shortfall

Joint work with Silvana Stefani, Enrico Moretto and Sergei Kulakov
(Published in *Risks*, 2020, Special Issue of Stochastic Modeling and Pricing in Energy Markets, Volume 8, pp. 118)

Abstract

This paper focuses on weather derivatives as efficient risk management instruments and proposes a more advanced approach for their pricing. An "hybrid" contract is introduced, combining insurance properties, specifically tailored for the region under study and introducing Value-at-Risk (VaR) and Expected Shortfall (ES) as appropriate measures for the strike price. The numerical results show that VaR and ES are both efficient ways for managing the so-called Tail Risk; further, being ES more conservative than VaR and due to its subadditivity property, it can be seen that seasonal contracts are generally better off than monthly contracts in reducing global risk.

Keywords: climate change, temperature, risk hedging, Value-at-Risk, Expected Shortfall, portfolio diversification

4.1 Introduction

In the last years, growing concern about climate changes and risk related to meteorological events has strongly entered into the agenda of governments and companies.

Companies whose profits rely heavily on certain weather conditions, for instance, not too hot or too cold temperatures throughout the year, could seek protection by diversifying their business by undertaking activities that are not sensitive to weather itself. Obviously, this strategy may not completely offset losses, due to adverse weather. In addition to this, opening a new business is very costly and might not be viable (Brix et al. (2005)). However, more importantly is that these dangers may not necessarily be obvious or of catastrophic nature. For instance, Kelly et al. (2005) demonstrate that global warming may gradually alter a firm's exposure to shocks. It is evident that, to adjust to changes, firms must bear additional costs.

Thus, many agents in the global economy are attempting to protect themselves from such risks. A vast variety of commodities can be heavily affected by, for instance, rises or drops in temperatures (see, e.g., Tang & Jang (2011) and Tang & Jang (2012)); such fluctuations may lead to substantial financial losses.

According to Zenghelis (2006), the dangers of climate change are reducible to financial terms and, thus, can be studied from an economic standpoint. Moreover, it is recognized that the magnitude of the impact may vary spatially, so a financial intervention is needed at the local level, by tailoring the hedging instruments.

Needless to say, both expected and unexpected fluctuations in temperature may lead to losses. Hence, the interest in hedging instruments that can offset these losses is high.

Thus, the aim of this paper will be placed upon examining weather derivatives as a capable risk management tool and upon developing a more advanced approach to weather derivatives pricing. Moreover, a new instrument, called "hybrid" is proposed, its aim being to serve as an efficient hedge against climate alterations. Naturally, a question may arise at this point: why does the subject of this study constitute a powerful alternative to conventional insurance tools?

Insurance tools have proven their efficiency in the case of a low-probability, high-impact event. Hence, such contracts tend to exhibit a superior performance when consequences of a serious disaster are to be mitigated. However, high-probability and low-impact risks may be better tackled with non-traditional hedging instruments Mills (2005) or Le Den et al. (2017). In Alaton et al. (2002), the authors mention the following reasons to illustrate the advantages of weather derivatives relative to insurance policies. First of all, in order to trigger an insurance claim, an insured agent has to have a proof of a significant loss. On the contrary, payout of weather derivatives is independent of such proof. Secondly, weather derivatives constitute a more flexible instrument, and they can be tailored to a vast range of specific weather events. Thirdly, just like other financial securities, weather derivatives are tradable assets and they can be used for speculative or hedging purposes. These activities are the basis for a somehow liquid market; a key point for obtaining fair prices. In turn, pricing weather derivatives may be an issue of complexity. A variety of pricing methods exists, and the academic community has no unity toward a 'universal' valuation model. A brief review of the most fundamental works upon which this paper is based is provided in the next subsection.

Contributions in the Literature

An overview paper that was written by Taib & Benth (2012) compared a number of popular approaches for weather derivatives pricing. This study is dedicated to the burn-analysis, the weather index approach, and the temperature modeling. A contribution by Alaton et al. (2002) first derives a stochastic model for pricing temperature weather derivatives. These authors rely on the Wiener process in order to describe variance and analyze the noise and trend components separately. Moreover, they incorporate mean-reversion into their model. They then let the price of a derivative be equal to a discounted expected value under martingale measure. In doing so, the authors suppose that the market price of risk is constant.

In order to forecast temperature and price weather derivatives, Benth & Saltyte-Benth (2005) develop a continuous-time mean-reverting Levy-based Ornstein-Uhlenbeck model with stochastic volatility. These authors rely on classical arbitrage pricing theory in the sense of Black & Scholes (1973); Merton (1973). It is though to be noted that their model is extended by means of a parameter measuring the market price of risk, because the valuation of derivatives on temperature indices cannot be only based on an hedging principle. Benth & Saltyte-Benth (2011) propose a continuous-time auto-regressive model for the temperature dynamics. These researchers treat volatility as the product of a seasonal function and a stochastic process, being the Barndorff–Nielsen and Shephard model applied to represent stochastic volatility. Thus, the setting of this model allows for the authors to take advantage of the classical derivatives pricing theory.

A step further is done in Saltyte-Benth & Benth (2012). Here, the authors derive a continuous-time stochastic model for temperature forecasting. The contribution of their study is threefold. First, they show that estimating each component of the model separately can be beneficial. Second, they argue that the average temperature plays a fundamental role in temperature modeling. Third, they call for using a product between a seasonal deterministic function and a classical GARCH process in order to estimate seasonal volatility.

In Stefani et al. (2018) authors rely on the studies by Benth & Saltyte-Benth (2005, 2011); Saltyte-Benth & Benth (2012) to price weather derivatives. They then proceed to evaluate an hybrid instrument that encompasses properties of insurance contracts and propose Value-at-Risk (denoted with VaR in what follows), as a threshold. VaR is a very well known risk measure, which is widely applied in finance.

The present paper further extends the approach by Stefani et al. (2018), as recalled above. The key novelty of this paper is the application of the Expected Shortfall (ES in what follows) instead of VaR. As a risk measure, Expected Shortfall (ES) follows what the Basel Committee on Banking Supervision suggests in order to efficiently measure risks (see, e.g., Barger & Adkins (2013)). The main features of ES are documented at length in papers as, e.g., Artzner et al. (1999), Acerbi & Tasche (2002),Acerbi & Szekely (2014). As opposed to VaR, ES turns out to be very useful when a financial market is 'under stress', as, in general terms, ES is coherent (in brief, a risk measure is coherent when it allows the risk manager to set aside an appropriate amount of money to cover and hedge a full range of risks) and conservative. More importantly, VaR and ES are both computed exploiting a given quantile-level q that seems a very appropriate way to deal with the so called tail risk. As shown in this paper, ES also allows to determine an appropriate strike price for Weather Derivatives. Moreover, as it is confirmed by the application, ES turns out showing its strength due to the subadditivity property, that is particularly useful when considering an entire set of derivatives contracts under a 'portfolio management perspective'. An insurance company willing to issue Weather Derivatives might benefit of the diversification effect that hedges its portfolio of such contracts leading to a reduction in its overall risk exposition. From the side of a producer, a farmer may be better off buying a seasonal contact rather than a bunch of monthly contracts. Further, the pricing of the proposed hybrid instruments includes risk loadings, which is the reward that an insurer requires to enter the contract.

Finally, generally speaking and, as an extension to the scope of this paper, note that Weather Derivatives can provide protection against various manifestations of climate change: abnormal temperatures' behavior, substantial changes in the quantity of precipitation per month, and the like. Therefore, the payout of a single contract may be prompted by multiple underlying variables. Pricing this contract can be incredibly complex, because, despite being improbable, several adverse weather conditions can occur simultaneously. ES can assess consequences of such improbable events more accurately when compared to VaR.

The present paper has the following structure. Section 4.2 recalls the main points in weather derivatives pricing and risk measure theory; Section 4.3 presents the data and performs an application of risk measures through historical and parametric methods. An illustration of the stochastic model that was used to capture the behavior of temperatures is provided in Section 4.4. Section 4.5 deals with the application of a worst case approach that is based on VaR and ES while Section 4.6 shows the valuation of hybrid weather derivatives. Section 4.7 is devoted to analyzing the effectiveness of ES-based approach for hedging meteorological risk. Section 4.8 concludes.

4.2 Weather Derivatives and Risk Measures

A Weather Derivative (WD) is a financial contract whose pay-off depends on the behavior of some meteorological underlying variable, such as, for instance, wind and rainfall.

This paper focuses on a temperature-based WD: the daily arithmetic mean between the maximum and minimum observed temperatures in a specific area, denoted with $T(i)$, $i = 1, 2, \dots,$ will be the relevant underlying variable.

The three most common contingent claims dealing with temperature are:

- Heating degree days (HDD) contracts defined, over some time interval $[t_1; t_2]$, as

$$\text{HDD} = \sum_{i=t_1}^{t_2} \max(K - T(i), 0) \tag{4.1}$$

  where threshold K is usually equal to 18 °C (65 °F) with respect to the Chicago Mercantile Exchange (CME) market (Alexandridis & Zapranis (2013)) while the European Environment Agency[1] (EEA) states that, in Europe, such a baseline temperature is K = 15.5 °C.

- Cooling degree days (CDD) contracts defined as

$$\text{CDD} = \sum_{i=t_1}^{t_2} \max(T(i) - K, 0), \tag{4.2}$$

  and act the opposite way with respect to HDD. According to EEA, K = 22 °C.

- Cumulative average temperature (CAT) is defined as

$$\text{CAT} = \sum_{i=t_1}^{t_2} T(i).$$

  that measures the sum of average temperature over the period $[t_1; t_2]$.

[1] https://www.eea.europa.eu/data-and-maps/indicators/heating-degree-days-2.

In order to determine a cash-flow, these quantities are multiplied by a amount λ called ‘tick size’.

Pay-offs generated by HDD and CDD contracts are non-symmetric; HDDs and CDDs produce cash-flows whenever temperatures are below, in the first case, or above, in the second, some threshold K. In derivative pricing terminology, this value is commonly known as strike price.

It is easy to see that cash-flows that are generated by HDDs and CDDs are, indeed, the sum of pay-offs generated by European call/put options. In the rest of the paper, due to their resemblance to European options, only HDDs and CDDs will be considered.

As said in the Introduction, the key point motivating this contribution is the flexibility in choosing strike prices. A K = 18 °C threshold could not be convenient for a number of cases. For instance, small farmers who operate in tiny areas (for instance wineries producing high quality wine in limited quantities) might not be able to access regulated WD markets or might find such derivatives too expensive or ineffective. On top of this, the reference temperatures for standardized WD cover a number of US and European towns whose climate might have little or no connection at all with the area toward which an agent seeks protection for.

This might lead to agents looking for WD that will not have a sufficiently liquid market to be traded into. Unfortunately, in order to have model capable of determining fair prices, a derivative market should be as liquid as possible, with the largest possible number of buyers and sellers acting on it at the same time.

This is the reason why this paper extends the above mentioned methodology introduced by Stefani et al. (2018) here, strike prices K are chosen so that WD, for whom a formula capable of evaluating them exists, produce cash-flows only when some relevant and well-suited events occur.

In order to opt for a "as-correct-as-possible" choice for K, a proper way to deal with this issue is to use quantiles. It is well known (Artzner et al. (1999)) that such quantities are very important in the theory of risk measures.

For the sake of compactness, a very brief and partial recall, with definitions of VaR and ES, of the theory proposed by Artzner et al. (1999) is presented in Appendix 4.8. Here, it suffices to recall that, even for both risk measures, closed-form expressions exist, VaR is not, unlike ES, a coherent risk measure. This difference is crucial: ES manages risks exploiting, if possible, some diversification effect. This turns out to be a fundamental feature for risk management.

Letting a WD strike price be equal to some quantile-derived level of, in this case, observed temperatures allows for pinpointing in a financially sound way which negative event the derivative is supposed to hedge.

4.3 Data Collection

The data-set used in this paper is composed of historical daily data for the temperature in Celsius (°C) from the weather station Molin Bianco in Arezzo, Tuscany spanning 47 years, from January 1970 to 2017 and available at Settore Idrologico Regionale (SIR)[2] web-site.

[2] https://www.sir.toscana.it.

Due to some discontinuities in data, in order to avoid biases or discrepancies, the numerical analysis performed in this article relates to the 1970–2017 period. Further, the pricing model is going to be calibrated using data from 1970 to 2016 while data of 2017 will be used to validate WD prices.

Traditionally, many areas in Tuscany are devoted to the production of high-quality agricultural produce. This is also the case of the province of Arezzo, so that, not only from a theoretical point of view, it is clear that a way to hedge against climatic risk is somehow required.

Daily average temperatures (DAT) have been calculated while using a simple arithmetic mean of max and min temperatures, as reported in the SIR web-site. Following Wang et al. (2015), data have been corrected for leap-years. The full sample consists of $16\,790$ daily observations. Table 4.1 presents some descriptive statistics for daily average temperatures.

Table 4.1: Descriptive Statistics of the daily average temperature in Arezzo (Tuscany) 1970–2016.

	Mean	**Var**	**Std Dev**	**Min**	**Max**	**Skewness**	**Kurtosis**
Temp. (°C)	14.26	51.34	7.17	-8.1	34	0.028	2.01

Financial literature carries a number of ways to determine the quantiles and related values out of a set of observed data. The first is simply to determine these quantities directly from observed data. In doing so, though, the methodology of a finite-sample approach for the estimation of VaR and ES described in Rockafellar and Uryasev (2002) is to be applied. The method that was implemented by these authors makes a correction, so that the tail probability is always consistent with the VaR level, a refinement that might be needed in the numerical analysis performed in this article.

Farmers are usually worried about the weather's behavior during some specific periods, particularly during winter or summer months. To address this point, a data-set from 1970 to 2016 has been divided into monthly subsets. For each of them, some quantile-related thresholds have been computed (see Figure 4.1).

Thresholds K for VaR (lhs-part of Figure 4.1) and ES (rhs-part of Figure 4.1) are depicted in Figure 4.1 (for the definition of VaR and ES see Appendix 4.8). VaR-related threshold (K_α^{VaR}) is the $\alpha-$percentile, $\alpha \in [0,1]$, of each subset. By definition, the ES-related threshold (K_α^{ES}) is the average of values of the subset that are identified by K_α^{VaR}.

In Figure 4.1, the relationship between strike prices, expressed in terms of temperature, (K_α^{VaR} and K_α^{ES}) (vertical axis) and percentiles, which vary from 0% to 100%, (horizontal axis) is depicted.

Strike values that correspond to percentiles from 0% to 50% are HDDs observed thresholds, while those corresponding to percentiles from 50% to 100% are CDDs observed thresholds.

As an illustration, $K_\alpha^{\text{VaR}} = 16.5$ °C and $K_\alpha^{\text{ES}} = 15.24$ °C are the thresholds for June when considering the 10% percentile. Therefore, these values identify thresholds applied in case of trading a HDD contract. On the other hand, $K_\alpha^{\text{VaR}} = 24.39$ °C and

$K_\alpha^{\text{ES}} = 19.97$ °C are similar threshold values for a CDD contract for June considering the 90% percentile.

From Figure 4.1, it can be observed that VaR/ES thresholds that are found in terms of quantiles are far away from the standard strike price suggested by CME (18 °C), but closer to the ones stated by EEA (15.5 °C). As an example, the value for VaR for June if temperature is 18° is approximately $\alpha = 0.2$.

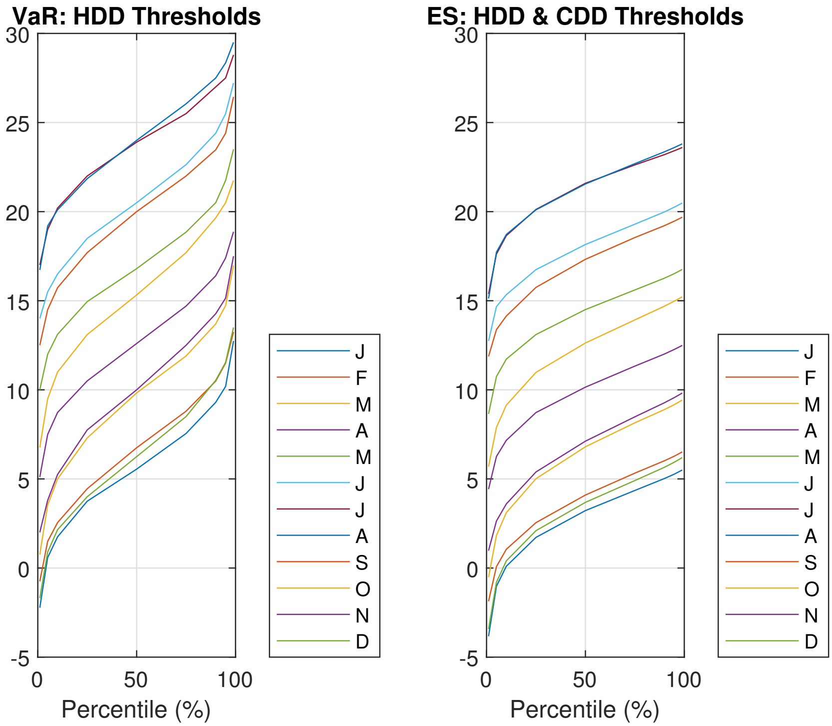


Figure 4.1: Thresholds K_α^{VaR} and K_α^{ES} for heating degree days (HDD) and cooling degree days (CDD) for different confidence levels and in different months (HDD corresponds to percentiles 0% to 50% while CDD corresponds to percentiles 50% to 100%). Note that the abbreviations J_1, A_1, J_2, J_3, and A_2 are used to identify the months of January, April, June, July, and August, respectively.

Estimation of VaR and ES Using Historical and Parametric Approaches

In this paper, a range of methods for estimating VaR and ES have been exploited (see, for instance, McNeil et al. (2005): these are the historical (see, e.g., Rockafellar and Uryasev (2002)), the normal, the "T5" and "T10" ones. While the historical method has no a priori assumptions, the normal, T5, and T10 are assumed parametric because they rely on some underlying distribution: the normal distribution for the normal method and a Student's-t distribution with 5 and 10 degrees of freedom for the remaining two.

However, while the normal and Student's-t distributions quite efficiently fit a number of financial time series, this is not the case for most meteorological data (see Figure 4.2).

As it is well known, historically speaking the first theoretical distribution applied in mathematical finance is the normal one. Empirical investigation has shown that real financial data are characterized almost ubiquitously by extreme events; in fact, large losses occur with a frequency that the normal distribution is unable to predict.

VaR and ES are risk measures developed to hedge such potential losses. Their numerical results depend, of course, on the assumed underlying distribution; for a number of distributions (for instance, Student's-t) explicit formulae exist.

When dealing with temperatures, at least in the area scrutinized in this article, it occurs that historical distribution is bimodal and with thin tails. This leads to the fact that, in this case, VaR and ES become very conservative and prudent ways of representing risk.

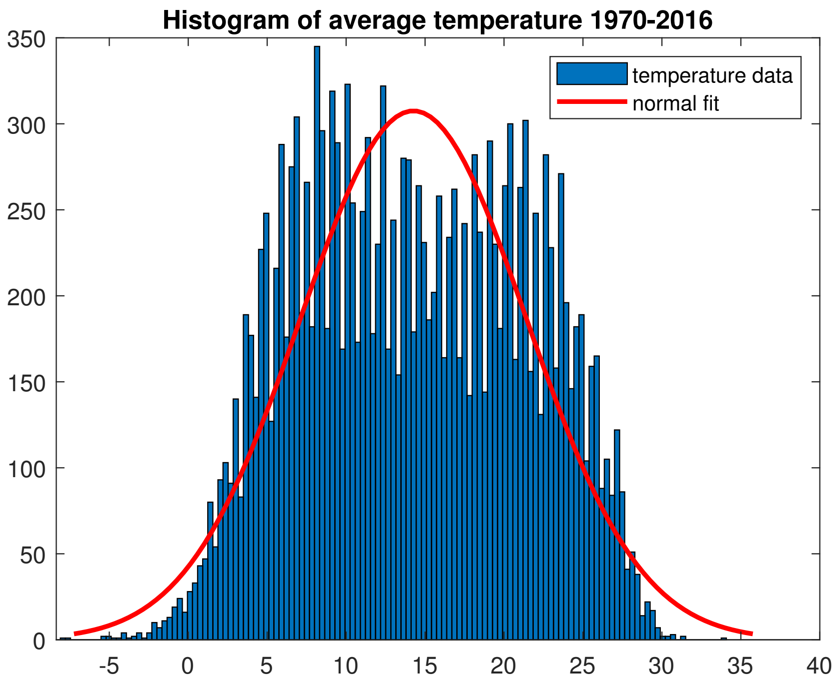


Figure 4.2: Historical distribution for daily temperature, Arezzo, 1970–2016.

The following analysis shows that meteorological data are usually characterized by 'thin', rather than 'fat', tails.

To this end, an analysis of the tails of the distribution of observed temperatures has been attempted. As it is known, the normal distribution is thin tailed, which is, its upper tail declines to zero faster than exponentially Pindyck (2011) and that the exponential distribution, whose probability density function is

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if} \quad x \geq 0 \\ 0 & \text{if} \quad x < 0 \end{cases},$$

is thin-tailed.

The right tail of such distribution is assumed to be composed of all temperatures that are larger or equal to the 90%th percentile (24 °C) of the temperature data-set.

In Figure 4.3, in order to apply the exponential distribution, the 90% percentile has been subtracted to all temperatures. The estimated λ is equal to 1.85997 with 95% confidence interval $[1.77605, 1.95002]$.

Similarly, the left tail of the distribution under scrutiny is assumed to be composed of all temperatures that are smaller or equal to the 10% percentile (5 °C) of the temperature data-set.

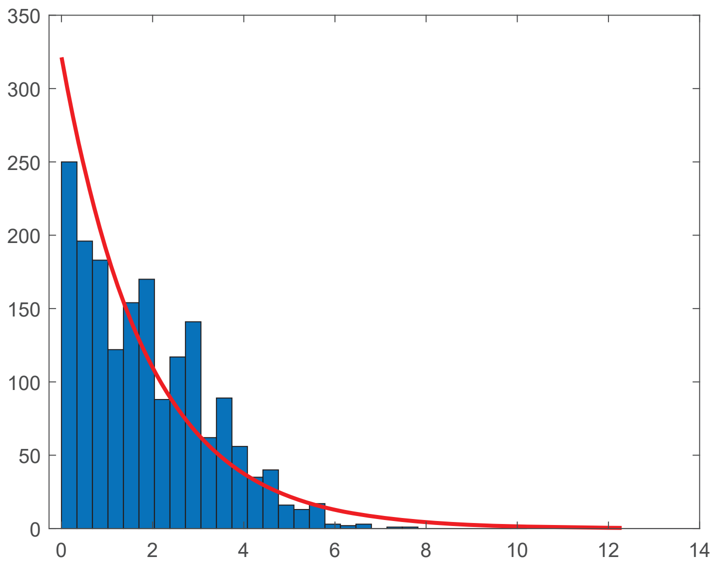


Figure 4.3: Exponential fit of temperature distribution—tail beyond 90% percentile.

In Figure 4.4, in order to apply again the exponential distribution, temperatures have been firstly changed in sign (obtaining a right tail). Further, in order to deal with positive numbers, the 10% percentile has been added to all temperatures. The estimated λ is here equal to 2.02396 with 95% confidence interval $[1.93481, 2.11945]$.

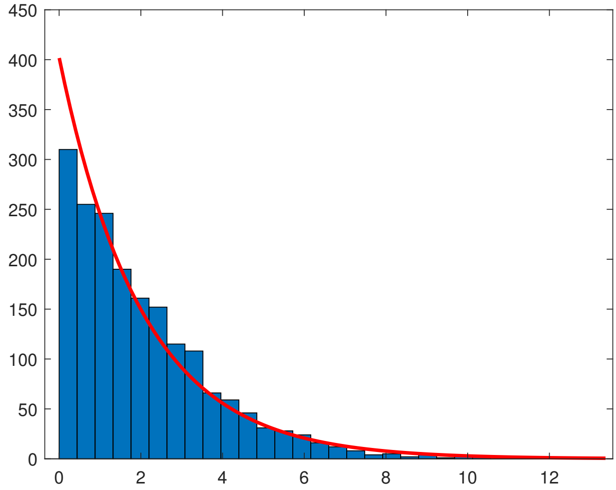


Figure 4.4: Exponential fit of temperature distribution—tail below 10% percentile.

In conclusion, the temperature distribution shows thin tails. Therefore, the econometric analysis on VaR and ES must keep proper account of this feature. This is achieved by considering the historical and normal methods.

The first method used is the historical one (see Figure 4.5). The estimation rolling window size is set equal to 365 days, so that a full year of data is used to estimate for both the historical VaR and ES; the sample window runs from the beginning of 2003 through the end of 2016. Following [**basel2**], a VaR confidence level of 97.5% is used.

According to Rockafellar & Uryasev (2002), the historical VaR and ES are computed, as follows:

$$VaR_\alpha = z(k)$$

and

$$ES_\alpha = \begin{cases} \dfrac{(k - \alpha \cdot N)z(k) + \sum_{i=k+1}^{N} z(i)}{N(1-\alpha)} & \text{if } k < N \\ z(k) & \text{if } k \geq N \end{cases}$$

where $k = \text{floor}\,(\alpha \cdot N)$, N is the length of the historical data under study, while z is the vector containing all sorted historical data.

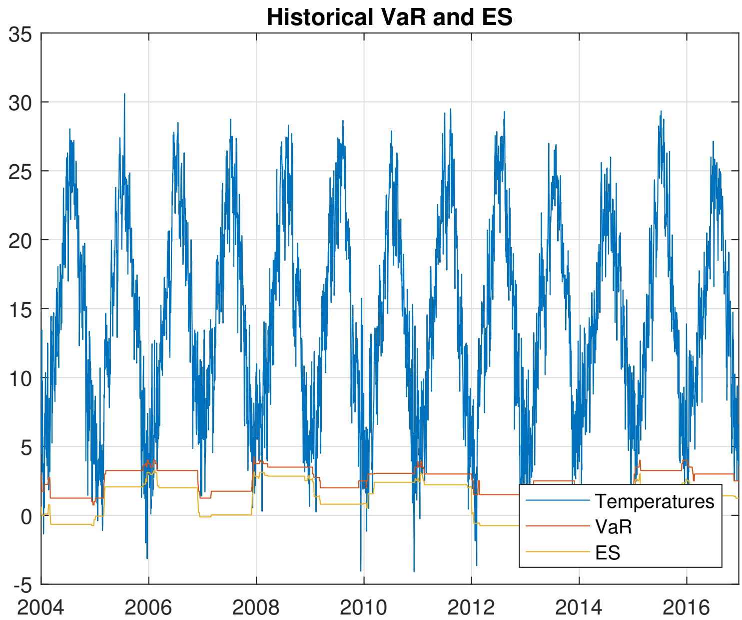


Figure 4.5: The daily average temperature, Value-at-Risk (VaR), and Expected Shortfall (ES) estimated with the historical method.

Figure 4.5 shows that the historical simulation curve has a piece-wise constant profile. The reason for this is that quantiles do not change for several days until extreme events occur. Thus, the historical simulation method is slow to react to changes in volatility.

Another estimation method uses parametric models (McNeil et al. (2005)); this approach requires computing the volatility of daily average temperatures.

Given this volatility, VaR and ES can be analytically computed assuming that temperatures follow a normal distribution with mean μ and variance σ^2:

$$\mathrm{VaR}_\alpha = \mu + \sigma\Phi^{-1}(\alpha)$$

and

$$\mathrm{ES}_\alpha = \mu + \sigma\frac{\phi(\Phi^{-1}(\alpha))}{1-\alpha}$$

where ϕ and Φ denote density of standard normal distribution and cumulative standard normal distribution, respectively. $\Phi^{-1}(\alpha)$ is the α-quantile of Φ. A non-zero mean is assumed and it is estimated as a sum of all yearly means. For the normal distribution, the estimated volatility is directly used to obtain the VaR and ES (see Figure 4.6).

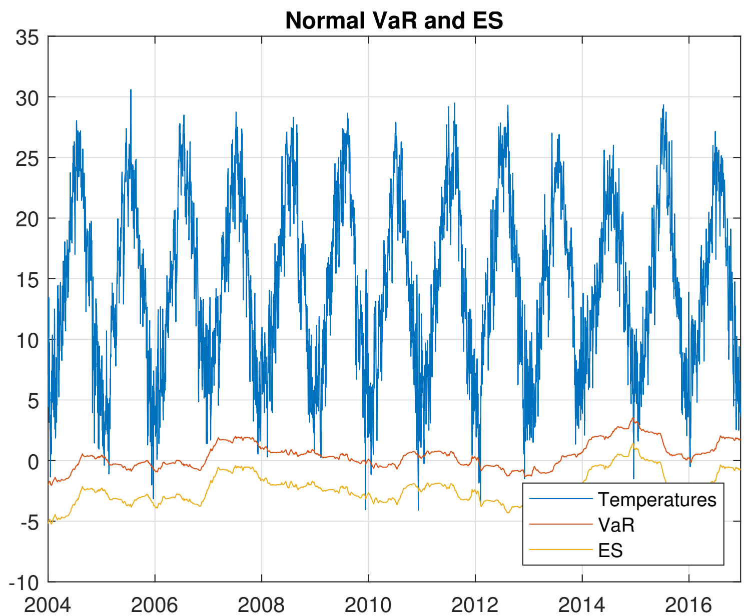


Figure 4.6: The daily average temperature, VaR, and ES estimated with the normal method.

The data-set under scrutiny has a kurtosis far away from 3, as can be seen from Table 4.1. To properly test this claim, Table 4.2 carries the results for two tests for normality of data:

Table 4.2: Normality tests of temperature data in Arezzo (Tuscany), 1970–2016.

	Jarque-Bera	**Lillefors (Kolmogorov-Smirnov)**
Statistic value	693.99	0.0564
p-value	0.001	0.001

The null hypothesis (data are normally distributed) is rejected by both of them.

To tackle this issue, a more flexible theoretical distribution can be applied. The key point here is to choose a distribution that could somehow be compatible with observed data. A plausible choice is the Student's-t one. In fact, denoting with ν the number of its degrees of freedom, its skewness is 0 for $\nu > 3$, while its kurtosis is $6/\left(\nu - 4\right)$ for $\nu > 4$, a number compatible with the kurtosis observed (see Table 4.1).

VaR and ES Formulas for a Student's-t with ν degrees of freedom $\left(T \sim t(\nu, \mu, \sigma^2)\right)$ distribution are

$$VaR_\alpha = \mu + \sigma t_\nu^{-1}(\alpha)$$

and

$$ES_\alpha = \mu + \sigma \left(\frac{g_\nu(t_\nu^{-1}(\alpha))}{1-\alpha} \cdot \frac{(\nu + (t_\nu^{-1}(\alpha))^2)}{\nu - 1} \right)$$

where t_ν is the standard t-distribution c.d.f, t_ν^{-1} is the standard t-distribution quantile and g_ν is the standard t-distribution p.d.f.

Figure 4.7 shows the daily temperatures, VaR and ES estimated values with Student's-t distribution with 10 and 5 degrees of freedom.

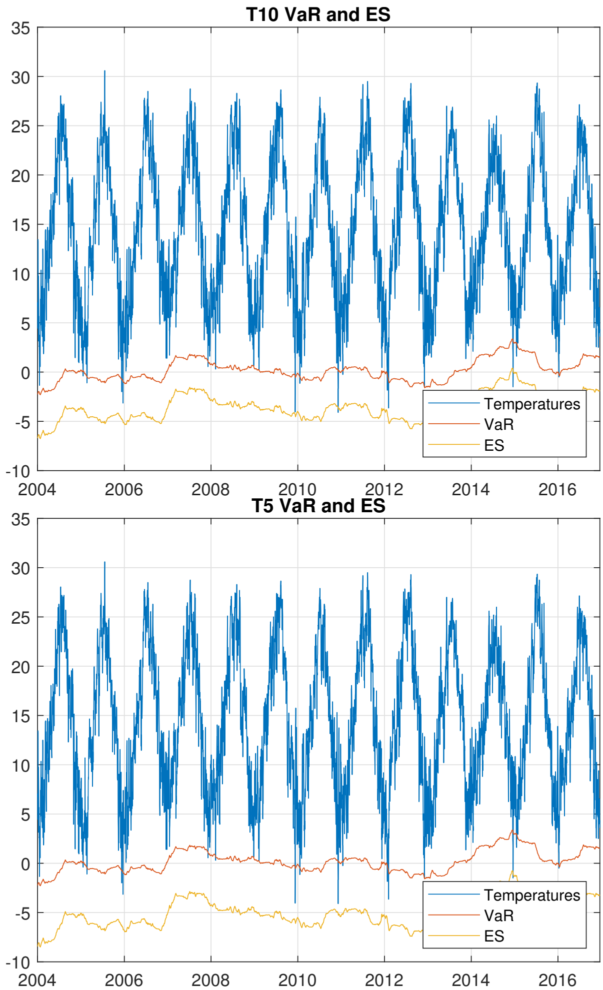


Figure 4.7: Daily average temperature, VaR and ES estimated with the T10 and T5 methods.

No theoretical model can fully catch reality, as is almost always the case. Still, financial literature has deeply analyzed the results obtained by theoretical models applied to real data. On top of this, closed formulas for VaR and ES, for the distributions used in this section, exist. Looking at data under scrutiny, it results that temperatures, at least in this case, have thin tails. Financial agents are, instead, worried about fat fails. This explains why the normal distribution performs better, when compared to the Student's-t distribution, with the latter being characterized by fatter tails. This observation leads to the idea, to be developed in a subsequent paper, to apply theoretical distributions with thinner tails when compared to the normal distribution ones.

A back-test on ES, which is based on available data, is in Appendix 4.8. Its main result is that the Historical and Normal models perform better and, therefore, will be used in what follows.

Of course, the fact that Arezzo's temperatures show thin tails could be the result of the fact that Tuscany is located in a very favorable area, meteorologically speaking.

As the aim of the paper is to present a novel approach in pricing weather derivatives by means of VaR and ES, its findings has to be standard and consistent. This leads to the claim that what has been presented here could be possibly used everywhere, regardless of the behavior of the observed meteorological data.

The historical method, of course, does not assume any underlying distribution. On the other hand, as said, the normal method, assuming tails fatter than the observed ones, will give results that are more conservative than taking a better fitting distribution. The claim is that the normal method is the best financial compromise in choosing an underlying distribution, according to the universal standards, without leaving the freedom of applying different methods for finding VaR and ES. In any case, the final results in pricing and hedging are found to be very similar using the historical or the normal method.

4.4 Temperature-Based Model

To model the temperature dynamics, a time series decomposition approach proposed by Benth & Saltyte-Benth (2013). Temperature $T(t)$ is decomposed into a mean component $\mu(t)$, which models the trend, and a residual component $\epsilon(t)$ that models the fluctuations around a trend over time. The temporal decomposition of $T(t)$ can be written as:

$$T(t) = \mu(t) + \epsilon(t) \tag{4.3}$$

where $\mu(t)$ is a deterministic function over time defined as:

$$\mu(t) = \theta(t) + \sum_{i=1}^{p} \alpha_i (T(t-i) - \theta(t-i)) \tag{4.4}$$

where α_i are the parameters of an AR(p) process, $\theta(t)$ is the (linear) trend, and seasonality in terms of the sine and cosine functions in $T(t)$ expressed is as follows:

$$\theta(t) = a_0 + a_1 t + a_2 \cos\left(\frac{2\pi t}{365}\right) + a_3 \sin\left(\frac{2\pi t}{365}\right).$$

The part containing trigonometric functions explains the seasonal variations in temperature. The variable $\epsilon(t)$ represents residuals at time t in the form:

$$\epsilon(t) = \sigma(t)\gamma(t) \tag{4.5}$$

where $\sigma(t)$ is a (possibly time–dependent) volatility function, $\gamma(t)$ is a zero-mean temporally independent Gaussian random process with standard deviation that is equal to one, and $\sigma(t)$ is a (possibly time–dependent) volatility function given by a truncated Fourier series as follows:

$$\sigma^2(t) = b_0 + \sum_{l=1}^{4} b_{1l} \cos\left(\frac{2\pi t}{365}\right) + b_{1l+1} \sin\left(\frac{2\pi t}{365}\right).$$

In order to predict average temperature for 2017, the dynamics of $\epsilon(t)$ is now to be found. To achieve this, parameters for the volatility function $\sigma(t)$, expressed using a Fourier series of order 2 and applying the standard OLS method, are determined.

The next step is to simulate 1000 trajectories for $\gamma(t)$; this leads to obtaining 1000 trajectories for average temperatures. The lhs of Figure 5.17 shows such trajectories and how the observed average temperatures (red curve) differ from the simulated average temperatures (blue curve). In the rhs of the same Figure 5.17, forecast errors for 1000 simulated trajectories with respect to average temperatures of 2017 are displayed.

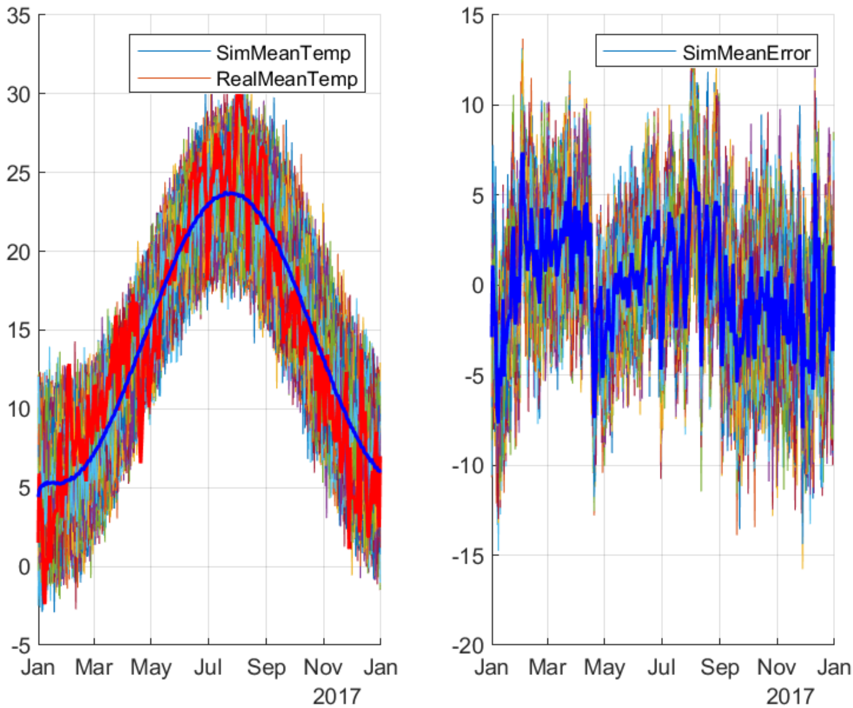


Figure 4.8: Simulated mean temperature and errors for each trajectory for year 2017.

For a more detailed analysis of errors due to simulation,the mean absolute percentage error (MAPE), mean absolute error (MAE), and root mean square error (RMSE) discrepancies of simulated data from the observed average 2017 temperatures have been computed and represented in Figure 4.9.

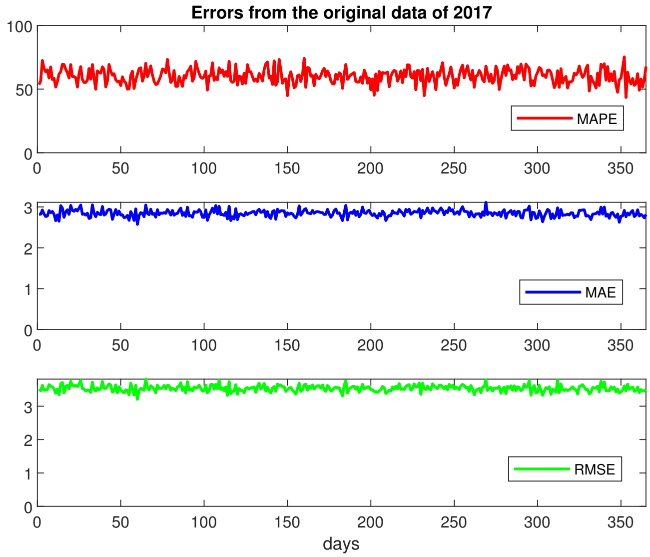


Figure 4.9: Daily errors of simulated data when compared to observed 2017 temperatures (mean absolute percentage error (MAPE), mean absolute error (MAE), and root mean square error (RMSE)).

These data represent the errors that result when simulated data are compared with the data observed in 2017. In order to at least partially reduce errors displayed above, a 'removing outliers' approach will be performed in the next section.

4.5 Worst Case Approach Based on VaR and ES

In order to reduce uncertainty when calculating WD prices, the method proposed by Benth, Kutrolli & Stefani (2019) is now applied.

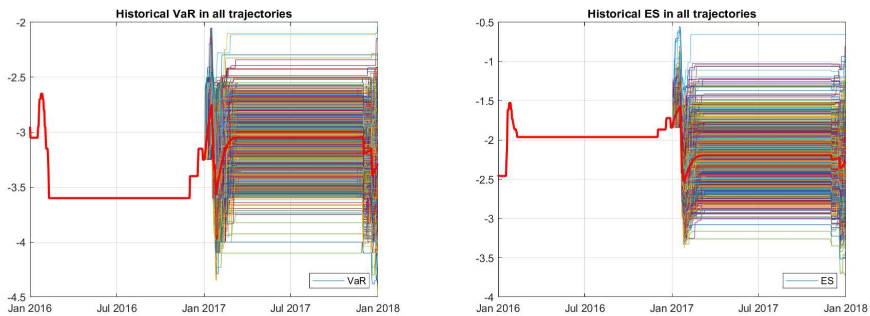


Figure 4.10: Simulated VaR and ES for 2017 obtained by means of 1000 trajectories using the historical method.

This is done inside a Monte Carlo framework, in which values for the stochastic component of the temperature model given in (4.5) are simulated.

At first, VaR and ES for all trajectories of volatility for year 2017 are computed. Figure 4.10 shows the VaR and ES obtained by applying the historical method (see Section 4.3).

Figure 4.11 shows VaR and ES obtained by applying the normal method; red curves identify the average of VaR and ES from all of the simulated trajectories.

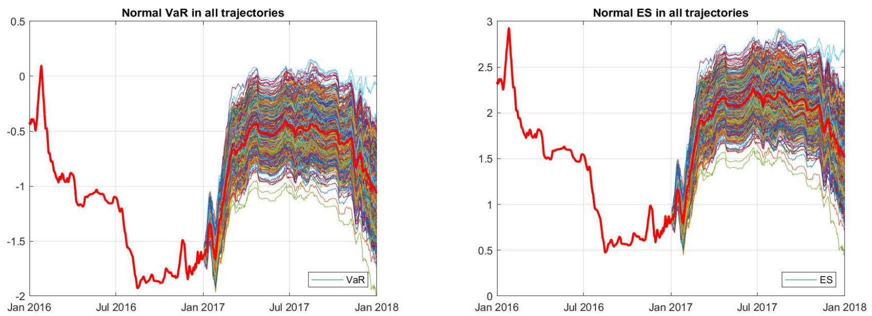


Figure 4.11: Simulated VaR and ES for 2017 applying in 1000 different trajectories using the normal method.

This being done, applying the worst case approach that is based on VaR and ES removes outliers/worst scenarios that, if not expunged, affect prices of weather derivatives and lead to an increased uncertainty. The elimination of extreme tails approach is based on a $\eta = 2.5\%$ level.

Figure 4.12 shows VaR and ES after having removed extreme tails by applying the historical method; red curves identify the averages of VaR and ES from the remaining simulated trajectories.

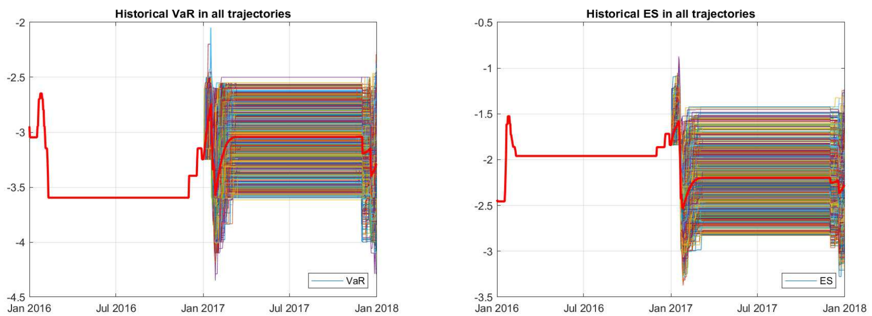


Figure 4.12: Simulated VaR and ES for 2017 upon removing extreme tails using historical method.

Figure 4.13 shows VaR and ES after having removed extreme tails by applying the normal method; red curves identify the averages of VaR and ES from the remaining simulated trajectories.

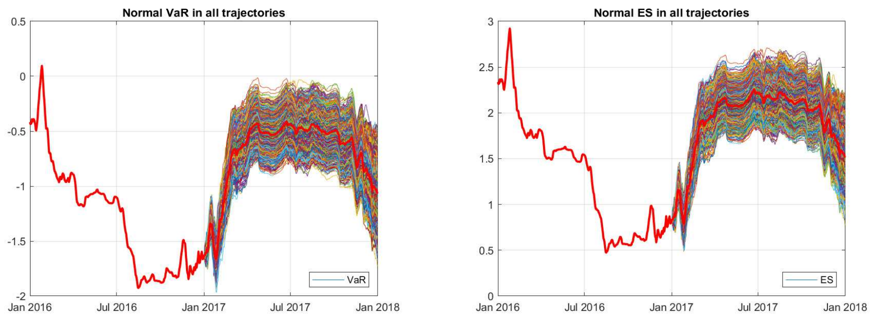


Figure 4.13: The simulated VaR and ES for 2017 after removing extreme tails using normal method.

By comparing Figures 4.12 and 4.13 with Figures 4.10 and 4.11, having removed outliers, the trajectories show a more compact behavior.

In the next section, pricing of pure financial derivatives has been computed removing (Table 4.3) or not removing (Table 4.4) extreme tails.

Table 4.3: Pure financial derivative values for heating degree days (HDD) and CDD computed from data simulated for 2017—thresholds 5% for HDD and 95% for CDD—tick size $\lambda = 20$ EUR—tail values have not been removed.

Month	**HDD (VaR/ES)**	**CDD (VaR/ES)**
January	$48.92/19.15$ EUR	$55.09/846.03$ EUR
February	$35.49/11.84$ EUR	$116.39/805.06$ EUR
March	$50.65/17.18$ EUR	$43.59/958.18$ EUR
April	$36.43/12.71$ EUR	$83.09/845.55$ EUR
May	$33.81/11.39$ EUR	$46.08/849.19$ EUR
June	$25.90/11.66$ EUR	$73.05/833.31$ EUR
July	$35.28/10.16$ EUR	$34.13/741.01$ EUR
August	$38.17/12.88$ EUR	$30.76/816.59$ EUR
September	$34.63/11.83$ EUR	$69.72/834.55$ EUR
October	$45.56/14.59$ EUR	$28.51/938.44$ EUR
November	$33.52/10.98$ EUR	$79.63/954.93$ EUR
December	$51.28/17.91$ EUR	$38.86/924.04$ EUR

Table 4.4: Pure financial derivative values for HDD and CDD computed from data simulated for 2017—thresholds 5% for HDD and 95% for CDD—tick size $\lambda = 20$ EUR—tail values have been removed.

Month	HDD (VaR/ES)	CDD (VaR/ES)
January	48.90/19.14 EUR	55.08/846.02 EUR
February	35.48/11.83 EUR	116.38/805.05 EUR
March	50.64/17.17 EUR	43.59/958.18 EUR
April	36.42/12.70 EUR	83.07/845.54 EUR
May	33.80/11.38 EUR	46.09/849.18 EUR
June	25.90/11.66 EUR	73.04/833.30 EUR
July	35.27/10.15 EUR	34.12/741.00 EUR
August	38.16/12.87 EUR	30.75/816.58 EUR
September	34.62/11.82 EUR	69.70/834.53 EUR
October	45.55/14.58 EUR	28.50/938.43 EUR
November	33.51/10.97 EUR	79.62/954.92 EUR
December	51.27/17.90 EUR	38.85/924.03 EUR

4.6 Pricing a Temperature-Based Weather Derivative

In this section, the pricing of HDD and CDD (both without and with risk loadings), computed for each month in 2017, are displayed. The temperatures for 2017 have been simulated as said before, while VaR/ES percentiles thresholds have been set equal to 5% for HDD and 95% for CDD.

Table 4.3 encompasses the ‘pure’ financial values of HDD and CDD computed according to formulae (4.1) and (4.2) and including tail values discussed above. Table 4.4 displays the same computations having, instead, removed tail values. The cut-off for considering a value an outlier has been set as equal to 2.5%.

A comparison between Tables 4.3 and 4.4 shows how the financial prices are not affected by the removal of extreme values. For this reason, and for sake of compactness, the rest of this numerical analysis is going to be performed when considering all values.

To depart from standard actuarial techniques, risk loadings are determined here as a fraction α of the $100 \cdot (1 - \alpha)$ percentile of the pay-off distribution of the derivative. Such values are obtained considering the historical temperature time series from 1970 to 2016. As a first step, temperatures are subdivided month by month. Out of these, for each of the twelve sub-sets the $100 \cdot (1 - \alpha)$ percentile is determined. Finally, these values become the threshold used to compute the historical pay-off distributions of either HDD and CDD. The complete graphical behavior of these thresholds is depicted above in Figure 4.1.

By looking at Tables 4.3 and 4.5, it is evident that, as HDD thresholds are lower when ES is applied instead of VaR, prices in the left part of column 2 are lower than those in the right part. A similar but opposite reasoning holds for CDD: here, thresholds are larger for ES rather than VaR so that prices in left part of column 3 are greater then corresponding prices in the right part.

Some descriptive statistics of option values treated in this section can be found in Appendix 4.8.

Table 4.5: Final prices (i.e., financial value + risk loadings) for HDD and CDD—simulations for 2017—tick size $\lambda = 20$ Euro—thresholds 5% for HDD and 95% for CDD—tail values have not been removed.

Months	HDD (VaR/ES)	CDD (VaR/ES)
January	61.09/24.09 EUR	65.08/926.09 EUR
February	41.95/16.70 EUR	170.62/889.76 EUR
March	67.43/25.57 EUR	57.31/1084.86 EUR
April	46.99/17.57 EUR	96.4/928.79 EUR
May	41.03/13.98 EUR	52.64/935.06 EUR
June	33.39/15.85 EUR	84.54/920.02 EUR
July	44.59/13.76 EUR	40.28/822.81 EUR
August	47.43/18.22 EUR	38.48/922.24 EUR
September	44.32/14.27 EUR	79.82/929.98 EUR
October	56.18/18.39 EUR	36.13/1048.33 EUR
November	43.19/14.08 EUR	92.49/1029.81 EUR
December	63.59/22.69 EUR	49.41/1006.49 EUR

4.7 Hedging Strategies with Hybrid Instruments Based on Value-at-Risk and Expected Shortfall

Consider a farmer who needs to hedge against negative weather events in the Arezzo's area by means of some tailor-made derivative contract.

According to computation that was performed above, the temperature risk for each month in 2017 can be hedged by means of either an HDD or CDD with monthly maturity.

In Benth & Saltyte-Benth (2013), it is stated that Chicago Mercantile Exchange (CME) offers HDD contracts from November to June, while CDD and CAT contracts are offered from May to September. This leads to the conclusion that, when considering WD, some months (i.e., summer and winter seasons) are more relevant than others.

When comparing final prices of these derivatives (see Table 4.5) with their actual 2017 daily pay-off obtained by using observed data leads to Table 4.6, where only winter and summer data are reported.

Table 4.6: HDD and CDD Profit/Losses for 2017—Prices in Table 4.5 have been compared with actual 2017 pay-offs.

Months	HDD (VaR/ES)	CDD (VaR/ES)
January	85.51/20.92 EUR	−65.08/−744.92 EUR
February	−41.95/−16.70 EUR	−120.62/354.61 EUR
March	−67.43/−25.56 EUR	−12.71/408.03 EUR
June	−33.39/−15.85 EUR	−52.54/828.07 EUR
July	−44.59/−13.76 EUR	−21.28/188.90 EUR
August	−38.43/−18.22 EUR	147.52/816.34 EUR

Aside monthly contracts, in order to test if the ES's diversification property holds in this context, the analysis has been extended to seasonal contracts.

For this reason, two seasonal contracts, the first encompassing the winter period ranging from January to March 2017, the second the summer one (June to August 2017) have been priced with the same methodology applied for monthly WDs.

Table 4.7 summarizes the pure financial prices (column 2), the full prices (i.e., adding risk loadings—column 3) and profit and losses (column 4) for HDD in the winter 2017 season.

Table 4.7: Comparison between HDDs pure financial prices, financial prices + risk loadings (RL) and Profit and Loss (P/L) − winter season 2017.

Months	**Pure Fin. Prices (VaR/ES)**	**Fin. Prices + RL (VaR/ES)**	**P/L (VaR/ES)**
Winter season	138.12/48.92 EUR	162.64/58.4 EUR	111.36/28.61 EUR
Jan-Feb-Mar	135.06/48.17 EUR	170.47/66.36 EUR	−23.87/−21.34 EUR

By looking at Table 4.7, it is evident, in terms of P/L, that a seasonal HDD contract leads to a profit for the insured farmer. Such profit would have turned into a loss in the case that the farmer had bought monthly contracts.

A somehow less clear-cut result occurs for the summer season (see Table 4.8), where CDD contracts are used. Final column in Table 4.8 shows that P/L for contracts priced while using VaR as a threshold is larger when considering a three-month duration derivative instead of three monthly CDDs. The opposite results, still leading, though, to a profit, occurs when ES is used.

Table 4.8: Comparison between CDDs pure financial prices, financial prices + RL and P/L − summer season 2017.

Months	**Pure Fin. Prices (VaR/ES)**	**Fin. Prices + RL (VaR/ES)**	**P/L (VaR/ES)**
Sum. season	88.60/2638.4 EUR	107.97/2877.9 EUR	242.03/1754.7 EUR
Jun-Jul-Aug	137.94/2390.9 EUR	163.30/2665.1 EUR	73.70/1833.3 EUR

To end this section, there is no need to say that some variables chosen in the above analysis, such as, for instance, tick sizes and contract maturities, can be chosen, so to match a large number of needs an agent might require to manage weather-related business risks. Insurance companies can, on their side, perform some fine tuning with respect to some quantities that are involved in WD so to avoid facing losses on their side, as displayed in Tables 4.7 and 4.8.

Further, according at least to data that were exploited in this article, the Expected Shortfall shows to maintain, even if not fully, its diversification property. Nevertheless, it has been numerically assessed that seasonal WD derivatives lead to better profit and loss results when compared with monthly contracts written on the same period.

4.8 Conclusions

This article has presented a way to hedge temperature risk exploiting weather derivatives contracts. This has been achieved when considering 'tail events' and the standard financial approach to tackle them: namely, Value-at-Risk and Expected Shortfall.

Using as a starting point historical time series for a very specific area, an 'hybrid' contract, composed of a Weather Derivative whose price also encompasses a risk loading, has been evaluated. The aim of this choice is to combine the efficient pricing of a derivative with the gains that are required by an insurer to carry a specific risk.

Pay-offs have been determined by means of a stochastic simulation technique and compared with those that were obtained observing 2017 temperatures.

According to risk measure theory, proved that Expected Shortfall captures diversification while Value-at-Risk does not, the numerical results presented above show that it is more convenient to enter a single contract that covers more months rather than monthly contracts spanned on the same period.

Research performed in this paper can be extended in a number of directions. As meteorological data are easily accessible, a possible extension might encompass different sites, letting for more sources of randomness to be explored. Another path is evaluating more complex derivatives instead of standard options. Beside this, the fact that temperatures present thin tails is an interesting, as well as a non-standard financial issue, research topic. Finally, a full comparison between hybrid and tailored-made Weather Derivatives and the financial contracts periodically issued on some regulated market, such as the Chicago Mercantile Exchange, might further shed some more light on this specific field.

Author Contributions

All authors have contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Aknownledgments

The authors of this paper gratefully thank the staff at the Servizio Idrologico Regionale della Regione Toscana (SIR) for its excellent management of data source and the referees whose precious reports helped to substantially improve and ameliorate this paper.

Declaration of Interest

The authors declare no conflict of interest.

Appendix A

Two of the most common quantile-based risk measures are Value-at-Risk (VaR) and Expected Shortfall (ES).

In general terms, let $\mathbf{X} \in \Omega$ be a random variable depicting the future outcome in $T > 0$ of a risk. In the ADEH framework, a risk manager is assumed to be always capable of distinguishing between acceptable and non acceptable risks.

Let $\mathcal{A}$ ($\mathcal{A}^C = \Omega - \mathcal{A}$) be the set containing all acceptable (non acceptable) risks. In order for the theory to be consistent, these sets must obey some "common sense" assumptions.

If $\mathbf{Y} \in \mathcal{A}^C$, a risk measure is some function $\rho : \Omega \to \mathbb{R}$ defined as

$$\inf_{\beta} \left\{ \beta v : (\mathbf{Y} + \beta) \in \mathcal{A} \right\}$$

where β is a positive amount of money while v is the risk-less discount factor in $[0; T]$.

A risk measure is, then, the smallest amount of money β to be added to a non acceptable random variable $\mathbf{Y}$ that makes it, as some future epoch, $\mathbf{Y} + \beta$ an acceptable one.

Let

$$F_{\mathbf{X}}(x) = P[\mathbf{X} \leq x] \quad \forall x \in \mathbb{R}$$

be the distribution function for a monetary-measured risk $\mathbf{X}$ and consider α, $0 < \alpha < 1$, a probability level.

- the Value-at-Risk (VaR) for a risk $\mathbf{X}$ with $1 - \alpha$ confidence level is defined as

$$\text{VaR}_{\alpha}(\mathbf{X}) = -\inf_{z} \left\{ F_{\mathbf{X}}(z) > \alpha \right\},$$

  If $\mathbf{X} \in \mathcal{A}^C$ ($\mathbf{X} \in \mathcal{A}$), then $\text{VaR}_{\alpha}(\mathbf{X}) > 0$ $\left(\text{VaR}_{\alpha}(\mathbf{X}) < 0\right)$

- the Expected Shortfall (ES) for a risk $\mathbf{X}$ with $1 - \alpha$ confidence level is expressed as

$$\text{ES}_{\alpha}(\mathbf{X}) = -\mathbb{E}\left[\mathbf{X} | \mathbf{X} \leq -\text{VaR}_{\alpha}(\mathbf{X})\right]$$

Even if such tail-based risk measures look appropriate in handling WD features, ADEH's work determines a fundamental result in risk management: some measures are unable to capture fully capture risks.

VaR, for instance is not a sub-additive measure. Let $\mathbf{X}, \mathbf{Y} \in \Omega$ be two risks.

It can be shown that

$$\text{VaR}_{\alpha}(\mathbf{X}) + \text{VaR}_{\alpha}(\mathbf{Y}) \leq \text{VaR}_{\alpha}(\mathbf{X} + \mathbf{Y}) \tag{4.6}$$

a feature that defies the notion of diversification: when analyzing two, or more, risks together, VaR might suggests not to hedge by means of a portfolio of random variables but to invest all the money into a single asset.

Otherwise said, the right hand side of inequality (4.6) reports that risks $\mathbf{X}$ and $\mathbf{Y}$, measured separately, are not as risky as the same risks included in a portfolio.

ES is instead sub-additive:

$$\mathrm{ES}_\alpha(\mathbf{X}) + \mathrm{ES}_\alpha(\mathbf{Y}) \geq \mathrm{ES}_\alpha(\mathbf{X} + \mathbf{Y}).$$

A complete discussion of risk measures can be found in Artzner et al. (1999) and Meucci (2009).

Appendix B

This section presents the ES back-test in order to assess the performance of the non-parametric and parametric models studied in the previous section. The ES back-test uses the unconditional test statistics proposed by Acerbi & Szekely (2014), given by

$$Z_{unc} = \frac{1}{Np_{\mathrm{VaR}}} \sum_{t=1}^{N} \frac{T_t I_t}{ES_t} + 1$$

where N is the number of observations, T_t is the temperature for period t, p_{VaR} is the probability of VaR failure defined as $1 -$ VaR level, I_t is the VaR failure indicator on period t with a value of 1 if $T_t < -\mathrm{VaR}_t$, and 0 otherwise and ES_t is the estimated expected shortfall for period t expressed as follows:

$$ES_t = -E_t \left[\frac{T_t I_t}{p_{\mathrm{VaR}}} \right],$$

where the operator E_t stands for the expected value conditioned to information available up to time t.

The expected value for this test statistics is 0; it should be negative when there is evidence of risk under-estimation. To determine how negative this estimate should be for the model to be rejected, some critical values are needed. These, of course, require some distributional assumptions on the observed variable T_t.

The unconditional test statistics turns out to be stable across a range of distributional assumptions for T_t, from thin-tailed distributions through the normal one and up to heavy-tailed distributions such as Student's t with a few degrees of freedom. Only very heavy-tailed Student's t-distributions lead to more noticeable differences in the critical values (see Acerbi & Szekely (2014)). The ES back-test takes advantage of the stability of the critical values of the unconditional test statistic and uses tables of pre-computed critical values to run ES back tests. The ES back-test has two sets of critical-value tables. The first set of critical values assumes that the temperature outcomes T_t follow a standard normal distribution; this is the unconditional normal test. The second set of critical values uses the heaviest possible tails, it assumes that the temperature outcomes T_t follow a t distribution with $\nu = 3$ degrees of freedom; this is the unconditional t-test.

The unconditional test statistics is sensitive to both the severity of the VaR failures relative to the ES estimate, and also to the number of VaR failures (how many times the VaR is violated, since, whenever losses exceed VaR, a VaR violation occurs). Expected average ratio is the average ratio of ES to VaR, over the period under scrutiny, with VaR failures while observed average severity ratio is the average ratio of loss to VaR over the periods with VaR failures.

Therefore, a single but very large VaR failure relative to the ES may cause the rejection of a model in a particular time window. A large increase/decrease of temperature on a day when the ES estimate is also large may not impact the test results as much as a large increase/decrease of temperature when the ES is smaller. A model can also be rejected in periods with many VaR failures, even if all the VaR violations are relatively small and only slightly higher than the VaR.

In Figure 4.14 the observed severity column shows the average ratio of increased/decreased temperature to VaR on periods when the VaR was violated. The expected severity column uses the average ratio of ES to VaR for the VaR violation periods. For the "Historical" and "Normal" models, the observed and expected severities are comparable. However, for the Historical method, the observed number of failures (red bars) is considerably higher than the expected number of failures (blu bars), about 2% higher. Both the "T5" and the "T10" models have observed severities much higher than the expected ones.

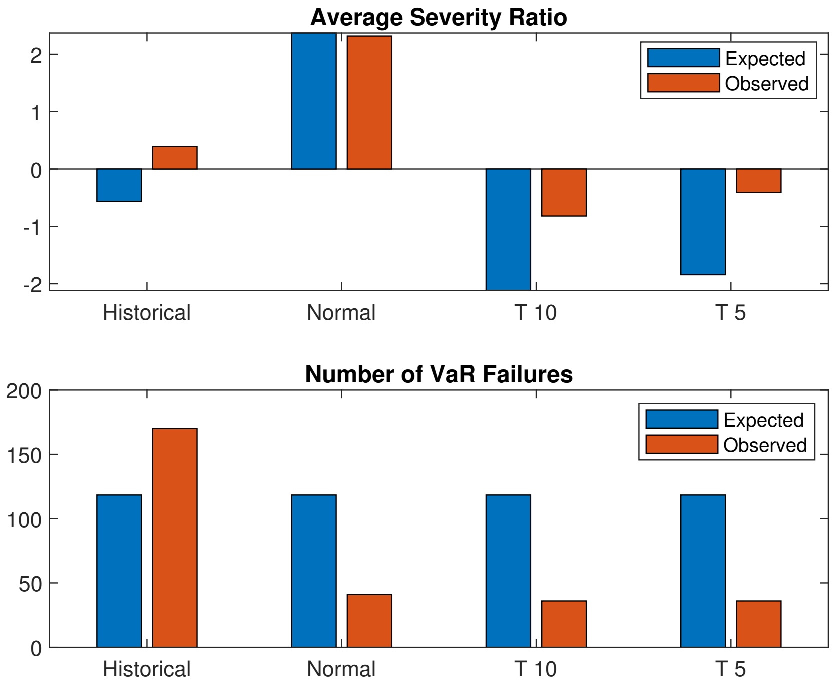


Figure 4.14: Severity of VaR failures relative to the ES estimate and to the number of VaR failures.

For the 2004–2016 window, all models pass both tests with a 95% level.

Even though all models pass both tests, it is clear from Figure 4.14 that T5 and T10 have a higher average severity ratio and number of VaR failures. To conclude this section, it results that from all the four methods (Historical, Normal, T5 and T10) under study, Historical and Normal VaR and ES are the ones that perform better when applied to the temperature data-set under scrutiny.

Appendix C

The following tables report some descriptive statistics for option data. Tables 4.9 and 4.10 show some descriptive statistics for seasonal HDD and CDD option contracts. Months considered for seasonal HDD are January to June, for seasonal CDD such months are May to September.

Table 4.9: Descriptive Statistics for seasonal HDDs for the period of time 1970–2016 and 10% percentile threshold.

Temp. (°C)	Mean	Var	Std	Min	Max	Skew.	Kurt.
VaR ($K = 4.25$ °C)	40.43	1 503.19	38.77	3.25	236.65	3.05	13.68
ES ($K = 2.27$ °C)	13.06	231.48	15.21	0	65.86	1.99	4.28

Table 4.10: Descriptive Statistics for seasonal CDDs for the period of time 1970–2016 and 90% percentile threshold.

Temp. (°C)	Mean	Var	Std	Min	Max	Skew.	Kurt.
VaR ($K = 26$ °C)	19.38	316.61	17.79	0	83.5	1.44	2.53
ES ($K = 20.26$ °C)	303.41	8 112.5	90.07	102.5	565.1	0.27	0.47

Tables 4.11 and 4.12 show descriptive statistics for June monthly HDD and CDD.

Table 4.11: Descriptive Statistics for monthly HDDs for the period of time 1970–2016 and 10% percentile threshold.

Temp. (°C)	Mean	Var	Std	Min	Max	Skew.	Kurt.
VaR ($K = 16.5$ °C)	40.43	3.44	17.46	4.18	0	1.77	2.78
ES ($K = 15.24$ °C)	1.08	4.33	2.08	0	9.04	2.44	5.62

Table 4.12: Descriptive Statistics for monthly CDDs for the period of time 1970–2016 and 90% percentile threshold.

Temp. (°C)	Mean	Var	Std	Min	Max	Skew.	Kurt.
VaR ($K = 24.4$ °C)	3.94	43.47	6.59	0	27.5	2.07	3.64
ES ($K = 19.97$ °C)	45.54	744.39	27.28	12.26	144.75	1.39	2.56

Bibliography

[1] Acerbi, Carlo, and Balazs Szekely. 2014. Back-testing expected shortfall. *Risk* 27: 76–81.

[2] Acerbi, Carlo, and Dirk Tasche. 2002. Expected shortfall: A natural coherent alternative to value at risk. *Economic Notes* 31: 379–88. [CrossRef]

[3] Alaton, Peter, Boualem Djehiche, and David Stillberger. 2002. On modelling and pricing weather derivatives. *Applied Mathematical Finance* 9: 1–20. [CrossRef]

[4] Alexandridis, Antonis K., and Achilleas D. Zapranis. 2013. The weather derivatives market. In *Weather Derivatives*. New York: Springer, pp. 1–20.

[5] Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath. 1999. Coherent measures of risk. *Mathematical Finance* 9: 203–8. [CrossRef]

[6] Barger, Norah, and Alan Adkins. 2013. Consultative Document: Fundamental Review of the Trading Book Rules—Further Response. Available online: https://www.gfma.org/wp-content/uploads/0/83/91/161/69270c81-740c-4447-bcdd-5ecbadd0a88a.pdf (accessed on 18 July 2020).

[7] Basel Committee on Banking Supervision. 2019. Minimum Capital Requirements for Market Risk. Available online: https://www.bis.org/bcbs/publ/d457.htm (accessed on 18 July 2020).

[8] Benth, Fred Espen, Gleda Kutrolli, and Silvana Stefani. 2019. Dynamic Probabilistic Forecasting with Uncertainty. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3405890 (accessed on 18 July 2020).

[9] Benth, Fred Espen, and Jurate Saltyte-Benth. 2005. Stochastic modeling of temperature variations with a view towards weather derivatives. *Applied Mathematical Finance* 12: 53–85. [CrossRef]

[10] Benth, Fred Espen, and Jurate Saltyte-Benth. 2011. Weather Derivatives and Stochastic Modelling of Temperature. *International Journal of Stochastic Analysis*. [CrossRef]

[11] Benth, Fred Espen, and Jurate Saltyte-Benth. 2013. *Modeling and Pricing in Financial Markets for Weather Derivatives*. Singapore: World Scientific, vol. 17.

[12] Black, Fisher, and Myron Scholes. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81: 637–54. [CrossRef]

[13] Brix, Anders, Stephen Jewson, and Christine Ziehmann. 2005. *Weather Derivative Valuation: The Meteorological, Statistical, Financial and Mathematical Foundations*. Cambridge: Cambridge University Press.

[14] Kelly, David L., Charles D. Kolstad, and Glenn T. Mitchell. 2005. Adjustment costs from environmental change. *Journal of Environmental Economics and Management* 50: 468–95. [CrossRef]

[15] Le Den, Xavier, Matilda Persson, Audrey Benoist, Paul Hudson, Marleen de Ruiter, Lars de Ruig, and Onno Kuik. 2017. Insurance of Weather and Climate-Related Disaster Risk: Inventory and Analysis of Mechanisms to Support Damage Prevention in the EU. Available online: https://op.europa.eu/en/publication-detail/-/publication/4f366956-a19e-11e7-b92d-01aa75ed71a1/language-en (accessed on 18 July 2020).

[16] McNeil, Alexander J., Rüdiger Frey, and Paul Embrechts. 2015. *Quantitative Risk Management: Concepts, Techniques and Tools—Revised Edition*. Princeton: Princeton University Press.

[17] Merton, Robert C. 1973. Theory of rational option pricing. *The Bell Journal of Economics and Management Science* 4: 141–83. [CrossRef]

[18] Meucci, Attilio. 2009. *Risk and Asset Allocation*. Berlin: Springer.

[19] Mills, Evan. 2005. Insurance in a climate of change. *Science* 309: 1040–44. [CrossRef] [PubMed]

[20] Pindyck, Robert S. 2011. Fat tails, thin tails, and climate change policy. *Review of Environmental Economics and Policy* 5: 258–74. [CrossRef]

[21] Rockafellar, Tyrrell R., and Stanislav Uryasev. 2002. Conditional Value-at-Risk for general loss distributions. *Journal of Banking and Finance* 26: 1443–71. [CrossRef]

[22] Saltyte-Benth, Jurate, and Fred Espen Benth. 2012. A critical view on temperature modeling for application in weather derivatives markets. *Energy Economics* 34: 592–602. [CrossRef]

[23] Stefani, Silvana, Enrico Moretto, Matteo Parravicini, Simone Cambiaghi, Adeyemi Sonubi, Gleda Kutrolli, and Vanda Tulli. 2018. Managing adverse temperature conditions through hybrid financial instruments. *Journal of Energy Markets* 11: 25–41.

[24] Taib, Che Mohd Imran Che, and Fred Espen Benth. 2012. Pricing of temperature index insurance. *Review of Development Finance* 2: 22–31. [CrossRef]

[25] Tang, Chun-Hung Hugo, and SooCheong Shawn Jang. 2011. Weather risk management in ski resorts: Financial hedging and geographical diversification. *International Journal of Hospitality Management* 30: 301–11. [CrossRef]

[26] Tang, Chun-Hung Hugo, and SooCheong Shawn Jang. 2012. Hedging Weather Risk in Nature-Based Tourism Business: An Example of Ski Resorts. *Journal of Hospitality & Tourism Research* 36: 143–63.

[27] Wang, Zhiliang, Peng Li, Lingyong Li, Chunyan Huang, and Min Liu. 2015. Modeling and forecasting average temperature for weather derivative pricing. *Advances in Meteorology* 2015: 1–8. [CrossRef]

[28] Zenghelis, Dimitri. 2006. The Stern Review on the Economics of Climate Change. Available online: http://videolectures.net/site/normal_dl/tag=12665/ccog07_zenghelis_tsr_01.pdf (accessed on 18 July 2020).

Part II

The Second Part

Chapter 5

Dynamic probabilistic forecasting with uncertainty

Joint work with Fred Espen Benth and Silvana Stefani

Abstract

We introduce a dynamical model for the time evolution of probability density functions incorporating uncertainty in the parameters. The uncertainty follows stochastic processes, thereby defining a new class of stochastic processes with values in the space of probability densities. The purpose is to quantify uncertainty that can be used for probabilistic forecasting. Starting from a set of traded prices of equity indices we do some empirical studies. We apply our dynamic probabilistic forecasting to option pricing, where our proposed notion of model uncertainty reduces to uncertainty on future volatility. A distribution of option prices follows, reflecting the uncertainty on the distribution of the underlying prices. We associate measures of model uncertainty of prices in the context of Cont (2006). As a further application we look at the Sharpe ratio and the VaR measure of market risk as well, proposing some decision rules for investors, regulators and risk managers.

Keywords: Probability density, model uncertainty, risk measure, volatility, option prices

5.1 Introduction and Motivation

Forecasting the value of a financial asset or position is typically based on a stochastic model, where forecasts are derived as the mean or the quantiles. In this paper we propose forecasting based on a stochastic model of the probability distribution. Indeed, we suggest to incorporate model uncertainty by considering forecasting using dynamical stochastic evolutions of the probability distribution of the model in question.

In science, forecasting is important, in particular in systems with uncertainty. A random variable X is supposed to span out the possible outcomes from an uncertain event, like for example a game of dices or the return from an investment. Although theoretically, X is a measurable map from the probability space $(\Omega, \mathcal{F}, P)$ into $\mathbb{R}$, the common approach is to describe it via its probability distribution, or rather, its density. Hence, often in statistics and probability, we do not specify the function

$$X : \Omega \to \mathbb{R}$$

in modeling, but rather its probability density p (assuming it exists)

$$\mathbb{R} \ni x \to \frac{d}{dx} P(X \leq x) := p(x)$$

Inference and forecasting is done with the use of p. A classical forecast is the mean,

$$m := \int xp(x)dx$$

Quantile forecasts are also important, finding x_q such that

$$\int_{-\infty}^{x_q} p(x)dx = q$$

for some $q \in (0,1)$. Sometimes the density is complex, like for example in hierarchical Bayesian modeling where Bayes' formula is applied in modeling with conditional densities (see Jokhadze & Schmidt (2020)). An alternative is simulation, however, simulation methods like Monte Carlo often rely on knowledge of the density.

In time series models, one starts out with a random variable as noise, and puts this into motion through some iterative scheme. For example, a simple AR(p)-time series is defined as first introducing a sequence of IID variables (ϵ_n), and then defining

$$X(n+p) + \alpha_1 X(n+p-1) + ... + \alpha_p X(n) = \epsilon_n \tag{5.1}$$

However, in the modeling the fundamental assumption is usually a specification of the *distribution* of ϵ_n, and *not* a specific representation as a real-valued function on Ω. A typical choice is to let ϵ_n be normally distributed. As such, we start out with a given distribution (e.g. the normal), and through some inductive iteration (as in (5.1)) we *move* the given distribution forward as a function of n. This gives, as a matter of fact, a *deterministic* evolution of the prescribed distribution of ϵ_n. In principle, we can completely determine the distribution of the system at any time step n, once specifying the distribution of ϵ_n. Indeed, for an AR(p) model with normally distributed ϵ_n, X will be again normally distributed, completely characterised by ϵ and the α's.

Stochastic differential equations (SDE) are, in some sense, time series models in continuous time. These provide a mathematically very detailed description of the dynamics of some phenomena. For example, SDEs have been used as models for prices of financial assets like stocks. But also SDEs, like time series models, starts out with a fundamental process, typically being the Brownian motion B, that is the driver in the dynamics. B has increments being independent and normally distributed, and is the time continuous analogue of the random walk ϵ_n. An SDE describes the future states of the system, which means for a stock price, the future possible stock prices. If the stock price at time t is X_t, given by an SDE, then again we can in principle completely determine the distribution of X_t at any future time $t > 0$, indeed $X_t := X_t(B)$, a functional of B and thus of the normal distribution. Hence, we can say that the distribution of X_t is *deterministically* determined from B.

In real-world situations, we do not know, of course, the exact distribution of ϵ, nor the driver in the SDE B. Our specific choice is only our (best) model of the situation, and sometimes not even that, but a pragmatic choice to allow for certain tractability of the model. But far more important, we do not know the exact dynamics forward in time, that is, we do not know how the system is evolving dynamically. This is indeed also prone to uncertainty, and is specified pragmatically and by insight into the system based on the information we have available. Going back to the stock price dynamics, we do not have available any physical law, nor economical law that can prescribe to us exactly how the prices dynamically move in time. We base the dynamics on empirical probabilistic properties of prices, that we have observed historically. As a modeler, we look at distributional properties, dependency properties over time, and select a model that can explain these stylized features. However, in principle, there may be many stochastic models able to explain our stylized facts. And even more, there may be other effects that we have not observed, and that we therefore do not take into account. The financial crisis around 2007 demonstrated this, where the existing models typically were far from being able to forecast the extreme moves that occurred in the markets.

SDEs can be very complex, and therefore it may be very hard to reveal the actual distribution and thereby make forecasts. For example, quantile forecasts may require simulation from the SDE, which is very time consuming if we look for tails of the distribution. We need many samples to accurately describe the tails of the distribution, along with iterations in time to reach the future forecasting point. Hence, although the SDE may feel attractive as a modeling device, it can lead to untractability when it comes to practical forecasting.

To keep the context of the financial example, we are interested in forecasting future prices. Our concern could be investment decisions, or risk evaluations of current positions, or, fashionably in mathematical finance, assessing financial derivatives contracts. All these questions rely on knowledge of the future probability distribution of prices. If we decide for an SDE as the model for future prices, we do, as explained above, move deterministically forward a given model, that is, a given distribution. There is no room for uncertainty about the distribution that we move forward, neither any uncertainty on the *moving forward dynamics*. As we are interested in the probability distribution, we propose in this paper an alternative approach: *A dynamical model of the probability distribution where we include uncertainty*, which follows stochastic processes.

Our modeling paradigm is in line with classical probability theory in that we describe mathematically a random event through its probability distribution. We consider probability densities that evolve in time, $t \mapsto p_t$. The added ingredient is that we want to allow for uncertainty in the distribution, and therefore consider p_t as a *random variable* in some appropriate state space of probability density functions, that is $t \mapsto p_t$ is a stochastic process with values in a state space of density functions.

Our modeling proposal cover classical SDEs. The solution, X_t, has a probability law p_t^X that evolves deterministically via the Chapman-Kolmogorov (forward) equation. Indeed, we cover all Markovian processes, as these have associated Chapman-Kolmogorov equations for the distribution density. It also covers model uncertainty by Cont (2006), where the model uncertainty is prescribed through uncertainty about the probability measure. We apply our dynamic probabilistic forecasting to option pricing, where the notion of model uncertainty it is reduced to uncertainty on future volatility and a distribution of option prices follows. This distribution reflects the uncertainty on the distribution of the underlying prices.

The idea put forward in this paper faces several challenges. First of all, we need to specify stochastic dynamical models for the density. Secondly, these models must be benchmarked against reality, that is, we need to be able to make inference on the model using data. On the other hand, we could possibly come up with very simple models capturing uncertainty in an easy way, and allowing for simpler procedures for forecasting. In addition, it also leads to interesting new models in more abstract situations, like the definition of a compound Poisson process for probability densities.

The paper is structured as follows. We start in Section 5.2 by discussing general density processes with values in $L^1(\mathbb{R}^d)$ and proposing various dynamical models introducing uncertainty in the probability density of a random event. A detailed empirical analysis of financial data from Europe and the US are performed in Section 5.3, including estimation of Ornstein-Uhlenbeck dynamical models assessing the uncertainty in volatility. In this section we focus on an application of our ideas to call option pricing. Model uncertainty measures are estimated by existing and proposed frameworks and some VaR analysis and backtesting is performed in Section 5.4. In Section 5.5 we point out some decision rules for measurement and management of risk. Section 5.6 concludes discussing our main findings.

5.2 Stochastic Probability Density Dynamics in $L^1(R^d)$

Assume we are given a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$. Furthermore, denote by $L^1(\mathbb{R}^d)$ the space of (equivalence classes) real-valued measurable functions on $\mathbb{R}^d, d \in \mathbb{N}$, which are integrable. It is well-known that $L^1(\mathbb{R}^d)$ is a separable Banach space with respect to the norm $|f|_1 := \int_{\mathbb{R}^d} |f(x)|dx$. Moreover, it is closed under convolution and forms a commutative Banach algebra under the convolution product.

We know that any probability density function p is a non-negative integrable function on $\mathbb{R}^d$ with $\int_{\mathbb{R}^d} p(x)dx = 1$. Moreover, as convolution is positivity preserving, and the $L^1(\mathbb{R}^d)$-norm of the convolution product of two positive functions is the product of their respective norms, the convolution of two densities is again a density. We can of course see this directly from the fact that the density of the sum of two independent random variables is the convolution of the marginal densities. A family $(g_t)_{t\geq 0}$ is called a $L^1(\mathbb{R}^d)$-valued stochastic process if it is adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$, that is, if g_t is an $\mathcal{F}_t$-measurable map from Ω into $L^1(\mathbb{R}^d)$ with the latter space equipped with the Borel σ-algebra.

Definition 1. *An $L^1(\mathbb{R}^d)$-valued stochastic process $(p_t)_{t\geq 0}$ is said to be a density process if for any $t \geq 0$, $p_t(x) \geq 0$ for all $x \in \mathbb{R}^d$ and $\int_{\mathbb{R}^d} p_t(x)dx = 1$.*

Consider now a parametric density function $f(\cdot, \theta)$, where $\theta \in C \subset \mathbb{R}^n$, $n \in \mathbb{N}$ with C being an open set.

Lema 2. *Assume that Θ is a C-valued random variable. If for a.e. $x \in \mathbb{R}^d$, $\theta \mapsto f(x, \theta)$ is continuous and for any neighborhood $U \subset C$ around θ there exists a function $h_U \in L^1(\mathbb{R}^d)$ such that $|f(x, \tilde{\theta})| \leq h_U(x)$ for $\tilde{\theta} \in U$, then $f(\cdot, \Theta)$ is an $L^1(\mathbb{R}^d)$-valued random variable.*

Proof. For each fixed ω, we have that $f(\cdot, \Theta(\omega)) \in L^1(\mathbb{R}^d)$. Moreover, as Θ is a random variable, it is a measurable map from Ω into the open set C. We must show that

f is a measurable map from Ω into $L^1(\mathbb{R}^d)$. Let θ_n be a sequence in C that converges to θ. Thus, for a given neighborhood U around θ, we find for sufficiently big n that $\theta_n \in U$. But then we find from dominated convergence theorem that

$$\lim_{n\to\infty} |f(\cdot,\theta_n) - f(\cdot,\theta)|_1 = \int_{\mathbb{R}^d} \lim_{n\to\infty} |f(x,\theta_n) - f(x,\theta)| dx = 0$$

and continuity in $L^1(\mathbb{R}^d)$ follows. The combination of a measurable map with a continuous function implies measurability.

We introduce next a sequence of $L^1(\mathbb{R}^d)$-valued random variables $(F_i)_{i\in\mathbb{N}}$, given as follows: Let $(\Theta_i)_{i\in\mathbb{N}}$ be an iid sequence of random variables. Define

$$F_i(x) := f(x, \Theta_i) \tag{5.2}$$

We have the following:

Lema 3. *Suppose that Θ has a density p_Θ. If $(\Theta_i)_{i\in\mathbb{N}}$ is an iid sequence of random variables distributed according to Θ, then $(F_i)_{i\in\mathbb{N}}$ in* (5.2) *is iid $L^1(\mathbb{R}^d)$-valued random variables with the same distribution as $f(\cdot,\Theta)$.*

Proof. Since Θ_i is distributed as Θ, F_i is distributed as $f(\cdot,\Theta)$ in $L^1(\mathbb{R}^d)$ for all $i \in \mathbb{N}$. By independence, we have that the density of (Θ_1,Θ_2) is $p_{\Theta_1}(\cdot)p_{\Theta_2}(\cdot)$. For any $A, B \in \mathcal{B}(L^1(\mathbb{R}^d))$, Borel sets of $L^1(\mathbb{R}^d)$, we have from conditioning

$$\begin{aligned} &P(F_1 \in A, F_2 \in B) \\ &\quad = \int_{C^2} P\left(F_1 \in A, F_2 \in B | \Theta_1 = \theta_1, \Theta_2 = \theta_2\right) p_{\Theta_1}(\theta_1) p_{\Theta_2}(\theta_2) d\theta_1 d\theta_2 \end{aligned}$$

But for given θ_1, θ_2, we find that

$$P(f_1(\cdot,\theta_1) \in A, f_2(\cdot,\theta_2) \in B) = P(f_1(\cdot,\theta_1) \in A) P(f_2(\cdot,\theta_2) \in B)$$

as these probabilities are zero-one probabilities (either f_i is in the set, or not). Thus, we find that

$$P(F_1 \in A, F_2 \in B) = P(F_1 \in A) P(F_2 \in B)$$

and the Lemma follows.

Let $N(t)$ be a Poisson process with values on $\mathbb{N} \cup \{0\}$, having an intensity $\lambda > 0$. Define the process

$$C(t) := g * \otimes_{i=1}^{N(t)} F_i \tag{5.3}$$

where g is a probability density function and $\otimes_{i=1}^{N(t)}$ signifies $N(t)$ times iterated use of the convolution product $*$ in $L^1(\mathbb{R}^d)$. We say that $C(t)$ is a *convolved Poisson process*, in some sense the natural analogue of a compound Poisson process for densities.

Lema 4. *$C(t)$ is a density process in $L^1(\mathbb{R}^d)$.*

Proof. First, for each $\omega \in \Omega$, we have that $F_1 * F_2 * \cdots * F_n \in L^1(\mathbb{R}^d)$, being positive by the property of the convolution product, and with integral equal to 1. Hence, it will be a density function. The convolution product is also a continuous function on $L^1(\mathbb{R}^d)$, so $C_n := g * \otimes_{i=1}^n F_i$ will be a random walk time series with values in $L^1(\mathbb{R}^d)$ for $n \geq 1$. Define $C(t) := C_{N(t)}$, which is then a subordination of C_n. As N is $\mathcal{F}_t$-adapted, we find that $C(t)$ is also $\mathcal{F}_t$-adapted. Thus the result follows.

The iid sequence F_i and the convolved Poisson process defined above constitute examples of dynamical models for the stochastic evolution of a probability density. More important for our empirical studies is a class of Gaussian models, that we consider next.

5.2.1 A particular Gaussian Model

Let M and Σ^2 be two random variables with values on $\mathbb{R}$ and $\mathbb{R}_+$, respectively. We use the notation $\mathbb{R}_+$ for the positive real line not including origin. Define f as

$$f(x, m, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) \tag{5.4}$$

and consider the $L^1(\mathbb{R})$-valued random variable $f(\cdot, M, \Sigma^2)$. Thus, we use the Gaussian density function for a variable with mean m and variance σ^2, combined with a parameter-valued bivariate random variable (M, Σ^2). We note that we can bound this function around a neighborhood of any (m, σ^2) by again a Gaussian function, and moreover, the map $(m, \sigma^2) \to f(x, m, \sigma^2)$ is continuous. Thus, $f(\cdot, M, \Sigma^2)$ is an $L^1(\mathbb{R})$-valued random variable by Lemma 2.

Consider now an example of pricing of call options. In the Black & Scholes paradigm, the price of a call option is given by the Black & Scholes formula (Black & Scholes (1973))

$$BS(S, K, r, T, \sigma) = S\Phi(d_1) - Ke^{-rT}\Phi(d_2) \tag{5.5}$$

where K is the strike price, T the exercise time, S the current stock price, r the risk-free interest rate and

$$d_{1,2} = \frac{\ln(S/K) + (r \pm \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}$$

This formula is based on a stock price with distribution at time T being log-normal, e.g., $\ln S(T) - \ln S$ is normally distributed with mean $(r - \frac{1}{2}\sigma^2)T$ and variance $\sigma^2 T$. However, if we are uncertain about the actual distribution at time T of the log-price, we can consider the log-price being distributed according to random density f. If we say that r is known with certainty, we have $M = r - \frac{1}{2}\Sigma^2$, and Σ^2 is some random variable distributed in $\mathbb{R}_+$. Then the option price will be $BS(S, K, r, T, \Sigma)$, with $d_{1,2}$ being random variables given in terms of Σ^2 (substituting σ with Σ). This means that $BS(S, K, r, T, \Sigma)$ also becomes a random variable. A challenge is then: what is the price of the option, given that in finance we should have *a price*, and not many prices! One can start discussing prices from the issuer point of view, or from the buyer, taking into account risk perception. We analyse this empirically in Sections 5.3 and 5.4.

Next, consider the convolved Poisson density process $(C(t))_{t\geq 0}$ defined in (5.3), with $F_i := f(\cdot, M_i, \Sigma_i^2)$ for (M_i, Σ_i^2) independent identically bivariate distributed random variables. Let us look at the mean forecast at time t: By definition, the mean forecast is

$$\begin{aligned}\widehat{M}(t) &= \int_{\mathbb{R}} xC(t)(x)dx \\ &= \int_{\mathbb{R}} x(g * \otimes_{i=1}^{N(t)} F_i(x))dx \\ &= \int_{\mathbb{R}} xg(x)dx + \sum_{i=1}^{N(t)} \int_{\mathbb{R}} xF_i(x)dx \\ &= \int_{\mathbb{R}} xg(x)dx + \sum_{i=1}^{N(t)} M_i\end{aligned}$$

In conclusion, the mean forecast is a compound Poisson process on $\mathbb{R}$, with intensity λ and jump sizes given by M.

Indeed, as the normal density is closed under convolution by summing mean and variances, we find that

$$C(t) = f(\cdot, m + \sum_{i=1}^{N(t)} M_i, \sigma^2 + \sum_{i=1}^{N(t)} \Sigma_i^2) \tag{5.6}$$

where $m = \int_{\mathbb{R}} xg(x)dx$, $\sigma^2 = \int_{\mathbb{R}}(x-m)^2 g(x)dx$ and $g(x)$ is the density function of $N(m, \sigma^2)$. Thus, we see that $C(t)$ is the density of a normal mean-variance mixture model, where the mean is given as the (real-valued) compound Poisson process $m + \sum_{i=1}^{N(t)} M_i$ and the variance by the compound Poisson process with values on the positive half-line $\sigma^2 + \sum_{i=1}^{N(t)} \Sigma_i^2$. This representation motivates a new type of density processes in $L^1(\mathbb{R}^d)$, given by the following definition:

Definition 5. *A conditional density process is given by $C(t) := f(\cdot, \Theta(t))$ where $f(\cdot, \theta) \in L^1(\mathbb{R}^d)$ is a probability density and $t \mapsto \Theta(t)$ is a stochastic process with values in the parameter space of f.*

If $f(\cdot, m, v^2)$ is the normal density function, we see from (5.6) that the convolved Poisson process $C(t) = g * \otimes_{i=1}^{N(t)} f_i(\cdot, M_i, \Sigma_i^2)$ can be represented as a conditional density process.

A natural extension of the convolved Poisson dynamics analysed above is to choose two stochastic processes $(X(t))_{t\geq 0}$ and $(Y(t))_{t\geq 0}$, with state spaces in $\mathbb{R}$ and $\mathbb{R}_+$, respectively.

Define a conditional density process by $C(t) = f(\cdot, X(t), Y(t))$, with f being the normal density function. Then we find that the mean forecast is

$$\begin{aligned}\widehat{M}(t) &= \int_{\mathbb{R}} xC(t)(x)\,dx \\ &= \int_{\mathbb{R}} xf(x, X(t), Y(t))dx \\ &= X(t). \end{aligned} \tag{5.7}$$

Thus the mean forecast follows the stochastic process X. If this is a stationary process, the forecasted mean will be stationary and is distributed according to the law of X. Later, in our empirical studies, we will use an Ornstein-Uhlenbeck process driven by a Brownian motion for X (see Subsection 5.3.1).

A probabilistic forecast of the variance at time t will be

$$\begin{aligned}\widehat{V}(t) &= \int_{\mathbb{R}} (x - X(t))^2 C(t)(x)\,dx \\ &= \int_{\mathbb{R}} (x - X(t))^2 f(x, X(t), Y(t))\,dx \\ &= Y(t). \end{aligned} \tag{5.8}$$

Also here we may choose Y as a stationary process, now naturally distributed on $\mathbb{R}_+$. In our empirical studies, we will consider an Ornstein-Uhlenbeck process driven by a subordinator (see Subsection 5.3.1), but also discuss a Gaussian Ornstein-Uhlenbeck model as a simple approximation of the uncertain variance.

Let us study a quantile forecast at time t, i.e., defined as $Q_\alpha(t)$ for $\alpha \in (0, 1)$ such that

$$\int_{-\infty}^{Q_\alpha(t)} f(x, X(t), Y(t))dx = \alpha$$

By changing variables $y = (x - X(t))/\sqrt{Y(t)}$, we find that

$$\int_{-\infty}^{Q_\alpha(t)} f(x, X(t), Y(t))dx = \int_{-\infty}^{(Q_\alpha(t)-X(t))/\sqrt{Y(t)}} f(y, 0, 1)dy$$

and therefore

$$Q_\alpha(t) = X(t) + \sqrt{Y(t)}q_\alpha \tag{5.9}$$

with q_α being the α-quantile of a standard normal distribution. This expression provides us with a dynamic for the VaR, incorporating uncertainty on the mean and variance of the underlying.

5.3 Empirical Analysis on Stock Indices

When examining financial assets it is most common to study the returns of stock prices rather than the actual raw asset prices because returns are approximately symmetrically distributed around 0 (or around somewhere close to 0) and behave roughly similar and uncorrelated of each other. For the purposes of this study we will examine the daily log-returns for FTSEMIB, S&P500 and FTSE100 equity indices,

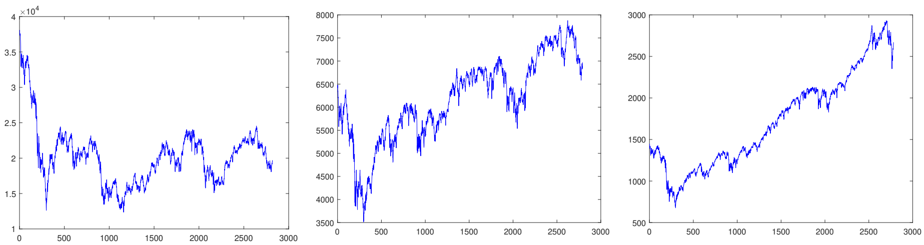
$$R(t) := \log(\frac{S_t}{S_{t-1}}). \tag{5.10}$$

Table 1 summarizes some statistics for the daily log-return series computed from the closing prices obtained from *Investing.com* and covering the period 1 January 2008 to 21 January 2019 (2820, 2781 and 2792 daily observations for FTSEMIB, S&P500 and FTSE100, respectively).

Table 5.1: Statistics of the daily log-returns of FTSEMIB, S&P500 and FTSE100, 2008 - 2019.

	Mean	Var	Std Dev	Skewness	Kurtosis	J-B
FTSEMIB	0.0002	0.0003	0.0169	0.1934	4.6182	2522.67
S&P500	-0.0002	0.0002	0.0127	0.3472	10.8043	13577.43
FTSE100	-0.00003	0.0002	0.0120	0.1459	8.3649	8147.03

The mean returns are almost identical for all series and close to zero. All series are slightly positively skewed and there is a presence of outliers in all of them. For daily returns, the null hypothesis of normality is strongly rejected by the Jarque-Bera (J-B) statistic (the 5% critical values of J-B is 5.99) for all indices.

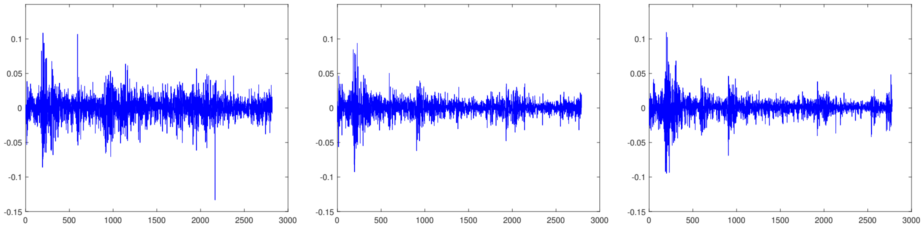


a. FTSEMIB b. S&P500 c. FTSE100

Figure 5.1: Time series of index values of FTSEMIB, S&P500 and FTSE100.

For a brief insight into the underlying returns we will first look at some of the plots of the index values and the log-returns of the data. Figure 1 shows a time series plot of all the three indices, where we see some similarity between S&P500 and FTSE100 whereas FTSEMIB has a slightly different evolution. We observe from Figure 5.2 that the log-returns fluctuate randomly around zero in a relatively symmetric fashion with occurring clustering of volatility. One may also observe a slight tendency towards a decreasing volatility with time in the indices S&P500 and FTSE100. We are going to focus on the log-returns for further examination.

A main application of this study is to analyse the model uncertainty in option pricing, where the notion of model uncertainty is reduced to uncertainty on future volatility when



a. FTSEMIB b. S&P500 c. FTSE100

Figure 5.2: Log-returns of FTSEMIB, S&P500 and FTSE100.

confining ourselves to the Black & Scholes Gaussian paradigm. Thus, starting from a set of traded prices of equity indices we calculate volatilities, and after that calibrate our pricing model and associate a measure of model uncertainty of prices. For this we first need to determine a dynamic for the volatility. Another application is uncertainty in forecasting VaR, for which we also need a dynamic for the mean. Recalling the analysis in Section 5.2, we focus on conditional density processes to assess the uncertainty in the distribution. As we can see from (5.7) and (5.8), the mean and variance require stochastic models in order to further study the uncertainty in option prices or forecasting VaR. For these purposes, we propose suitable Ornstein-Uhlenbeck processes in the next Subsection.

5.3.1 Dynamic Modeling with Ornstein-Uhlenbeck Processes

Representing the mean and variance in the normal distribution of the log-return indices by Ornstein-Uhlenbeck processes allow us to model the uncertainty by different speeds of mean reversion and incorporating a mixture of jump and diffusional behaviour. Based on empirical findings, we suggest a Gaussian Ornstein-Uhlenbeck process for the mean and a non-Gaussian Ornstein-Uhlenbeck process with a gamma limiting distribution for the volatility.

The *Gaussian* Ornstein-Uhlenbeck process $(X(t))_{t\geq 0}$ satisfies the following stochastic differential equation:

$$dX(t) = \alpha(\mu - X(t))dt + \beta dB(t) \tag{5.11}$$

where α and β are positive constant parameters, μ is the long-term mean of the process, and $(B(t))_{t\geq 0}$ is a Brownian motion. As can be seen in the equation above, the process is expected to exhibit some form of regression to the mean, because deviations from the mean μ effectively induce restoring forces. It is explicitly given by (see Benth, Šaltytė Benth & Koekebakker (2004))

$$X(t) = X(0)e^{-\alpha t} + \mu(1 - e^{-\alpha t}) + \beta \int_0^t e^{-\alpha(t-s)} dB(s). \tag{5.12}$$

The limiting distribution is normal, with mean μ and variance $\beta^2/2\alpha$.

A discretization of the Ornstein-Uhlenbeck process connects it to an AR(1) process, i.e., a first-order auto-regressive model. With our notation, the AR(1) process is defined recursively as follows for time steps of length $\Delta > 0$:

$$x(t) = \mu(1 - e^{-\alpha\Delta}) + e^{-\alpha\Delta}x(t - \Delta) + \epsilon(t) \tag{5.13}$$

where $\epsilon(t)$ is iid Gaussian random variables with zero mean and variance $\beta^2(1 - e^{-2\alpha\Delta})/2\alpha$. Equation (5.13) can be represented in a more general form as:

$$x(t) = a + bx(t - \Delta) + \epsilon(t) \tag{5.14}$$

Some parameter constraints are necessary for the model to remain wide-sense stationary. In particular, AR(1) processes with $|b| > 1$ are not stationary.

For the *non-Gaussian* Ornstein-Uhlenbeck model we assume the dynamics

$$dY(t) = -\xi Y(t)dt + dL(t) \tag{5.15}$$

where ξ is a positive constant and $(L(t))_{t\geq 0}$ is a subordinator (i.e., a non-decreasing Lévy process). This model is indeed motivated from the Barndorff-Nielsen and Shephard volatility model, see Barndorff-Nielsen & Shephard (2001). From Benth, Šaltytė Benth & Koekebakker (2004) we know the explicit representation as

$$Y(t) = Y(0)e^{-\xi t} + \int_0^t e^{-\xi(t-s)}dL(s)$$

As we will see, it is convenient to model the uncertain volatility by such a process, and in particular using a compound Poisson process as $(L(t))_{t\geq 0}$ with exponential jumps. This yields a limiting distribution for Y being in the class of Gamma distributions. I.e., let

$$L(t) = \sum_{i=1}^{N(t)} J_i$$

for a Poisson process $(N(t))_{t\geq 0}$ with jump intensity $\lambda > 0$ and jumps $(J_i)_{i=1}^{\infty}$ being iid exponentially distributed random variables with parameter ζ. The limiting distribution of $Y(t)$ as $t \to \infty$ will be Gamma distributed, that is, $\lim_{t\to\infty} Y(t) \sim \Gamma(1/\zeta, \lambda/\xi)$ (see Benth, Šaltytė Benth & Koekebakker (2004)).

The AR(1)-models above raise the immediate question of generalizations, and the link to more sophisticated models like GARCH, say, comes to mind (see e.g., Primiceri (2005) and Bolleslev (1986)). However, we emphasise here that we do not aim at modeling the volatility processes, but the randomness in the probability distribution dynamics. On the other hand, in future studies one may try out more sophisticated models for the mean and variance uncertainty in the probability distribution, following for example recent studies as in (Gneiting & Ranjan (2013); Amisiano & Giacomini (2007); Billio, Casarin, Ravazzolo, & van Dijk (2013), and Geweke & Amisano (2010)). Here one can find that a probabilistic forecast can be represented in the form of a predictive cumulative distribution function which can be discrete, discrete-continuous or continuous. However, still the focus in these papers is *not* on random probability distributions.

5.3.2 Mean and Variance of Log-Returns

As part of our empirical examination, we study the uncertainty in the historical mean and variance of log-returns. So, after we have chosen to work on three time series of stock prices the next step is considering the running mean of log-returns for those time series, defined over a window of length N. Recalling the definition of log-returns in (5.10), the running mean is

$$\hat{m}_N(t_i) := \frac{1}{N} \sum_{k=0}^{N-1} R(t_{i-k})$$

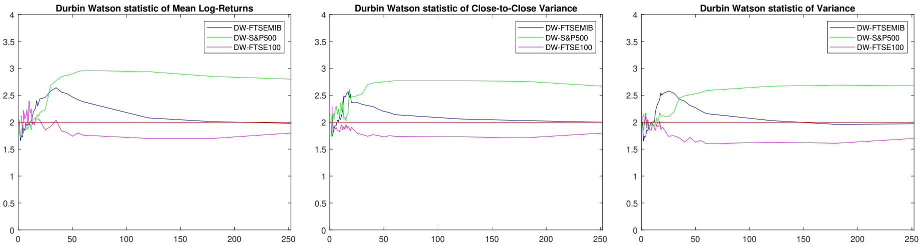
We also look at a variance estimator over a window in a similar way:

$$\hat{\sigma}^2_N(t_i) := \frac{1}{N-1} \sum_{k=1}^{N} (R(t_{i-k}) - \hat{m}_N(t_i))^2$$

where N is the number of observations or the so called rolling window size. We will start by studying the mean and variance time series as $AR(1)$-processes, to fit OU-processes $X(t)$ and $Y(t)$ (see Avellaneda & Lee (2010)). In fact, we shall in a first attempt model the variance Y *also* as a Gaussian Ornstein-Uhlenbeck process, as this is providing a simple link to empirical analysis of $AR(1)$ processes. Admittedly, a Gaussian Ornstein-Uhlenbeck process may give negative values, which is unreasonable for the variance. However, such a model is very simple from the empirical point of view, and provide some insight we believe. We will eventually define a non-Gaussian model for the variance, circumventing the problem of negative values. However, this model requires a slightly different estimation procedure.

Before going for AR(1) processes for mean and variance the very first problem we have to resolve is to find the most appropriate length of the rolling window we are going to use for calculating the running mean of log-returns and the variance estimator over that window. Hull (Hull et al (2009)) suggests that a good rule is to set the number of observations, N to the same amount of days that the volatility is to be applied to. This means that, if we want to estimate the price of an option with 30 days left to the expiration then we have to measure the historical volatility based on 30 days too. In our examination we try for different values of N that means for different historical days of index returns and we have found that N should be set for a time horizon from one to two weeks considering just trading dates. Figure 3 illustrates how the Durbin Watson statistic changes for different N and where to search for the most appropriate number. This statistic is a number that tests for auto-correlation in the residuals from a statistical regression analysis and it is always between 0 and 4 (see Durbin & Watson (1950) and Durbin & Watson (1951)). A value of 2 means that there is no auto-correlation in the sample. Values from 0 to 2 indicate positive auto-correlation and values from 2 to 4 indicate negative auto-correlation. In our case the critical values around 2 in order to conclude residuals have positive or negative correlation are 2 ± 0.15. By close inspection of the numbers behind Figure 5.3, we conclude that for FTSEMIB we will have no auto-correlation for $N = 11$ trading days, for S&P500 for $N = 7$ trading days while for FTSE100 for $N = 9$ trading days[1]. It is clear to see by a close inspection of

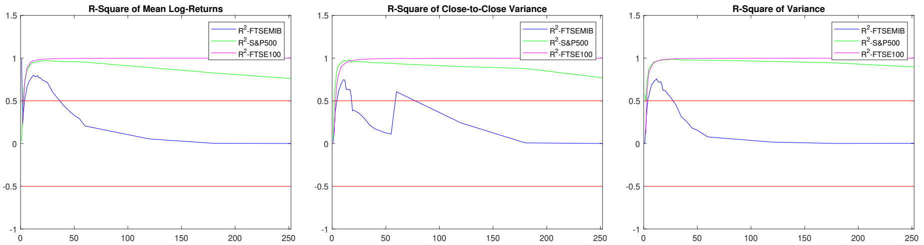
[1] This is admittedly not easy to conclude from the Figure, but read off from the estimated numbers



a. FTSEMIB b. S&P500 c. FTSE100

Figure 5.3: Durbin Watson statistic of mean log-returns, close-to-close variance and variance of FTSEMIB, S&P500 and FTSE100.

Figure 5.4 that the corresponding R^2 of all those results is statistically significant, being more than 50%.



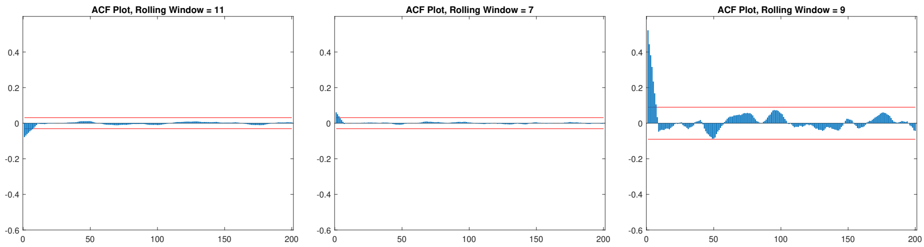
a. FTSEMIB b. S&P500 c. FTSE100

Figure 5.4: R^2 of mean log-returns, close-to-close variance and variance of FTSEMIB, S&P500 and FTSE100.

In general close-to-close variance is an approximation of variance (still a second moment but in the case when the mean is very close to zero and one does not consider it in calculating variance) but as it is said in Bennett & Gil (2012) close-to-close variance is suggested to be used for relatively short time periods (daily, weekly), cases in which the drift should be close to zero and can be ignored. In our case we are going to work with variance in order to have more accuracy in calculations.

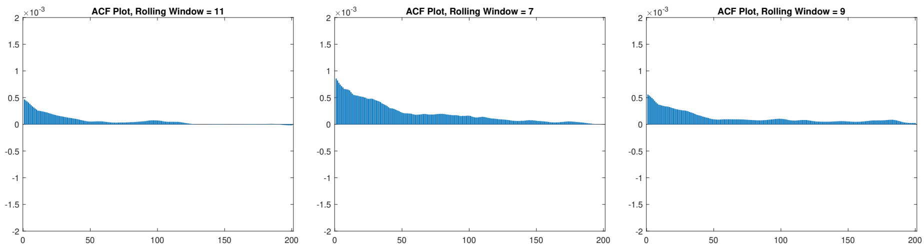
The same results we get also from the auto-correlation function. Figure 5.5 and 5.6 show some results on the auto-correlation function corresponding to the mean and variance of log-returns. The behaviour of mean and variance is typically as that of an auto-regressive process. Each value that rises above or falls below the red lines is considered to be statistically significant. Therefore, if a value is significantly different from zero, that is evidence of auto-correlation. A value that is close to zero is evidence against auto-correlation.

The values of the auto-correlation function for the variance are quite small compared to the mean, and almost all the values are inside the confidence interval. However, we still observe in Figure 5.6 a positive but decaying auto-correlation structure, which is evidence of an autoregression dynamics.



a. FTSEMIB b. S&P500 c. FTSE100

Figure 5.5: Auto-correlation function of mean log-returns for all indices.



a. FTSEMIB b. S&P500 c. FTSE100

Figure 5.6: Auto-correlation function of variance of log-returns for all indices.

After arguing for evidence of auto-correlation for the moving window mean and variance we have to go further on modeling the uncertainty in order to forecast the future. We propose as already indicated to model the mean and variance uncertainty by Gaussian Ornstein-Uhlenbeck processes described in Section 5.3 in a first attempt. Later, we will expand on the variance model to a more realistic non-Gaussian dynamics.

To reinforce the fact why we go for an AR(1) and not with a higher-order autoregressive model AR(p) for $p > 1$, we present some more results from an econometric analysis we have done using variable selection techniques. Below we display the corresponding results based on the "All possible Regressions" technique. This algorithm fits all regressions involving one regressor, two regressors, three regressors, and so on (we have chosen to illustrate cases of autoregressive processes until the $10th$ order). The selection criterion is recorded for each regression. Once the procedure finishes, we analyze our results based on some criteria and then determine which order of autoregressive process is optimum for our cases under study. As a criterion we have chosen to use R^2, R^2_{adj}, AIC or BIC, and look for the model where this value stabilizes for R^2, R^2_{adj} and the smallest value for AIC or BIC. Below we have presented the results of the all possible regressions procedure for all indices under study.

One can see from the table 5.2, 5.3, and 5.4 that from order 1 to order 10 it is an increase of 3.6% of R^2 and 3.5% of R^2_{adj} for mean and 0.28% of R^2 and 0.25% of R^2_{adj} for variance in case of FTSEMIB index, 7.9% of R^2 and 7.8% of R^2_{adj} for mean and 1.12% of R^2 and 1.09% of R^2_{adj} for variance in case of S&P500 index and 8.95% of R^2 and 8.9% of R^2_{adj} for mean and 1.45% of R^2 and 1.43% of R^2_{adj} for variance

Table 5.2: 'All possible Regressions' procedure for mean and variance of log-returns of FTSEMIB, 2008-2019.

FTSEMIB	Mean				Variance			
p	R^2	R^2_{adj}	AIC	BIC	R^2	R^2_{adj}	AIC	BIC
1	0.7759	0.7759	-15007.4	-14992.9	0.9257	0.9257	-22623.2	-22621.8
2	0.7798	0.7797	-14978.5	-14965.5	0.9258	0.9257	-22604.7	-22601.8
3	0.7850	0.7848	-14908.2	-14896.6	0.9259	0.9258	-22586.8	-22582.5
4	0.7853	0.7849	-14900.4	-14890.3	0.9264	0.9263	-22576.1	-22570.3
5	0.7873	0.7869	-14912.2	-14903.5	0.9273	0.9272	-22572.2	-22565
6	0.7885	0.7881	-14918.7	-14911.5	0.9278	0.9276	-22561.4	-22552.7
7	0.7890	0.7885	-14920.6	-14914.8	0.9278	0.9277	-22543	-22532.9
8	0.7927	0.7921	-14933.2	-14928.8	0.9280	0.9278	-22526.7	-22515.1
9	0.8050	0.8044	-14917.1	-14914.2	0.9281	0.9279	-22510	-22496.9
10	0.8116	0.8109	-14904	-14902.5	0.9285	0.9282	-22496	-22481.6

Table 5.3: 'All possible Regressions' procedure for mean and variance of log-returns of S&P500, 2008-2019.

S&P500	Mean				Variance			
p	R^2	R^2_{adj}	AIC	BIC	R^2	R^2_{adj}	AIC	BIC
1	0.6162	0.6161	-14656.69	-14642.28	0.9069	0.9068	-22127.66	-22126.21
2	0.6291	0.6287	-14655.80	-14642.83	0.9071	0.9070	-22111.62	-22108.74
3	0.6361	0.6357	-14660.91	-14649.38	0.9081	0.9080	-22104.78	-22100.45
4	0.6548	0.6543	-14490.46	-14480.36	0.9086	0.9085	-22087.94	-22083.61
5	0.6568	0.6562	-14461.96	-14453.31	0.9090	0.9088	-22076.14	-22070.37
6	0.6616	0.6609	-14457.02	-14449.8	0.9091	0.9089	-22064.06	-22056.85
7	0.6735	0.6726	-14462.65	-14456.88	0.9095	0.9092	-22045.55	-22036.9
8	0.6897	0.6889	-14408.98	-14404.65	0.9169	0.9167	-22032.15	-22022.06
9	0.6917	0.6908	-14391.42	-14388.53	0.9172	0.9170	-22117.59	-22104.63
10	0.6948	0.6938	-14363.84	-14362.4	0.9181	0.9177	-22109.4	-22095

in case of FTSE100 index. These changes show only a minor increase compared to the number of variables we have to consider in the model. Related to AIC and BIC statistics the best model from the set of plausible models being considered is the one with the smallest AIC/ BIC value. The fact that we have negative values of AIC and BIC indicates less information loss than a positive AIC/ BIC and therefore a better model (Baguley, (2012)). As a result from tables 2, 3 and 4 it is clear that the smallest values of AIC/ BIC corresponds to the first model (AR(1)). Aslo if we focus in the values of AIC and BIC the model selection will be done related to the lowest values of AIC and BIC. Therefore, we select the AR(1) model as our final model in all three cases.

Table 5.4: 'All possible Regressions' procedure for mean and variance of log-returns of FTSE100, 2008-2019.

FTSE100	Mean				Variance			
p	R^2	R^2_{adj}	AIC	BIC	R^2	R^2_{adj}	AIC	BIC
1	0.706	0.706	-15426.2	-15411.8	0.9428	0.9428	-23401.3	-23399.8
2	0.725	0.725	-15245.8	-15232.8	0.9432	0.9431	-23388.8	-23385.9
3	0.732	0.732	-15220.8	-15209.3	0.9438	0.9437	-23381.4	-23377.1
4	0.733	0.733	-15209.8	-15199.7	0.9441	0.9440	-23367.8	-23362
5	0.745	0.745	-15218.6	-15209.9	0.9443	0.9442	-23352.8	-23345.6
6	0.747	0.746	-15224.9	-15217.7	0.9472	0.9471	-23398.9	-23390.2
7	0.748	0.747	-15177.1	-15171.3	0.9475	0.9473	-23385.3	-23375.2
8	0.753	0.752	-15183.5	-15179.2	0.9486	0.9484	-23394	-23382.5
9	0.760	0.759	-15166.4	-15163.5	0.9487	0.9486	-23381.2	-23368.2
10	0.795	0.795	-15085.7	-15084.2	0.9573	0.9571	-23383.1	-23368.6

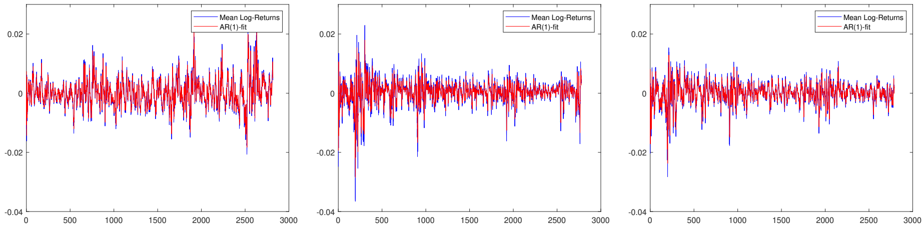
In conclusion, we identify the uncertainty dynamics of the mean and variance as AR(1) processes (for more details see Table 5).

Table 5.5: Fitted AR(1) processes of mean and variance for FTSEMIB, S&P500 and FTSE100, 2008-2019.

	Fitted Mean/ Var	Sig(Fisher)	R^2	DW
FTSEMIB	$x(t) = 0.8812x(t-1) + \epsilon_1(t)$	0.000	77.59%	1.99
	$y(t) = 0.9622y(t-1) + \gamma_1(t)$	0.000	92.57%	1.95
S&P500	$x(t) = 0.7809x(t-1) + \epsilon_2(t)$	0.000	61.62%	2.04
	$y(t) = 0.9522y(t-1) + \gamma_2(t)$	0.000	90.69%	2.05
FTSE100	$x(t) = 0.8392x(t-1) + \epsilon_3(t)$	0.000	70.58%	2.02
	$y(t) = 0.9777y(t-1) + \gamma_3(t)$	0.000	94.28%	1.95

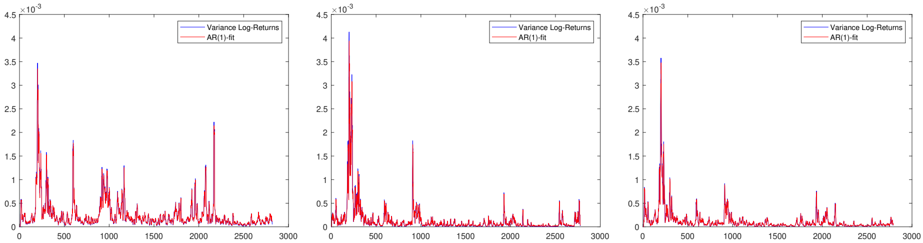
As is seen from Table 5 the goodness of fit for all models is quite good since the significance level from the Fisher test is less than 0.05. Also these results are further emphasized by R^2 values that are all bigger than 50%. For each model, it is associated the Durbin Watson statistic which is close to 2 for all cases. Now let us see how the AR(1) fit the mean and variance graphically and study the behaviour of residuals. Figure 5.7 and 5.8 illustrate how the AR(1) processes that we identify with "AR(1)-fit" line fit the mean and variance of log-returns. As a result they match very well the original time series.

We build Q-Q plot to test for normality of residuals (see Figure 5.9 and 5.10). The deviations from the straight line are minimal except the tails. The points in the Q-Q plot form a relatively straight line since the quantiles of the time series nearly



a. FTSEMIB b. S&P500 c. FTSE100

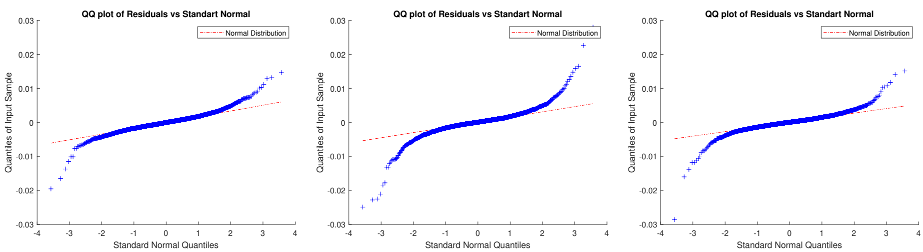
Figure 5.7: Fitted mean of log-returns for all indices.



a. FTSEMIB b. S&P500 c. FTSE100

Figure 5.8: Fitted variance of log-returns for all indices.

match what the quantiles of the time series would theoretically be if the time series was perfectly normally distributed. Compared to the normal distribution there is more data concentrated in the center of the distribution and less data in the tails. These “fat tails” correspond to the first quantiles occurring at less than expected values and the last quantiles occurring at larger than expected values. This can be clearly seen by the values of skewness and kurtosis for each index. We notice symmetry in distribution for European stock indices and some deviations of symmetry for the American one associated with presence of outliers. "Fat tails" affects the variance of the distribution.

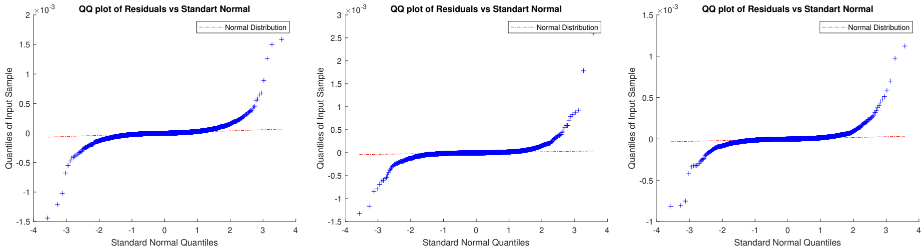


a. FTSEMIB b. S&P500 c. FTSE100

Figure 5.9: Residuals: mean of log-returns for all indices.

As it is seen from the normal Q-Q plots the centre of the data displayed normality while it was in the tails that we viewed deviation from the normal.

For a further estimation of residuals processes for mean and variance one can find



a. FTSEMIB b. S&P500 c. FTSE100

Figure 5.10: Residuals: variance of log-returns for all indices.

Table 5.6: Residual's normality for mean and variance log-returns of FTSEMIB, S&P500 and FTSE100, 2008-2019.

Mean	Skewness	Kurtosis	Variance	Skewness	Kurtosis
FTSEMIB	0.0174	6.8653	FTSEMIB	0.0612	103.6371
S&P500	0.5664	13.9104	S&P500	-3.0542	183.1418
FTSE100	0.5436	21.8828	FTSE100	-0.2795	100.01

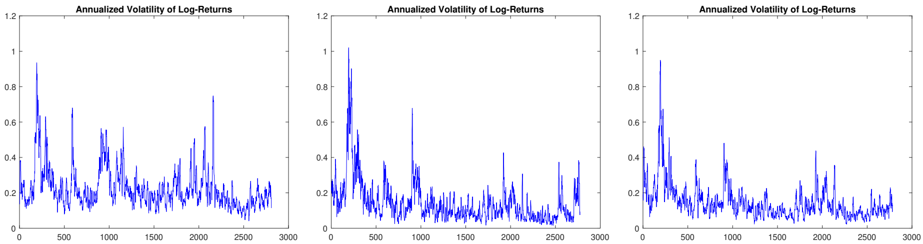
further analysis in the Appendix A.

5.3.3 Volatility

Uncertainty on volatility leads to model uncertainty and model risk. Since the volatility of an asset changes over time the measurement of the historical volatility is merely an estimate of the future volatility of the asset.

$$\text{Volatility} = \hat{\sigma}_N(t_i) := \sqrt{\frac{1}{N-1}\sum_{k=1}^{N}(R(t_{i-k}) - \hat{m}_N(t_i))^2}$$

Note that the number we got from this formula ($\widehat{\sigma}_N(t_i)$) is 1-day historical volatility. In our calculations we are interested in the annualized volatility which we get by multiplying the 1-day volatility by the square root of the number of (trading) days in a year (the average of trading days in a year is 252) (see Figure 11). If the mean is the average return of a stock index, variance and volatility can give us a sense of how much that stock index over/ under performs. If values of volatility are large a stock index value can potentially be spread out over a larger range of values. Therefore the price of the stock index can change dramatically over a short time period in either direction, which means it is riskier. Otherwise if values of volatility are small then a stock index value does not fluctuate dramatically, and tends to be more steady.

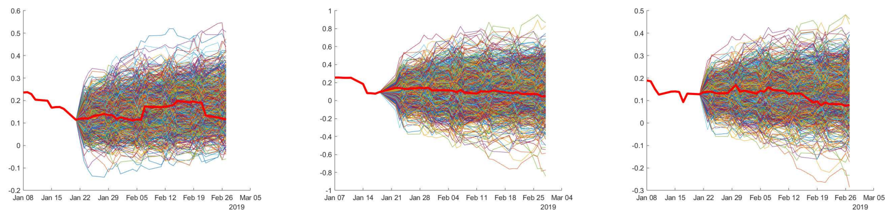


a. FTSEMIB b. S&P500 c. FTSE100

Figure 5.11: Annualized volatility of log-returns for all indices.

5.3.3.1 Simulation of Volatility by Gaussian Stationary Ornstein-Uhlenbeck Processes

We simulate volatilities $V(t)$ by using the approximated form of OU-process (equation 5.11) of an AR(1) form (equation 5.14). For this, at first we need to find the parameters a and b by regression and simulating the residuals as Gaussian iid variables. In Figure 5.12 it is shown the case of simulating 1000 trajectories of volatility by using a Gaussian OU-process where normal distributed residuals allow for negative values of errors which can result in negative volatilities as well. Indeed the probability and magnitude of these negative values are rather large.



a. FTSEMIB b. S&P500 c. FTSE100

Figure 5.12: Simulations of volatilities by using OU-process driven by a BM.

To avoid the problem of negative values in volatilities we propose to go for another model where volatilities are simulated by using another class of autoregressive models ensuring positive values by means of a non-Gaussian error term.

5.3.3.2 Simulation of Volatility by non-Gaussian Stationary Ornstein-Uhlenbeck Processes

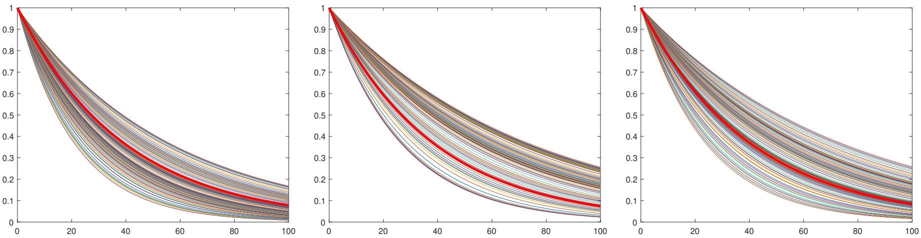
We simulate volatilities $V(t)$ by using equation 5.15. For this, at first we need to find a proper value of ξ, and we look for it referring to the k-lagged autocorrelation function as follows:

$$\text{Corr}(V(t), V(t+k)) = \exp(-\xi k), k = 1, 2, 3...$$

If we take the logarithm of both sides we will have:

$$\ln(\text{Corr}(V(t), V(t+k))) = -\xi k, k = 1, 2, 3...$$

In this way we will obtain a series of ξ-s from which we have to choose the proper one for our model. To decide which ξ is the best we are going to refer to the optimal value by minimizing the negative log-likelihood (NLL) function. We will choose as the best one the ξ which corresponds to the smallest value of NLL (see Figure 5.13).

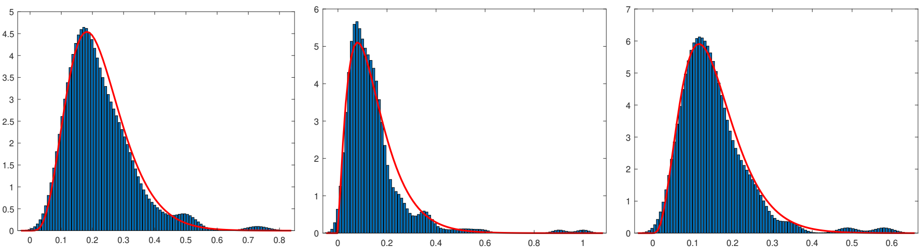


a. FTSEMIB b. S&P500 c. FTSE100

Figure 5.13: Autocorrelation function of OU process to fix the speed of mean-reversion ξ of $V(t)$.

The line in red in Figure 5.13 indicates the best ξ chosen from this calibration process.

Next we analyse the residuals. Let us consider every ℓ-th data from the original volatility time series. What we are interested in, is to obtain a new sequence of data which will be closely to zero-correlated (and treated as independent in order to enforce fitting the new sequence of data to a gamma distribution). As a consequence, by trying for different values of ℓ we found that the most proper value of ℓ for each index will be 25, 16, and 30, respectively. Let us plot a histogram for the new time series obtained and look for a fit of the gamma distribution (see Figure 5.14).



a. FTSEMIB b. S&P500 c. FTSE100

Figure 5.14: Fitted gamma distribution for new sample (based on every $25th$, $16th$, $30th$ data of the original data for each index, respectively).

Since in our cases we found a gamma distribution that fits well the new time series this means that $L(t)$ is a compound Poisson process with jumps exponentially distributed (see Schoutens (2003)). The estimated parameters for gamma distribution are shown in Table 5.7, where $\alpha = 1/\zeta$ and $\beta = \lambda/\xi$ according to section 5.3.1.

Table 5.7: Estimated parameters of gamma distribution for FTSEMIB, S&P500 and FTSE100, 2008-2019.

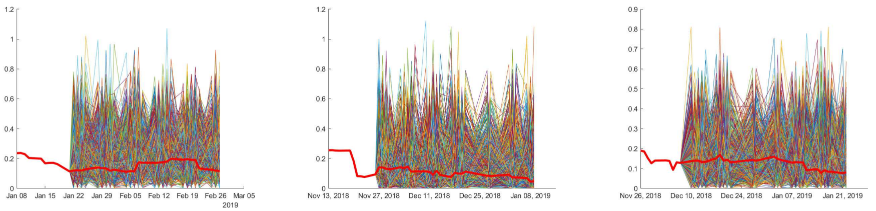
	α	β
FTSEMIB	5.4682	0.0408
S&P500	2.2008	0.0666
FTSE100	4.0794	0.0374

Therefore, to simulate new volatilities referring to model 5.15 we need to discretize the time interval $[0, T]$ by homogeneous time intervals of length $\triangle > 0$. Then an Euler approximation gives,

$$\begin{aligned} V(t) &= \exp\left(-\xi\Delta\right)V(t-\Delta) + \int_{t-\Delta}^{t} \exp\left(-\xi(t-s)\right)dL(s) \\ &\approx \exp\left(-\xi\Delta\right)V(t-\Delta) + \exp\left(-\xi\Delta\right)\Delta L(t) \end{aligned}$$

where $\Delta L(t) = L(t) - L(t-\Delta)$

In Figure 5.15 is displayed the simulated volatility by using the methodology described above:



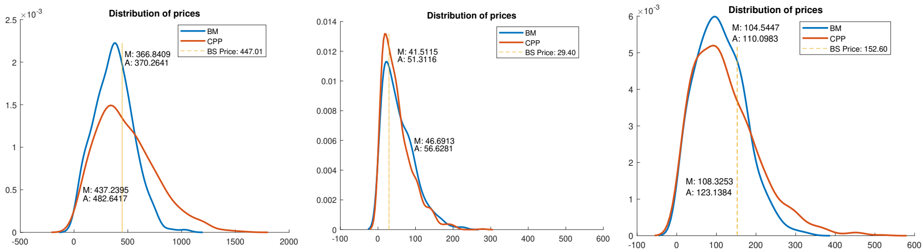
a. FTSEMIB b. S&P500 c. FTSE100

Figure 5.15: Simulated volatility paths based on CPP OU-process.

As one sees in this case we will have just positive values of volatility. The red line indicates the real values of volatility while the other trajectories are the simulated ones. Note that the variability here is seemingly very big but is based on historical variations.

5.3.4 Option Pricing by the Black-Scholes Formula with Uncertain Volatility

Let us consider the case of pricing an at-the-money (ATM) call option when the volatility is uncertain, as we discussed in Section 5.2.1 (see equation 5.5). We fix the time to maturity to be one month for illustration (the date under study will be $19th$ February, 2019), and consider the Gaussian and non-Gaussian Ornstein-Uhlenbeck models for the volatility empirically analyzed in the previous Subsections. The risk-free interest rates used in calculations are collected from [25], [26] and [27] as the reported value of December 2018, for each index, respectively.



a. FTSEMIB　　b. S&P500　　c. FTSE100

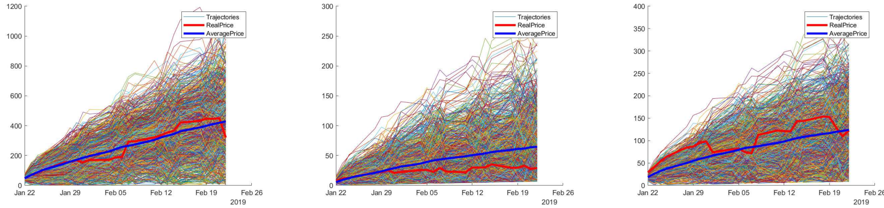
Figure 5.16: Distribution of prices by simulated volatilities using OU-processes, Gaussian (blue) and non-Gaussian (red). (Note that M indicates median while A indicates the average.)

Referring to the two methods we used to simulate volatilities, in Figure 5.16 one sees the distribution of prices calculated on simulated volatilities using an OU process considering BM which allows for negative values and considering the Levy process which we assumed to be a compound Poisson process with exponentially distributed jumps. Referring to the results we obtained for FTSEMIB and FTSE100 the non-Gaussian model gives more spread out prices than the Gaussian ones, while in case of S&P500 the Gaussian and non-Gaussian models perform almost in the same way. In case of FTSEMIB the non-Gaussian gives much more spread out prices. We have included the B&S price based on the historical volatility from the complete set of log-return data as a vertical line in each plot as a reference point. Interestingly, the uncertainty in the volatility yields for all three indices a large variation in option prices around this reference point, with median and average relatively far away. Remark that the negative prices shown in Figure 5.16 in reality are an effect of the kernel smoother used in depicting the distributions.

An illustration on how the price changes in each date until time to maturity, for each index and model applied is done below. Figures 5.17 and 5.18 show how the prices are changing in different trajectories and how the "real prices" (B&S prices based on historical volatility, red line) differ from the average price obtained from these trajectories (blue line). Please recall that the prices are one month ahead contracts.

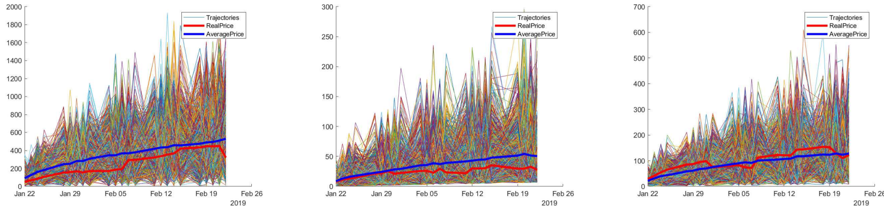
5.4 Model Uncertainty

Cont (2006) and Avellaneda et al (1995) propose to handle model risk by a worst case approach: Let C be a set of contingent claims. Given that a payoff $X \in C$ has a well-defined value in all pricing models $Q \in \mathbb{Q}$, where $\mathbb{Q}$ is the set of equivalent martingale measures, one can define an upper and lower bound for the price of X by $\overline{\pi}(X) = \sup_{i=1..n} E^{Q_i}[X]$ and $\underline{\pi}(X) = \inf_{i=1..n} E^{Q_i}[X]$, respectively. This definition clearly quantifies the extremes of model uncertainty, which is the main



a. FTSEMIB b. S&P500 c. FTSE100

Figure 5.17: Distribution of prices referring to the simulated volatility by using OU process considering BM (negative values of volatilities are generated in this case).



a. FTSEMIB b. S&P500 c. FTSE100

Figure 5.18: Distribution of prices referring to the simulated volatility by using an OU process with CPP.

purpose of Cont (2006). The model uncertainty measure is represented by

$$\mu(X) = \overline{\pi}(X) - \underline{\pi}(X).$$

In our study in Subsection 5.3.4 above, we have simulated 1000 different volatility samples which resulted in 1000 different call option prices in order to analyze the implied distribution from model uncertainty (recall Figure 5.16).

Table 5.8: Model uncertainty ($\mu(X)$) for the call option.

	Gaussian OU	non-Gaussian OU
FTSEMIB	0.105	0.158
S&P500	0.231	0.274
FTSE100	0.327	0.509

According to Avellaneda et al (1995) the upper bound $\overline{\pi}(X)$ is the lowest price that can be charged for the derivative such that, by following an appropriate hedging strategy, the seller can be sure to avoid making a loss on hedging (e.g. the so-called super-hedging strategy (see Bayraktar & Zhou (2017))). The lower bound, on the other hand, is the highest price that can be paid for the derivative such that, by following an appropriate hedging strategy, the buyer can be sure to avoid making a loss on hedging.

Table 5.8 presents how the model uncertainty differs from one model to another for each index. In the table, OU-BM and OU-CPP refer to the Gaussian and non-Gaussian Ornsten-Uhlenbeck models for the volatility, respectively. We observe the smallest model uncertainty values corresponding to the Gaussian Ornstein-Uhlenbeck model, which is in accordance with Figure 5.16.

Rather than using the worst case upper and lower prices to asses model uncertainty, one may ask if it makes sense to calculate the upper and lower bounds of prices using a Value-at-Risk (VaR) measure instead? Since we know that, for example, with 1,000 trials the 1%VaR (1 percentile) is the $10th$ worst case then we will use this methodology to try to define the upper and lower bounds in order to calculate model uncertainty. This means that we calculate the $\alpha\%$-VaR for each trajectory in the date we are interested in and from this vector of VaR values we choose the lower and upper bounds as the corresponding values of simulated volatilities in the trajectories where we find the minimal and maximal values of VaR. Hence, we will calculate $\overline{\pi}_\alpha(X) = BS(S, K, r, T, \overline{s})$ and $\underline{\pi}_\alpha(X) = BS(S, K, r, T, \underline{s})$, where $\overline{s} = \Sigma_{max\{\alpha\%-VaR_i, i=1..n\}}$ and $\underline{s} = \Sigma_{min\{\alpha\%-VaR_i, i=1..n\}}$, $BS(\cdot)$ is the Black & Scholes formula in (5.5), and n is number of simulations. Consequently, in this case the model uncertainty will be represented by

$$\mu_\alpha(X) = \overline{\pi}_\alpha(X) - \underline{\pi}_\alpha(X).$$

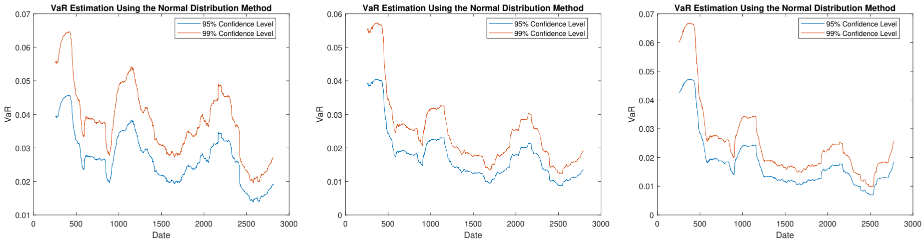
To apply this approach we first need to understand which of the VaR-methods is the most proper one for our data. In the sections below we will see results from *three* methods of calculating VaR and a Backtesting VaR which will instruct us which method we have to use to continue with the corresponding calculations of our proposal. Notice that in case of estimating the model uncertainty of a portfolio of indices it is suggested to substitute VaR by Expected Shortfall since VaR is not sub-additive or convex, and it can lead to anomalous values for a portfolio of options (for more see Artzner, Delbaen, Eber & Heath (1999) and Hull et al (2006)).

5.4.1 VaR Estimation of Log-Returns

In this section, we present the analysis and results for all three indices based on three most common VaR calculation methods focusing on strengths and weaknesses of each method (for more see Nieppola et al (2009) and Danielsson et al (2011)).

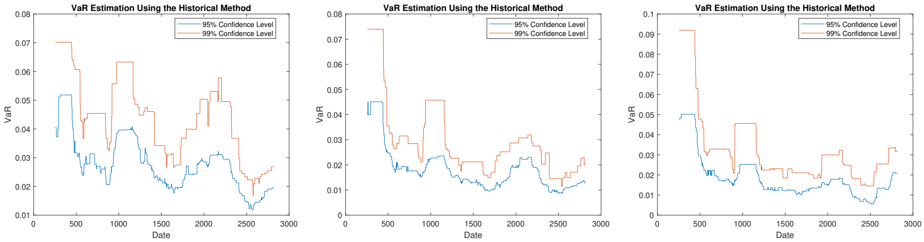
Using the *normal distribution method*, the profit and loss of a stock return is assumed to be normally distributed. Under this assumption, VaR is computed by multiplying the z-score, at each confidence level by the standard deviation of the returns. Because VaR backtesting looks retrospectively at data, the VaR "today" is computed based on values of the returns in the last $N = 252$ days (e.g., one year of daily data) leading to, but not including today. In this case it is assumed that all past returns carry the same weight.

Unlike the normal distribution method, *historical simulation* is a non-parametric method. It does not assume a particular distribution of the asset returns. Historical simulation forecasts risk by assuming that past profits and losses can be used as the distribution of profits and losses for the next period of returns. The VaR "today" is computed as the p^{th} - quantile of the last N returns prior to "today".



a. FTSEMIB b. S&P500 c. FTSE100

Figure 5.19: VaR estimation using the normal distribution method for all indices.



a. FTSEMIB b. S&P500 c. FTSE100

Figure 5.20: VaR estimation using the historical simulation method for all indices.

Figure 5.20 shows that the historical simulation curve has a piece-wise constant profile. The reason for this is that quantiles do not change for several days until extreme events occur. Thus, the historical simulation method is slow to react to changes in volatility.

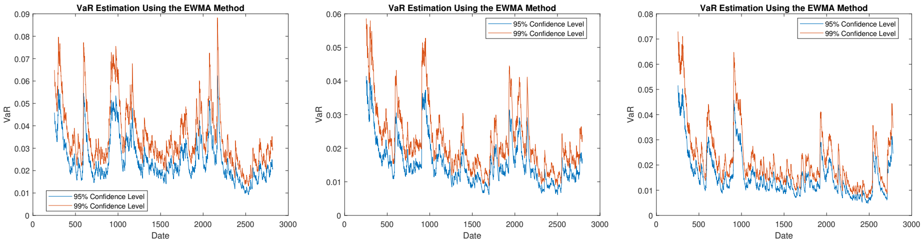
The *exponential weighted moving average (EWMA) method* assigns non-equal weights, particularly exponentially decreasing weights. The most recent returns have higher weights because they influence "today's" return more heavily than returns further in the past. The variance referring to the EWMA method over an estimation window of size W and weight $0 < \lambda < 1$ is

$$\sigma^2 := \frac{1}{c} \sum_{k=1}^{W} \lambda^{k-1} Y_{t-k}^2$$

where c is the normalizing constant

$$c := \sum_{k=1}^{W} \lambda^{k-1} = \frac{1-\lambda^W}{1-\lambda} \overset{W\to\infty}{\longrightarrow} \frac{1}{1-\lambda}.$$

In the Figure 5.21, the EWMA reacts very quickly to periods of large (or small) returns and it can account for fat tails of the return distributions. Note that the value of λ in case of EWMA is chosen to be 0.94.



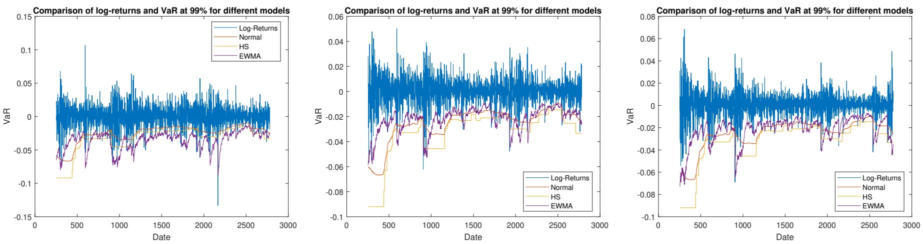
a. FTSEMIB b. S&P500 c. FTSE100

Figure 5.21: VaR estimation using the EWMA method for all indices.

5.4.2 VaR Backtesting

In backtesting, VaR is estimated over the test window with three different methods and at two different VaR confidence levels. The goal of VaR backtesting is of course to evaluate the performance of the three VaR models. A VaR estimate at 99% confidence should be violated only about 1% of the time, and VaR failures should not cluster. Clustering of VaR failures indicates the lack of independence across time because the VaR models are slow to react to changing market conditions.

A common first step in VaR backtesting analysis is to plot the volatility of returns and the VaR estimates together. We plot all three methods at the 99% confidence level and compare them to the volatility of returns.

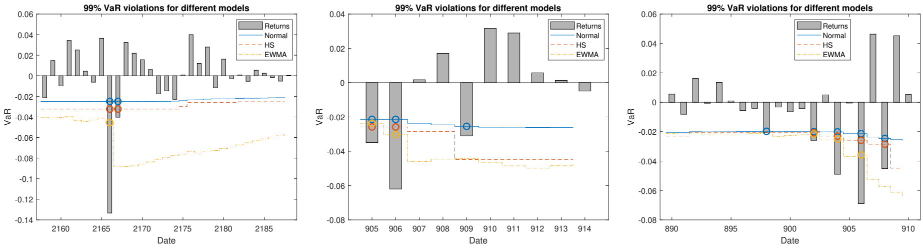


a. FTSEMIB b. S&P500 c. FTSE100

Figure 5.22: Comparisons of log-returns and VaR at 99% with different models.

To highlight how the various approaches react differently to changing market conditions, we have zoomed in (see Figure 5.23) on the time series where there is a large and sudden change in the value of returns. For example $2166th$ date which corresponds to "24 June 2016" for FTSEMIB, $906th$ date which corresponds to "6 June 2011" for S&P500 and FTSE100.

A VaR failure or violation happens when the returns have a negative VaR. A closer look around specified dates for each index shows a significant dip in the returns. It is seen that the EWMA follows the trend of the log-returns closely and more accurately than the other approaches. Consequently, EWMA has fewer VaR violations compared to the normal distribution and historical simulation approaches. As a result, we are going to use EWMA method to calculate VaR for different scenarios of volatilities and



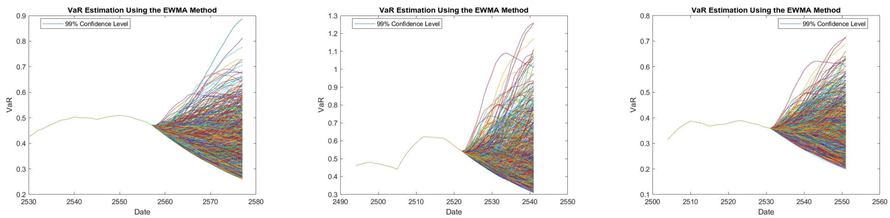
a. FTSEMIB b. S&P500 c. FTSE100

Figure 5.23: VaR violations with different models and for all indices.

to highlight the upper and lower price bounds for model uncertainty.

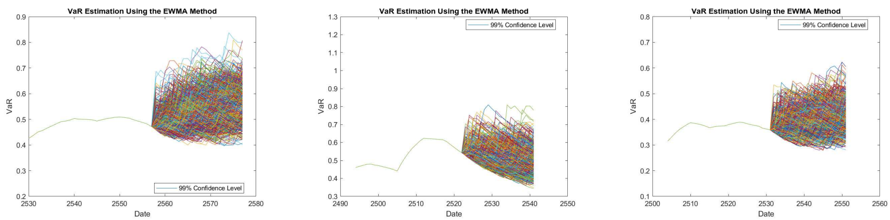
5.4.3 Sensitivity of Model Uncertainty by Using VaR Approach

Now let us apply EWMA approach in 1-month ATM call option prices. In our study we have chosen to deal with a daily 1% VaR and $\lambda = 0.94$. In Figures 5.24 and 5.25 one sees the VaR values calculated for 1000 scenarios of volatilities simulated from the Gaussian and non-Gaussian Ornstein-Uhlenbeck processes, respectively.



a. FTSEMIB b. S&P500 c. FTSE100

Figure 5.24: VaR scenarios for volatilities simulated by Gaussian OU-process.



a. FTSEMIB b. S&P500 c. FTSE100

Figure 5.25: VaR scenarios for volatilities simulated by non-Gaussian Ornstein-Uhlenbeck process.

As a result, for Gaussian and non-Gaussian Ornstein-Uhlenbeck processes the boundaries founded include the real value of volatility. It is clear from the Figures 5.24

and 5.25 that the non-Gaussian Ornstein-Uhlenbeck process define a tighter interval of pricing values

Table 5.9: Model uncertainty ($\mu_\alpha(X)$) for the call options by using VaR.

	Gaussian OU	non-Gaussian OU
FTSEMIB	0.095	0.029
S&P500	0.180	0.069
FTSE100	0.275	0.115

From Table 5.9 it is clear to see that the results on model uncertainty corresponding to the non-Gaussian Ornstein-Uhlenbeck process are the smallest ones. Furthermore, comparing with the results in Table 5.8, we see a clear reduction in model uncertainty using the VaR approach compared to the extreme measure of Cont (2006). This is particularly evident for the non-Gaussian model, where apparently the extreme tails are ignored using the VaR methodology.

5.5 Decision Rules

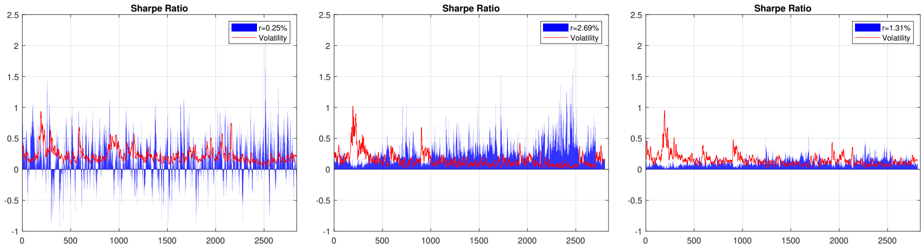
In practice, investors, regulators and risk managers need some decision rules for judging whether an investment has a good performance or has a model uncertainty and model risk value that is too high. In this Section we apply our fitted mean log-return $X(t)$ and volatility $\sigma(t)$ from all three stock indices, to analyse the forecasted risk-adjusted excess return, known as the Sharpe ratio:

$$Sh(t) = \frac{X(t) - r}{\sigma(t)}$$

Here, $r > 0$ is some benchmark return, for example the risk-free one. The risk-free rate is different for each index and it is collected from [**RfRItaly**], [**RfRUSA**] and [**RfRUK**] as the reported value of December 2018.

The Sharpe ratio indicates how well an investment performs in comparison to the rate of return on a risk-free investment. Usually, any Sharpe ratio from 0.5 to 1 is considered acceptable to good by investors. A ratio higher than 2 is rated as very good, and a ratio of 3 or higher is considered excellent (for more see Bayley, & López de Prado (2012)). When analyzing the Sharpe ratio, the higher the value, the more excess return investors can expect to receive for the extra volatility they are exposed to by holding a riskier asset. Similarly, a risk-free asset with no excess return would have a Sharpe ratio of zero. As it is shown in the Figure 5.26, in general, lower volatility results in a better (higher) Sharpe ratio.

It is clear to see that on February 2019, risk-adjusted returns are negative for FTSEMIB and close to zero for S&P500 and FTSE100. A negative Sharpe ratio is hard to interpret. An investor wants to increase a positive Sharpe ratio, by increasing returns and decreasing volatility. However, a negative Sharpe ratio can be brought closer to zero by either increasing returns (something recommended to do) or increasing volatility

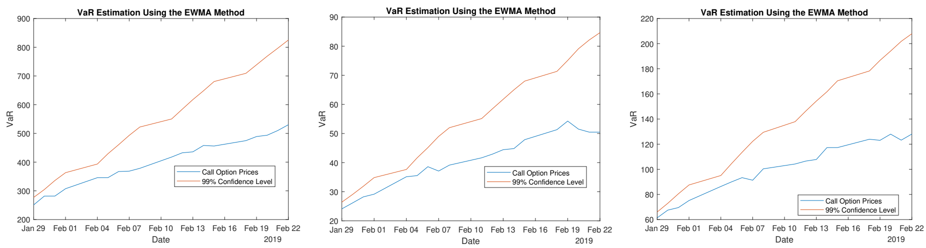


a. FTSEMIB b. S&P500 c. FTSE100

Figure 5.26: Sharpe ratio for all indices.

(something highly not recommended to do). Thus, the Sharpe ratio is not a particularly useful for negative returns. The average Sharpe ratio using calendar daily returns over the considered period (see Figure 5.26) is 0.2, 0.3, and 0.1 for FTSEMIB, S&P500, and FTSE100, respectively. As a summary, we focus on the following investment decision problem under model uncertainty: Reject the investment with payoff X if $Sh(X) < s$, i.e., do not invest in X if the Sharpe ratio is less than s. As we explained above, a reasonable threshold is $s = 0.5$.

Regarding to model risk, we are going to concentrate to the VaR measure which is used by risk managers to measure and control the level of risk undertaken and to ensure it is within limits.



a. FTSEMIB b. S&P500 c. FTSE100

Figure 5.27: VaR measure for ATM call option using EWMA method.

For example, considering Figure 5.27, FTSEMIB index has a one-day 1%VaR of 709.20 EUR, S&P500 index as a one-day 1%VaR of 71.35 USD, and FTSE100 index as a one-day 1%VaR of 178.26 GBP for $19th$ February, 2019. So, based on VaR only, it is suggested to make the following decision: Reject the investment X if $\alpha\% VaR(X) > t_\alpha$, as a consequence do not buy X if the lowest/ smallest loss of X in the worst $\alpha\%$ of scenarios is greater than t_α. In the same way we can construct a decision rule for the model uncertainty measure $\mu(X)$, such as reject X if $\mu(X) > t_\mu$, i.e., do not buy X if the model uncertainty $\mu(X)$ is greater than t_μ. A large value of model uncertainty will result in a significant implication, such as largest effect on losses. The aim of regulators is to minimize the model uncertainty. The value of t_μ should depend on the investor's risk preferences or there could be industry-standard values set by regulators. The obvious generalisation is to aggregate the performance, market risk and model uncertainty with a combined rule such as reject X if $Sh(X) + \alpha\%VaR(X) + \mu(X) > t_M$ for some t_M or as it is suggested by Deng, Dulaney, McCann & Wang (2013) instead of considering VaR and Sharpe ratio separately we can use a VaR Adjusted Sharpe ratio $\alpha\%VaRSh(X) + \mu(X) > t_M$. As a result we notice that, if $Sh(X) + \alpha\%VaR(X) + \mu(X) < 0$ or $\alpha\%VaRSh(X) + \mu(X) < 0$, we would always buy X, since positive returns are made in all possible combinations of market scenarios and models.

5.6 Conclusion

We have proposed a dynamical model of the probability distribution of random events adding uncertainty in the distribution. Further we have defined and discussed model uncertainty and risk measures in a stochastic manner illustrated by various examples applied to stock market indices such as FTSEMIB, S&P500 and FTSE100. It is demonstrated that model and parameter risk and uncertainty play a prominent role. In our study the notion of model uncertainty it is reduced to uncertainty on future volatility. We have considered and compared the results of two different classes of autoregressive models where the so-called stationary distribution is and is not normal (where the OU-process considered is driven by the Brownian motion and Levy process being a compound Poisson process, respectively). The uncertanty on volatility is applied to option pricing, where a study of ATM call options on the three indices imply a distribution of prices rather than a unique price. This is a quantification of the volatility uncertainty in terms of its implied option prices. Next we have implemented the quantitative framework for measuring the impact of model uncertainty on pricing options proposed by Cont et al (2006). We have compared the uncertainty measure based on worst cases with our proposal using a VaR-based measure giving tighter estimates. Finally, some decision rules were discussed referring to VaR, model uncertainty and the Sharpe ratio.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Appendix A. Estimation of the residual process

After we first run a linear regression on log-returns on a rolling window to get the residuals ϵ_t, we define an auxiliary process e_k

$$e_k = \sum_{j=1}^{k} \epsilon_j, k = 1, 2, ...N$$

where N is the length of the residuals. We view e_k as a discrete version of $E(t)$ being an OU-pricess as defined in equation (5.11).A regression estimation of e_k is run to obtain parameters α, μ, β of the Ornstein-Uhlenbeck processes for the respective two cases (see Table 5.10), based on the method of (Avellaneda & Lee (2010)).

Table 5.10: Parameters of OU-process for mean and variance of residuals of FTSEMIB, S&P500 and FTSE100, 2008-2019.

Mean	α	μ	β	β_{eq}	z_score
FTSEMIB	0.3283	-0.0394	0.4888	0.6033	0.0652
S&P500	0.2271	-0.0322	0.5219	0.7744	0.0416
FTSE100	0.3747	-0.0130	0.4219	0.4874	0.0266
Var	α	μ	β	β_{eq}	z_score
FTSEMIB	0.2525	-0.0026	0.1012	0.1425	0.0183
S&P500	0.2271	-0.0322	0.5219	0.7744	0.0416
FTSE100	0.1906	-0.0015	0.0766	0.1241	0.0124

The parameter β_{eq} in Table 5.10 is the standard deviation of the limiting Gaussian distribution and z-score is the standardized version of $E_X(t)$[2] and $E_Y(t)$[3] measuring how far $E_X(t)$ or $E_Y(t)$ deviate from their mean level. This measure is valid measure across all securities being dimensionless and it is used as a trading signal (the idea is that, depending on the strength of the mean-reversion signal (the value of the z-score), we decide on day N to buy/sell at tomorrow's open (or close an existing position)).

Fast mean-reversion (i.e. compared to the 365-day time horizon (estimation sample requires that $\alpha > 252/365$) requires $\alpha > 0.0895$ for FTSEMIB, $\alpha > 0.09071$ for S&P500, $\alpha > 0.0904$ for FTSE100. In all cases, $0 < b_r$[4] < 1 and the above formulas make sense otherwise the mean-reversion time is too long and the model is rejected for the stock under consideration. For all our cases b_r are inside the allowed interval.

[2] $E_X(t)$ indicate the regression equation on the residuals of mean

[3] $E_Y(t)$ indicate the regression equation on the residuals of variance

[4] b_r is the coefficient of the linear regression on the accumulated sum of residuals of AR(1) processes to obtain parameters of the residuals OU-process and z_score.

Bibliography

[1] Abdymomunov, A., Blei, S. and Ergashev, B. (2011). Worst-case scenarios as a stress testing tool for risk models. *SSRN E-print* 10.2139/ssrn.1867365.

[2] Artzner, P., Delbaen, F., Eber, J. and Heath, D. (1999). Coherent measures of risk. *Math. Finance*, **9**(3), pp. 203–228. (https://doi.org/10.1111/1467-9965.00068)

[3] Avellaneda, M. (1997). An introduction to option pricing and the mathematical theory of risk. *Seminario Mat. e. Fis. di Milano*, **67**(1), pp. 225–250. (https://doi.org/10.1007/BF02930502)

[4] Avellaneda, M. and Lee, J. H. (2010). Statistical arbitrage in the US equities market. *Quantit. Finance*, **10**(7), pp. 761–782. (DOI: 10.1080/14697680903124632)

[5] Avellaneda, M., Levy, A. and Paras, A. (1995). Pricing and hedging derivative securities in markets with uncertain volatilities. *Appl. Math. Finance*, **2**, pp. 73—88. (https://doi.org/10.1080/13504869500000005)

[6] Baguley, T. (2012). *Serious stats: A guide to advanced statistics for the behavioral sciences*. Macmillan International Higher Education.

[7] Bannör, K. F. and Scherer, M. (2013). Capturing parameter risk with convex risk measures. *Europ. Actuarial J.*, **3**(1), pp. 97–132. (DOI: 10.1007/s13385-013-0070-z)

[8] Bannör, K. F. and Scherer, M. (2014). Model risk and uncertainty illustrated with examples from mathematical finance. In: C. Klüppelberg , D. Straub, I. Welpe (eds.) *Risk - A Multidisciplinary Introduction*. Springer, Cham. pp. 279–306. (https://doi.org/10.1007/978-3-319-04486-6_10)

[9] Barndorff-Nielsen, O. E. and Shephard, N. (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *J. Royal Statist. Soc.*, Series B (Statist. Meth.), **63**(2), pp. 167–241. (https://www.jstor.org/stable/2680596)

[10] Bayley, D. and López de Prado, M. (2012). The Sharpe ratio efficient frontier, *J. Risk*, **15**(2), pp. 3–44.

[11] Bayraktar, E., & Zhou, Z. (2017). Super-hedging American options with semi-static trading strategies under model uncertainty. *International Journal of Theoretical and Applied Finance*, **20**(06), 1750036.

[12] Bennett, C. and Gil, M. A. (2012). Measuring historical volatility. Research market paper. *Santander Investment Bolsa, Sociedad de Valores, S.A.*

[13] Benth, F. E., Šaltytė Benth, J., and Koekebakker, S. (2004). *Stochastic Modeling of Electricity and Related Markets*, World Scientific.

[14] Billio, M., Casarin, R., Ravazzolo, F., & Van Dijk, H. K. (2013). Time-varying combinations of predictive densities using nonlinear filtering. *Journal of Econometrics*, **177**(2), 213–232.

[15] Black, F. and Scholes, M. (1973). The Pricing of Options and Corporate Liabilities, *Journal of Political Economy* **81** (3), pp. 637–654.

[16] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, **31**(3), 307–327.

[17] Cont, R. (2006). Model uncertainty and its impact on the pricing of derivative instruments. *Math. Finance*, **16**(3), pp. 519–547. (https://doi.org/10.1111/j.1467-9965.2006.00281.x)

[18] Danielsson, J. (2011). *Financial Risk Forecasting: The Theory and Practice of Forecasting Market Risk with Implementation in R and Matlab*. Hoboken: John Wiley.

[19] Deng, G., Dulaney, T., McCann C. and Wang, O. (2013). Robust portfolio optimization with Value-at-Risk-adjusted Sharpe ratios. *J. Asset Manag.*, **14**(5), pp. 293–305. (https://doi.org/10.1057/jam.2013.21)

[20] Durbin, J., and Watson, G. S. (1950). Testing for Serial Correlation in Least Squares Regression, I. *Biometrika*. **37**(3–4), 409-–428. (doi:10.1093/biomet/37.3-4.409)

[21] Durbin, J., and Watson, G. S. (1951). "Testing for Serial Correlation in Least Squares Regression, II". *Biometrika*. **38**(1–2), 159-–179. (doi:10.1093/biomet/38.1-2.159)

[22] Ellsberg, D. (1961). Risk, Ambiguity and the savage axioms. *Quarterly Journal of Economics*, 75, pp. 643–669.

[23] Epstein, L. G. and Wang, L. (1995). Uncertainty, risk neutral measures and asset price booms and crashes. *J. Econ. Theory*, **67**(1), pp. 40–80.

[24] Geweke, J., Amisano, G. (2010). Comparing and evaluating Bayesian predictive distributions of asset returns. *International Journal of Forecasting*, **26**(2), 216–230.

[25] Gilboa Schmeidler. (1989). Maxmin expected utility witn non-unique prior. *Journal of mathematical economics*, 18, pp. 141–153.

[26] Gneiting, T., Ranjan, R. (2013). Combining predictive distributions. *Electronic Journal of Statistics*, 7, 1747–1782.

[27] Hoeting, J. A., Madigan, D., Raftery, A. E. and Volinsky, C. T. (1999). Bayesian model averaging: a tutorial. *Stat. Sci.*, **14**(4), pp. 382–417.

[28] Hull, J. (2006). *Risk Management and Financial Institutions*, 4th Edition, Prentice Hall.

[29] Hull, J. (2009). *Options, Futures and Other Derivatives*, 7th Edition, Prentice-Hall.

[30] Jokhadze, V., & Schmidt, W. M. (2020). Measuring model risk in financial risk management and pricing. *International Journal of Theoretical and Applied Finance*, **23**(02), 2050012.

[31] Kerkhof, F. L. J., Melenberg, B. and Schumacher, J. M. (2010). Model risk and regulatory capital. *J. Banking Finance*, **34**(1), pp. 267–279. (https://doi.org/10.1016/j.jbankfin.2009.07.025)

[32] Knight, F. H. (1921). *Risk, Uncertainty, and Profit*. (New York: Houghton Mifflin).

[33] Nieppola, O. (2009). Backtesting Value-at-Risk models, M.S. Thesis, *Helsinki School of Economics*, Finland. Retrieved from https://aaltodoc.aalto.fi/handle/123456789/181

[34] Primiceri, G. E. (2005). Time varying structural vector autoregressions and monetary policy. *The Review of Economic Studies*, **72**(3), 821–852.

[35] Raftery, A. E. (1993). Bayesian model selection in structural equation models. In *Testing Structural Equation Models*, K. Bollen and J. Long (eds.), Newbury Park , CA : Sage, pp. 163–180.

[36] Rigotti, L. and Shannon, C. (2005). Uncertainty and risk in financial markets. *Econometrica*, **73**(1), pp. 203–243. (https://doi.org/10.1111/j.1468-0262.2005.00569.x)

[37] [RfRItaly] Risk-free rate for Italy: http://www.market-risk-premia.com/it.html

[38] [RfRUSA] Risk-free rate for USA: http://www.market-risk-premia.com/us.html

[39] [RfRUK] Risk-free rate for UK: http://www.market-risk-premia.com/gb.html

[40] Schoutens, W. (2003). *Levy processes in finance: Pricing financial derivatives*. John Wiley & Sons Ltd. Bognor Regis.

Chapter 6

An Application of Functional Data Analysis to Forecast Weather Variables.

Joint work with Fred Espen Benth

Abstract

Functional data analysis (FDA) has emerged as a new area of statistical research with a wide range of applications. In this paper, we propose some functional linear models in which both the response and the covariate variables are functions. These models enable to regularize curves observed over a specific time period and predict curves at unobserved period of time. The proposed models are illustrated in the analysis of temperature and wind speed data. The first model use FDA to estimate temperature from climate zones applying functional principal components analysis (FPCA) at all meteorological stations of Lithuania and predicting full log wind speed and annual log wind speed from temperature climate zones resulted from FPCA. The second model predicts log wind speed directly from temperature observations. The validation procedure based on smoothed functional data of log wind speed shows that the proposed models are reliable and can be used for various practical applications.

Keywords: functional data analysis, linear concurrent model, temperature, wind speed, smoothing

6.1 Introduction

So far, the focus has been on modeling and pricing temperature and rainfall derivatives. The aim of this study is to determine if changes in temperature and wind of Lithuanian cities are detectable using methods of functional data analysis (FDA). More concretely, we propose the functional linear model observed in different patterns in order to predict wind speed from temperature data in which both these variables are functions. Predicting wind play a key role in businesses which are sensitive from wind, such as wind farms, transportation and construction companies, electricity producer companies based on wind mills production prediction, pricing wind derivatives and so on. Many research papers are focused in modeling the dynamics of wind speed. For example, Šaltytė Benth and Benth (2010) propose an ARMA time-series model for the wind speed at a single spatial location, and estimate the model with data based on three different wind farm regions in New York. By a comparison between daily average and three-hourly wind speed predictions the authors find that more accurate predictions are obtained from modeling aggregated data directly rather than at the finer time scale.

Šaltytė Benth and Šaltytė (2011) propose a spatial–temporal model for the wind speed (WS) applied on daily WS records from 18 meteorological stations in Lithuania. The model contains seasonality, a higher-order autoregressive component, a variance describing the remaining heteroskedesticity in residuals and is estimated at the single spatial meteorological station independently on spatial correlations. The spatial dependencies are modeled by a Gaussian random field.

Alexandridis and Zapranis (2013) model the dynamics of the wind generating process using a non-parametric non-linear wavelet network which is validated in New York. The proposed methodology is compared against alternative methods, proposed in prior studies. Their results indicate that wavelet networks can model the wind process very well and consequently constitute an accurate and efficient tool for wind derivatives pricing. These authors provide pricing equations for wind futures written on two indices, the cumulative average wind speed index and the Nordix wind speed index. The characteristics of the wind speed process are very similar to the process of daily average temperatures. It is indicated a slight downward trend and seasonality in the mean and variance. In addition the seasonal variance is higher in the winter while it reaches its lower values during the summer period.

Benth and Pircalabu (2018) propose a non-Gaussian Ornstein–Uhlenbeck model for the wind power production index. The model allows for an analytical formula for pricing wind power futures. Generally, the authors find a negative risk premium whose magnitude decreases as the length of the delivery period increases. The result suggests that wind power producers are willing to accept a lower price when selling wind power futures. Moreover, the market price of risk is more volatile for shorter delivery periods and it is argued that this behavior might be related to liquidity aspects and the information contained in short-term weather forecasts, which the proposed model does not incorporate.

Since many traded financial contracts are based on the daily average wind speed index and from recent research it is seen a similarity in the dynamics of temperature and wind, in our paper we have chosen to focus on modeling the dynamics of the daily average wind speed by applying FDA based on temperature data.

FDA is increasingly used in a wide range of fields including weather derivatives as well. In FDA, the data units are functions or curves, where the observed discrete data are converted to functions using various smoothing procedures. These data are then analyzed using traditional statistical methods to extract information from the functions. The key challenge in FDA is to develop effective and well-suited methodologies for retrieving information within and across curves. Ramsay and Silverman (1997, 2002) present the basic principles of FDA. Besse, Cardot and Stephenson (2000) and Aguilera, Ocana and Valderrama (1999) develop autoregressive forecasting models for climatic variations. A considerable effort is being made in order to adapt some standard statistical methods for functional data. For example, the case of principal component analysis by Boente and Fraiman (2000), Dauxois et al (1982), Locantore et al (1999), Pezzulli and Silverman (1993) and Silverman (1996), discriminant analysis by Ferraty and Vieu (2003) and regression by Cardot et al (1999), Cuevas et al (2002), Ferraty and Vieu (2002). Many other interesting examples can be found in Bosq (1991), Brumback and

Rice (1998), Ramsay and Dalzell (1991), and Rice and Silverman (1991). The FDA approach has become the object of an increasing amount of attention on the part of many researchers in recent years and it is applied in different areas of study. Goia, May and Fusai (2010) consider the short-term peak load forecasting for a district-heating system by applying a functional clustering procedure to classify the daily load curves and functional linear regression models for each cluster.

Hardle and Osipenko (2012) use functional principal component analysis of the variation curves of temperature in order to compute the risk premium involved in the spatial derivative price distribution.

Later, Marron, Ramsay, Sangalli and Srivastava (2015) took attention to the concept of phase variability which is present in functional data and why it is important to not ignore it in statistical analysis.

Fan, James and Radchenko (2015) propose functional additive regression (FAR) method, which extends the usual linear regression model involving a functional predictor and a scalar response and uses a penalized least squares optimization approach to deal with high-dimensional problems.

Guo, Zhou, Huang and Härdle (2015) develop a functional data analysis approach to jointly estimate a family of generalized regression quantiles assuming that the generalized regression quantiles share some common features that can be summarized by a small number of principal component functions.

A recent overview on functional regression can be found in Morris (2015). Some more developments of functional linear model are proposed in Kneip, Poß and Sarda (2016) and Brockhaus, Scheipl, Hothorn, and Greven (2015). Kneip, Poß and Sarda (2016) consider functional linear regression, where scalar responses are modeled in dependence of iid random functions. While Brockhaus, Scheipl, Hothorn, and Greven (2015) propose the functional linear array model which is a unified model class for functional regression models including function-on-scalar, scalar-on-function and function-on-function regression.

Wang, Chiou and Müller (2016) provide an overview of FDA, starting with simple statistical notions such as mean and covariance functions, then covering functional principal component analysis, functional linear regression, as well as clustering and classification of functional data.

Yu, Du and Zhang (2020) propose a flexible single-index partially functional linear regression model, which combines single-index model with functional linear regression model, where all the unknown functions are estimated by B-spline approximation.

Boente, Salibian-Barrera and Vena (2020) construct robust estimators for semi-functional linear regression models by combining splines in order to approximate both the functional regression parameter and the nonparametric component with robust regression estimators based on a bounded loss function and a preliminary residual scale estimator.

The functional linear model is an important model for FDA for which there has been a vast literature on its estimation and prediction. Methods of estimating the slope function were studied by, for instance, Cardot et al. (2003), Yao et al. (2005), Crambes et al. (2009), minimax convergence rates of estimation were established by Hall and Horowitz (2007), Cai and Hall (2006) using functional principal components regression.

Usually, in the problem of functional linear model the data $(X_1, Y_1), ..., (X_n, Y_n)$ are observed, where the $X_i^{'}$s are independent and identically distributed as a random function X, defined on an interval I, and the $Y_i^{'}$s are generated by the regression model,

$$Y_i = \alpha + \int_I \beta X_i + \epsilon_i$$

where α is a constant, denoting the intercept in the model, and β is a square integrable function on I, representing the slope function. The majority of attention usually focuses on estimating β, typically by methods based on functional principal components. Note that in our analysis Y will identify wind speed and X identify temperature data.

The functional linear model in our study is used in various ways:

- The functional linear model is used to predict wind speed from climate zone and a functional covariate constructed by removing climate effects from temperature.
- Annual wind speed, a scalar dependent variable, is fitted by using temperature as a functional covariate. Harmonic acceleration roughness in the regression coefficient function is penalized.
- The full wind speed function is fitted by the regressing on the full temperature profile, and various levels of smoothing are used to show the effects of smoothing over both arguments of the regression coefficient function.

For this analysis, the first step is to inspect and smooth the data with the appropriate level of smoothing for the analyses being considered. The idea of a "smart" roughness penalty is introduced right away in the form of harmonic acceleration, which is especially important for periodic data such as these with a variation that is dominated by a sinusoid plus a constant signal, or shifted harmonic variation. However, in order to keep the analysis simple and to economize on computational effort, we use a saturated basis capable of interpolating the data combined with a roughness penalty, and instead opt for a Fourier series basis system with 65 basis functions and no roughness penalty.

Nevertheless, there is a smoothing section below that use 35 Fourier basis functions combined with a harmonic acceleration penalty where we estimate the smoothing parameter by minimizing the generalized cross-validation (GCV) parameter. Smoothing is followed by the display of various descriptive statistics, including mean and standard deviation functions, and covariance and correlation surfaces. A short section illustrates the principal components analysis of temperature, and these analysis is repeated for wind speed as well.

6.2 Functional linear model

We consider the concurrent linear model introduced in Malfait and Ramsay (2003) in order to relate the value of a functional response (log-wind speed $(Y_i(t))$) to the current value of functional covariate(s) (in our case temperature $(T_i(s))$). We have chosen Ω_t to contain the range of values of argument s over which T_i is considered to influence response Y_i at time t, and the subscript t on this set indicates that this set can change from one value of t to another. For example, when both s and t are time, using $T_i(s)$ to predict $Y_i(t)$ when $s > t$ may imply backwards causation. In order to avoid this nonsense, we consider only values of T_i before time t. We may also add a restriction on how far back in time the influence of T_i on Y_i can happen. This leads us to restrict the integral to

$$\Omega_t = \{s \geq T_0 | t - \delta \leq s \leq t\}$$

where T_0 is the initial data and $\delta > 0$ specifies how much history is relevant to the prediction. In our case we have chosen $\delta = 9855$ days corresponding to 27 year historical temperature data. Note that in our analysis we display snapshots from the latest year 2003 in order to have clearer results.

A more general version for a single functional covariate (temperature) and an intercept is

$$Y_i(t) = \beta_0(t) + \int_{\Omega_t} \beta_1(t,s) T_i(s) ds + \epsilon_i(t) \tag{6.1}$$

where i indicates the location/city considered. The bivariate regression coefficient function $\beta_1(s,t)$ defines the dependence of $Y_i(t)$ on covariate $T_i(s)$ at each time t and is defined as follows:

$$\beta_1(s,t) = \sum_{k=1}^{K} \sum_{l=1}^{L} b_{kl} \phi_k(s) \psi_l(t) = \phi^{'}(s) B \psi(t) \tag{6.2}$$

where the coefficients for the expansion are in the K x L matrix B. We therefore need to define two bases $\phi(s)$ and $\psi(s)$ for the coefficient functions: $\phi(s)$ for β_1, as well as $\psi(s)$ for the intercept function β_0. The intercept function β_0 can be expressed as follows:

$$\beta_0(t) = \sum_{l=1}^{L} a_l \psi_l(t) = \psi^{'}(t) a \tag{6.3}$$

For the bivariate regression coefficient we have chosen Fourier basis while for the intercept some constant basis. In this case $T_i(s)$ need not be defined over the same range, or even the same continuum, as $Y_i(t)$.

6.3 Functional smoothing and descriptive statistics

6.3.1 Data Collection

The analysis on the daily temperature and wind speed (WS) (m/s) data for 16 Lithuanian meteorological stations, provided by the Lithuanian Hydrometeorological Service (LHS) [1] in Vilnius, Lithuania, is developed in this study. We have available records of WS starting earlier than 1 January 1977 for 16 meteorological stations for WS and temperature data starting from 1 January 1977 available for some more meteorological stations but in order to have time series of equal length, as a starting point we choose 1 January 1977 and 16 meteorological stations for which we have records for temperature

[1] http://old.meteo.lt/english/

and WS, simultaneously. Data series continue until 31 December 2003, resulting in 9855 daily observations corresponding to 27 years. Note that in order to have clear visualisations of our analysis, in the following sections are shown just snapshots of the last year 2002 taken in analysis and results of year 2013 from the predictive model.

6.3.2 Smoothing functional data of temperature and wind speed

This analysis starts with smoothing raw data of temperature and wind speed using a technique of fitting models to data by minimizing the sum of squared errors. This approach consist on fitting the discrete observations t_j and y_j over time, $j = 1, ..., n$, for temperature and wind speed, respectively, using the following models:

$$t_j = x_1(l_j) + \epsilon_j$$

$$y_j = x_2(s_j) + \eta_j$$

and a basis function expansion for $x_1(l)$ and $x_2(s)$ of the form

$$x_1(l) = \sum_{k=1}^{K} c_k \phi_k(l) = c^{'} \phi$$

$$x_2(s) = \sum_{m=1}^{M} d_m \psi_m(s) = d^{'} \psi$$

where vectors c and d of length K and M contain the coefficients c_k and d_m which determine the expansions and assume that the residuals ϵ_j and η_j about the true curve are independently and identically distributed with mean zero and constant variance σ^2. Let define the n x K matrix Φ as containing the values $\phi_k(l_j)$ and the n x M matrix Ψ as containing the values $\psi_k(s_j)$. Then, a simple linear smoother is obtained if the coefficients of the expansions c_k and d_k are determined by minimizing the least squares criterions

$$\begin{aligned} MLSSE(t, c) &= \sum_{j=1}^{n} (t_j - \sum_{k=1}^{K} c_k \phi_k(l_j))^2 \\ MLSSE(y, d) &= \sum_{j=1}^{n} (y_j - \sum_{m=1}^{M} d_m \psi_m(s_j))^2 \end{aligned} \tag{6.4}$$

which in matrix form are expressed as:

$$\begin{aligned} MLSSE(t, c) &= (t - \Phi c)^{'} (t - \Phi c) \\ MLSSE(y, d) &= (y - \Psi d)^{'} (y - \Psi d) \end{aligned} \tag{6.5}$$

Taking the derivative of criterions $MLSSE(t, c)$ and $MLSSE(y, d)$ with respect to c and d yield the equations

$$2\Phi\Phi^{'} c - \Phi^{'} t = 0$$

$$2\Psi\Psi^{'}d - \Psi^{'}y = 0$$

and solving this for c and d provides the estimators $\widehat{c}$ and $\widehat{d}$ that minimizes the least squares solution,

$$\widehat{c} = (\Phi^{'}\Phi)^{-1}\Phi^{'}t$$

$$\widehat{d} = (\Psi^{'}\Psi)^{-1}\Psi^{'}y$$

The vectors $\widehat{t}$ and $\widehat{y}$ of fitted values are

$$\widehat{t} = \Phi(\Phi^{'}\Phi)^{-1}\Phi^{'}t$$

$$\widehat{y} = \Psi(\Psi^{'}\Psi)^{-1}\Psi^{'}y$$

From equation (6.4) the functional observation for temperature and wind speed are expressed by:

$$x_1(l) = \sum_{k=1}^{K} c_k\phi_k(l)$$

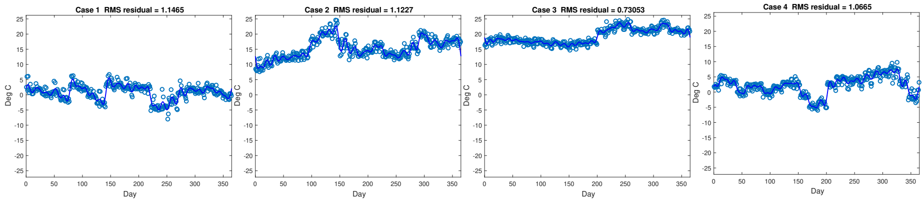
$$x_2(s) = \sum_{m=1}^{M} d_m\psi_m(s)$$

The smoothness of the fit can be controlled by the choice of K and M, which indicates the number of basis functions. The smaller the number of basis functions, the smoother the fit, and the larger the number of basis functions, the closer the fit will be to the data. The basis functions employed in this analysis, are Fourier basis functions since they perfectly represent periodic data. Fourier basis functions are useful for examining annual trends with seasonal variation. The set of basis functions for Fourier series includes one constant function and then pairs of sine and cosine functions to capture the variation in phase (the number of basis must always be odd):

$$\begin{aligned} \phi_1(l) &= 1 & \psi_1(s) &= 1 \\ \phi_2(l) &= sin(l\omega) & \psi_2(s) &= sin(s\omega) \\ \phi_3(l) &= cos(l\omega) & \psi_3(s) &= cos(s\omega) \\ \phi_k(l) &= sin(\frac{k}{2}l\omega) & \psi_m(s) &= sin(\frac{m}{2}s\omega) \\ \phi_{k+1}(l) &= cos(\frac{k}{2}l\omega) & \psi_{m+1}(s) &= cos(\frac{m}{2}s\omega) \end{aligned}$$

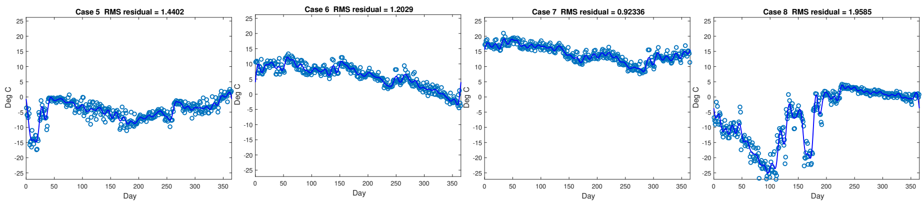
where ϕ_k and ψ_m are the k and mth basis functions and $\omega = 2\pi/T$ where T is the period of the function.

However, as it will be seen from the results, some regularization in the function $x = (x_1(l), x_2(s))$ by attaching to the least squares fitting criterion an additional term that controls the roughness of some derivative of the fit, is needed. Below we start presenting results obtained in case of smoothing functional data by applying the least squares technique. In figures 6.1-6.4, are shown the smoothed temperature functional data for all Lithuanian meteorological stations.



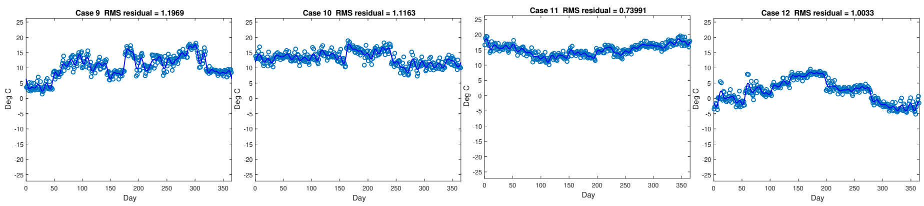
a. Vilnius b. Nida c. Raseiniai d. Kybertai

Figure 6.1: Smoothing functional data by least squares: Temperature curves and values for the Lithuanian weather stations.



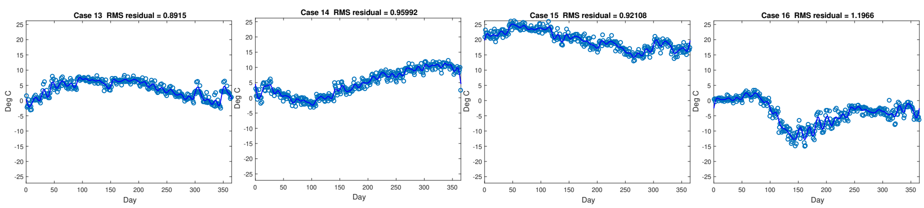
a. Kaunas b. Ukmerge c. Silute d. Birzai

Figure 6.2: Smoothing functional data by least squares: Temperature curves and values for the Lithuanian weather stations.



a. Laukuva b. Telsiai c. Panevezys d. Utena

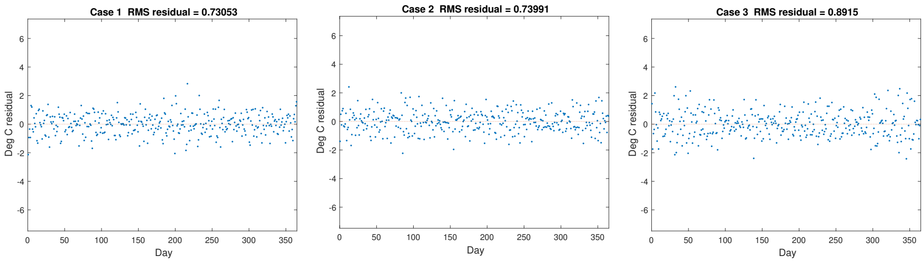
Figure 6.3: Smoothing functional data by least squares: Temperature curves and values for the Lithuanian weather stations.



a. Varena b. Klaipeda c. Siauliai d. Dotnuva

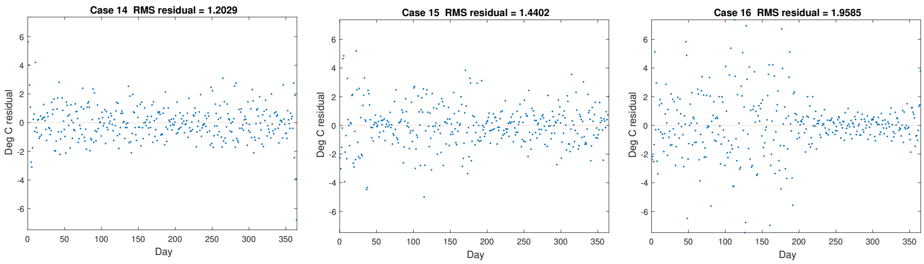
Figure 6.4: Smoothing functional data by least squares: Temperature curves and values for the Lithuanian weather stations.

In order to understand better the variation of residuals for temperature curves, we have chosen to display residuals of the three meteorological stations with the best fits (see figure 6.5) and residuals of the three stations with the worst fits (see figure 6.6).



a. Raseiniai b. Panevezys c. Varena

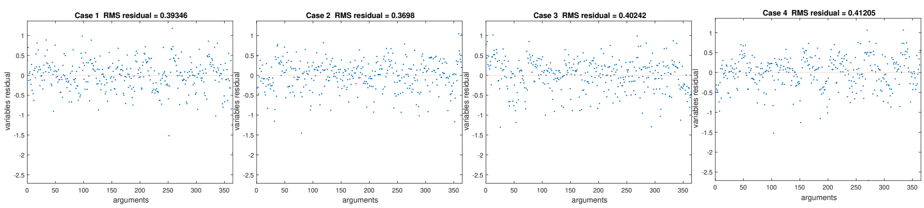
Figure 6.5: Smoothing functional data by least squares: Residuals for three best fits.



a. Ukmerge b. Kaunas c. Birzai

Figure 6.6: Smoothing functional data by least squares: Residuals for three worst fits.

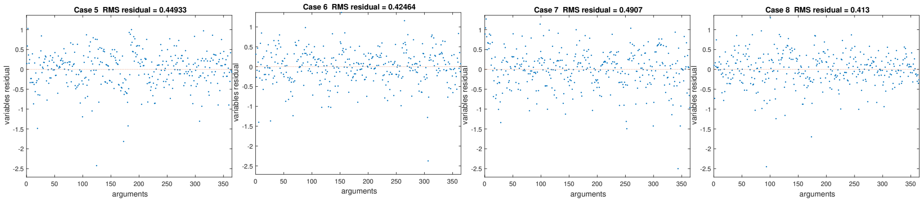
In figures 6.7-6.10, are shown the smoothed wind speed functional data for all Lithuanian meteorological stations. As it is seen the values RMS [2] on residuals are smaller than the ones computed in case of temperature data.



a. Vilnius b. Nida c. Raseiniai d. Kybertai

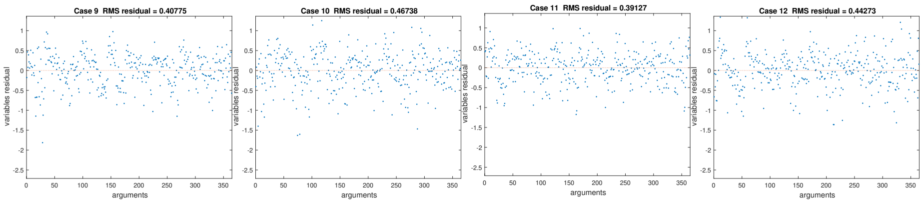
Figure 6.7: Smoothing functional data by least squares: Wind curves and values.

[2]stands for Root mean square error



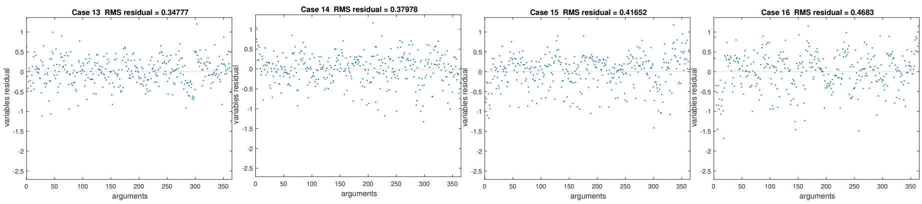
a. Kaunas b. Ukmerge c. Silute d. Birzai

Figure 6.8: Smoothing functional data by least squares: Wind curves and values.



a. Laukuva b. Telsiai c. Panevezys d. Utena

Figure 6.9: Smoothing functional data by least squares: Wind curves and values



a. Varena b. Klaipeda c. Siauliai d. Dotnuva

Figure 6.10: Smoothing functional data by least squares: Wind curves and values.

The functions are definitely too rough with the basis functions chosen, especially for temperature, which has a much higher noise level. These smoothing parameter values probably undersmooth the data, but we can impose further smoothness on the results of our analyses by a regularization approach as mentioned before.

Therefore, the fit to the data vectors t and y in equation (6.4) are regularized by minimizing the criterions

$$\begin{aligned} PENSSE &= MLSSE(t, c) + \lambda_1 * PEN(x_1) \\ PENSSE &= MLSSE(y, d) + \lambda_2 * PEN(x_2) \end{aligned} \tag{6.6}$$

where the second term on the right side penalizes some form of roughness in x. In our case the following criterion is applied:

$$PEN(x) = \int [D^2 x(t)]^2 dt \tag{6.7}$$

which measures the roughness of the function x by integrating the square of its second derivative D^2x, which is the total curvature of x. The smoothing parameter $\lambda = (\lambda_1, \lambda_2)$ is important; the larger λ, the more heavily roughness in x is penalized, and ultimately as λ increases without limit, x is forced towards a straight line, for which the second derivative is everywhere 0. On the other hand, as λ is reduced to zero, the roughness of x matters less and less, and finally when $\lambda \to 0$, x will be just as rough as y since it will pass exactly through the data points. In the following figures one can see the smoothed data and original data for temperature and wind, respectively. The goodness of fit is given by RMS on residuals. This approach permits the direct smoothing of the raw discrete data without any regularization which means λ is zero and the basis of Fourier functions is chosen: about one basis function per week which means for the frame of one year 65 basis fourier functions are applied.

However, even more sophistication in the definition of roughness can be obtained by defining a linear differential operator of the form:

$$Lx = \beta_0 x + \beta_1 Dx + ... + \beta_{m-1} D^{m-1} x + D^m x$$

where the m weight functions $\beta_j, j = 0, ..., m-1$, may be either constants or themselves functions. In our case it is applied the following operator (the degree of the operator is chosen to be 3):

$$Lx = w^2 Dx + D^3 x$$

where the data are periodic with period $2\pi/w$ and we smooth towards a vertically shifted sinusoid. The regularization penalty (6.7) then becomes

$$PEN(x) = \int [Lx(t)]^2 dt \tag{6.8}$$

The reason for considering this wider family of penalties is that by the appropriate choice of L, we can force the smooth as $\lambda \to \infty$ to be toward a linear combination of m functions u_j that we choose (Heckman and Ramsay (2000)). The results of this approach are shown in the figures below. In this case, 65 basis fourier functions are considered and $\lambda = 10$ and $\lambda = 100\ 000$ for temperature and wind speed, respectively.

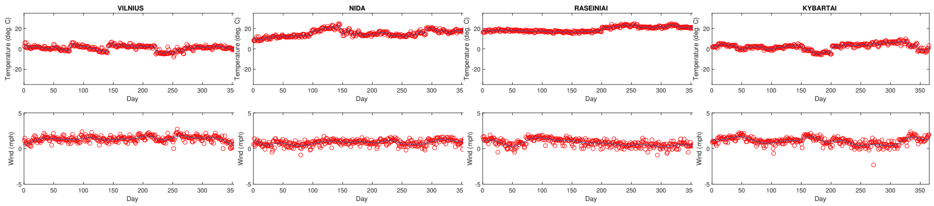


Figure 6.11: Smoothing functional data with a roughness penalty: Each pair of functions along with raw data

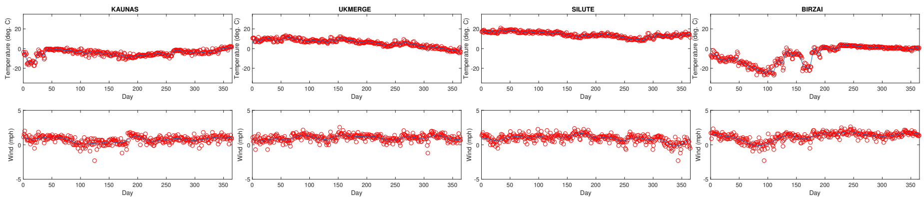


Figure 6.12: Smoothing functional data with a roughness penalty: Each pair of functions along with raw data

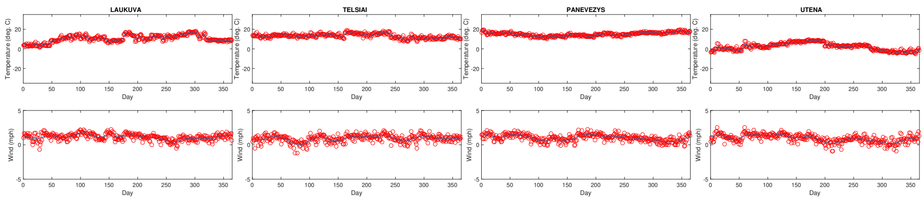


Figure 6.13: Smoothing functional data with a roughness penalty: Each pair of functions along with raw data

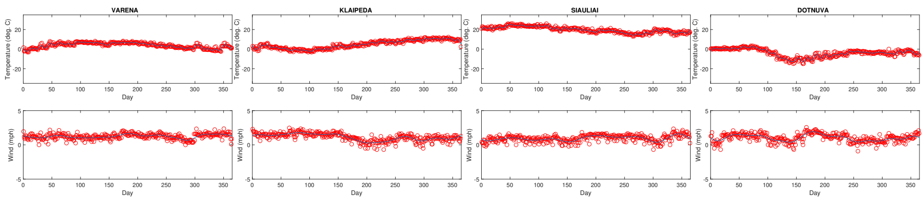


Figure 6.14: Smoothing functional data with a roughness penalty: Each pair of functions along with raw data

The general visualisation of all smoothed temperature and wind speed functional data for all meteorological stations are shown in figure 6.15.

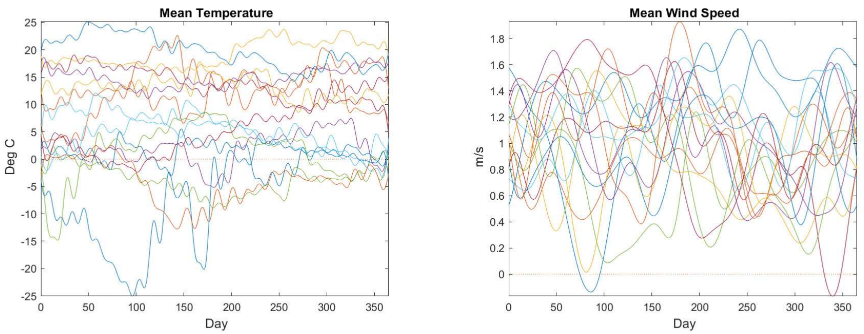


Figure 6.15: All functions: Mean Temperature and Mean Wind Speed in 16 meteorological stations

Note that the behaviour of temperature and wind speed in Lithuania is quite similar in all stations, which is coming likely because of the Lithuania's simple terrain.

When we fit data using a roughness penalty instead of least squares, we switch from defining the smooth in terms of degrees of freedom to defining the smooth in terms of the smoothing parameter λ. In the following section the procedure of selecting the level of smoothing is described by observing the generaized cross-validation criterion.

6.3.2.1 Choosing the level of smoothing using the generalized cross-validation criterion

We choose level of smoothing using the generalized cross-validation criterion with smoothing functional data with a roughness penalty. The generalized cross-validation (GCV) measure developed by Craven and Wahba (1979) is designed to locate the best value for the smoothing parameter λ. The criterion is given as follows:

$$GCV(\lambda) = (\frac{n}{n - df(\lambda)})(\frac{SSE}{n - df(\lambda)})$$

where SSE identifies the sum of squared estimate of errors and df denotes degree of freedom. The right second factor is the unbiased estimate of error variance σ^2, and thus represents some discounting by subtracting $df(\lambda)$ from n. The right first factor further discounts this estimate by multiplying by $n/(n - df(\lambda))$.

The left hand side of Figure 6.16 shows the variation of the generalized cross-validation statistic GCV for temperature over a range of $log10(\lambda)$ values in its top panel. The minimizing value of λ is about 10^0, and at this smoothing level, the degrees of freedom measure has the value of $df(\lambda) = 120$ while $GCV(\lambda) = 30$. The right hand side of Figure 6.16 shows the variation of the generalized cross-validation statistic GCV for wind over a range of $log10(\lambda)$ values in its top panel. The minimizing value of λ is about 10^4, and at this smoothing level, the degrees of freedom measure has the value of $df(\lambda) = 27$ while $GCV(\lambda) = 2.54$.

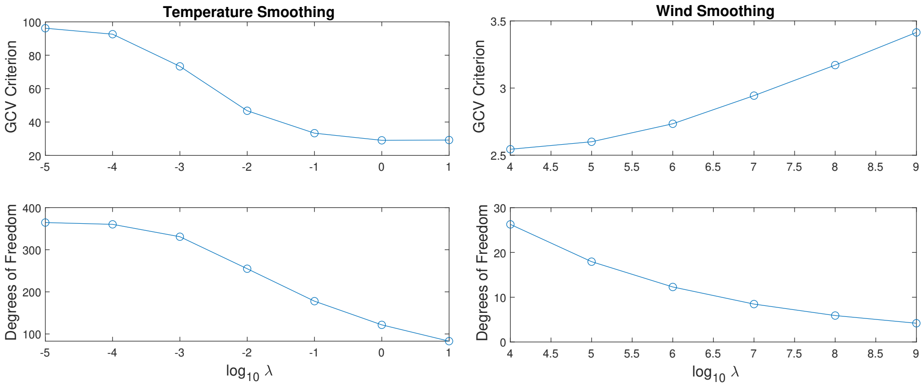
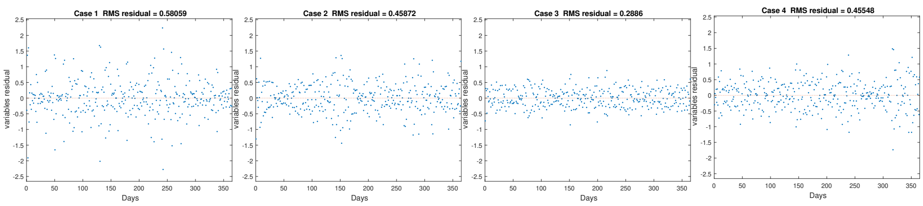


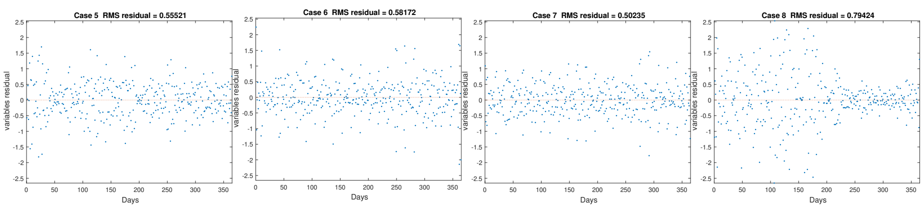
Figure 6.16: Temperature and wind speed: Plot of degrees of freedom and GCV criterion: The top panel displays the relation between the GCV statistic and smoothing level for simulated temperature and wind speed records. The bottom panel displays the degrees of freedom.

By fixing these parameters to be used in the process of generating the smoothed functional curves corresponding to temperature and wind speed, the results changes as it is shown in the following figures. In figures 6.17-6.20 are shown the results corresponding to temperature for all meteorological stations.



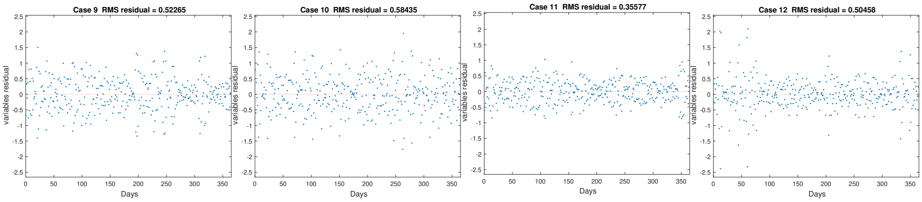
a. Vilnius b. Nida c. Raseiniai d. Kybertai

Figure 6.17: Temperature: Plot of data and fit after applying the final smooth with minimum GCV value.



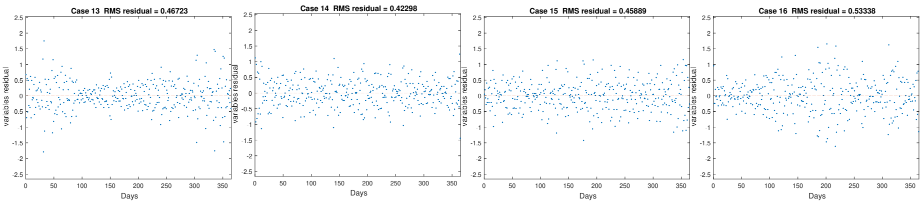
a. Kaunas b. Ukmerge c. Silute d. Birzai

Figure 6.18: Temperature: Plot of data and fit after applying the final smooth with minimum GCV value.



a. Laukuva b. Telsiai c. Panevezys d. Utena

Figure 6.19: Temperature: Plot of data and fit after applying the final smooth with minimum GCV value.

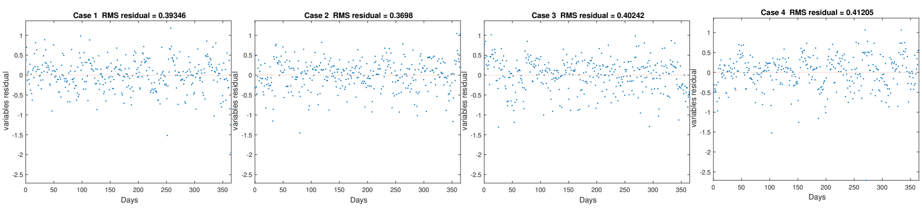


a. Varena b. Klaipeda c. Siauliai d. Dotnuva

Figure 6.20: Temperature: Plot of data and fit after applying the final smooth with minimum GCV value.

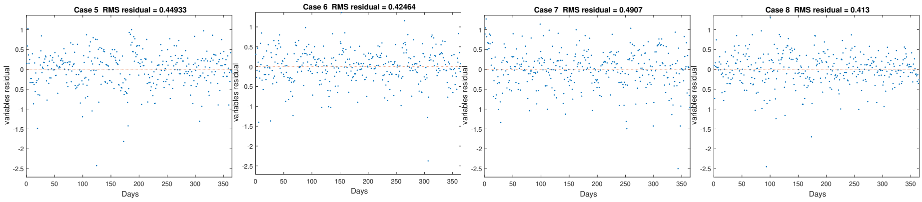
It is clear to see from the pictures that the RMS on residuals is smaller than the one generated by applying the least square technique.

In figures 6.21-6.24 are shown the results corresponding to wind speed for all meteorological stations. It is clear to see that the results generated in this case are improved compared with ones generated by applying the least square technique (RMS errors for all meteorological stations are less than 0.5).



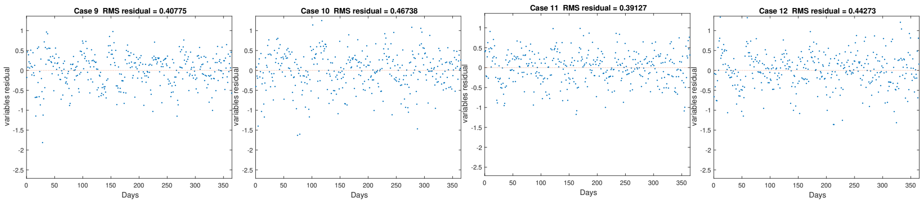
a. Vilnius b. Nida c. Raseiniai d. Kybertai

Figure 6.21: Wind Speed: Plot of data and fit after applying the final smooth with minimum GCV value.



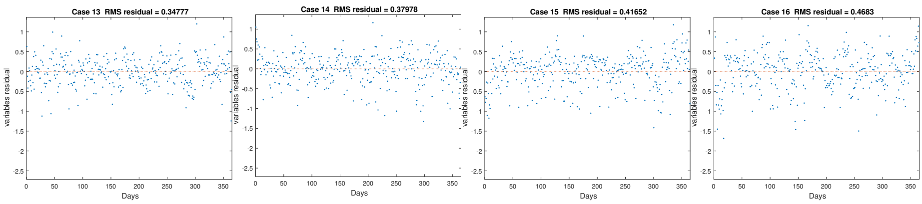
a. Kaunas b. Ukmerge c. Silute d. Birzai

Figure 6.22: Wind Speed: Plot of data and fit after applying the final smooth with minimum GCV value.



a. Laukuva b. Telsiai c. Panevezys d. Utena

Figure 6.23: Wind Speed: Plot of data and fit after applying the final smooth with minimum GCV value.



a. Varena b. Klaipeda c. Siauliai d. Dotnuva

Figure 6.24: Wind Speed: Plot of data and fit after applying the final smooth with minimum GCV value.

Before applying the functional linear model on the smoothed data generated let first see some descriptive statistics about temperature and wind speed.

6.3.3 Descriptive Statistics for temperature and wind speed

We start by computing mean and standard deviation of temperature and log wind speed for all 16 stations of Lithuania. Let $x_{ij}, i = 1, 2$ and $j = 1, ..., 16$ be a sample of functions fit to data. The sample mean and variance functions for each day, are defined as follows:

$$\overline{x}_i(l) = \frac{1}{N} \sum_j x_{ij}(l)$$

and

$$s_i(l) = \frac{1}{N-1} \sum_j (x_{ij} - \overline{x}_i(l))^2$$

The results for temperature and log wind speed are shown in the following figures:

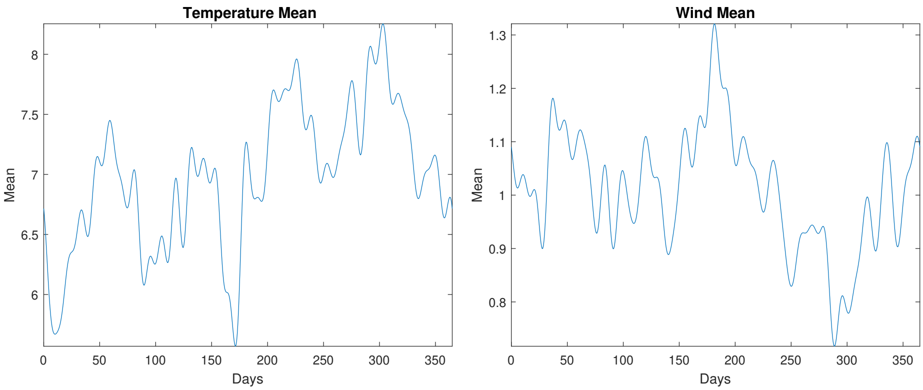


Figure 6.25: Functional mean for temperature and wind speed.

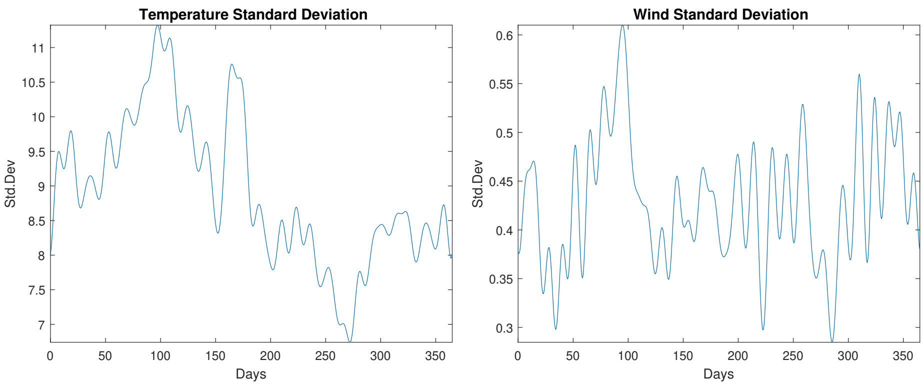


Figure 6.26: Functional standard deviation for temperature and wind speed.

The distribution of wind speed is strongly skewed, and by logging these data, we effectively work with the geometric mean of wind as a more appropriate measure of location in the presence of substantial skewness.

The functional standard deviation focuses on the intrinsic variability between observations, such as Lithuanian weather stations, after removing variations that are believed to represent measurement and replication error not attributable to the variability between observations.

The bivariate covariance function $cov(l, s)$ specifies the covariance between curve values $x_1(l)$ and $x_2(s)$ at times l and s, respectively. It is estimated by

$$cov(l, s) = \frac{1}{N-1} \sum_j (x_{1j}(l) - \overline{x}_1(l))(x_{2j}(s) - \overline{x}_2(s))$$

while functional covariance for temperature or log wind speed, are expressed by the following expression, where $i = 1$ states for temperature, $i = 2$ states for log wind speed.

$$cov_i(l, l) = \frac{1}{N-1} \sum_j (x_{ij}(l) - \overline{x}_i(l))^2$$

The variance-covariance of temperature and log wind speed functions are shown in Figure 6.27 as the height of the diagonal running from $(0, 0)$ to $(53, 53)$ (365 days are translated in weeks, 53 weeks for a year). There is much more variation in wind in the winter months. While regarding temperature the highest values are during the first half of the year.

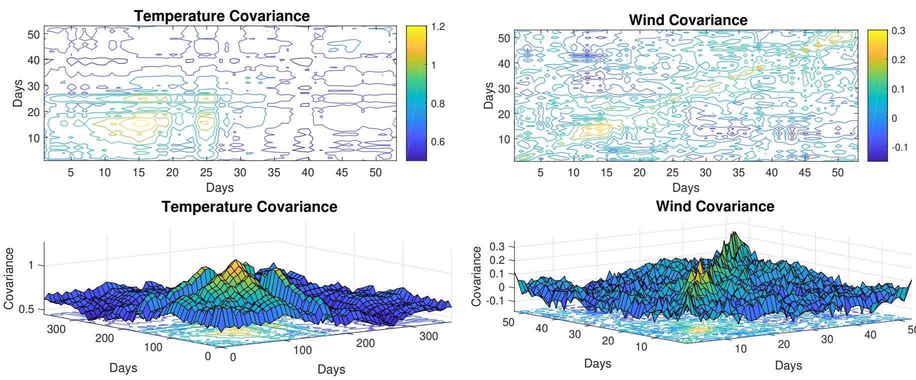


Figure 6.27: Variance-covariance bivariate function for temperature and log wind speed.

The variance-covariance surface shown in Figure 6.27 indicate that variance across weather stations is larger in the winter than it is in the summer.

The bivariate correlation function $corr(l, s)$ specifies the correlation between curve values $x_1(l)$ and $x_2(s)$ at times l and s, respectively. It is estimated by

$$corr(l, s) = \frac{cov(l, s)}{\sqrt{cov(l, l)cov(s, s)}}$$

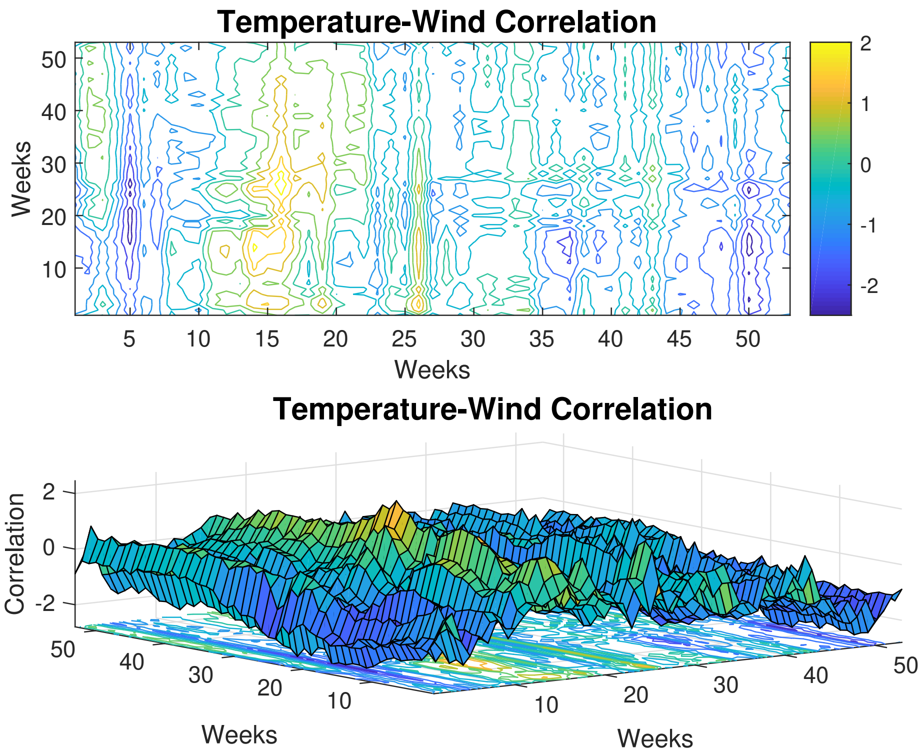


Figure 6.28: Correlation function between temperature and log wind speed.

From figure 6.28 it is seen that there is a correlation structure between temperature and log wind speed.

6.4 Exploring variation by functional principal components analysis (PCA)

Let now look at how observations vary from one replication or sampled value to the next and see what modes of variation are in the data, and how many of them seem to be substantial. As in multivariate statistics, eigenvalues of the bivariate variance-covariance function are indicators of the importance of principal components, and plotting eigenvalues is a method for determining how many principal components are required to produce a reasonable summary of the data. In functional PCA, there is an eigenfunction associated with each eigenvalue, rather than an eigenvector, where eigenfunctions describe major variational components and we apply it on smoothed data (Ramsay and Dalzell (1991); Foutz and Jank (2010)). Applying a rotation to them often results in a more interpretable picture of the dominant modes of variation in the functional data, without changing the total amount of common variation.

The objective in principal components analysis is the orthogonal decomposition of the variance-covariance function to isolate the dominant components of functional variation. We calculate eigenfunctions ψ_j of the bivariate covariance function $cov(s,t)$ as solutions of the functional eigenequation:

$$\int cov(s,t)\psi_j(t)dt = \mu_j\psi_j(s)$$

where the eigenvalues μ_j indicate the amount of variance attributable to each component and ψ_j are known as the principal component functions or harmonics which satisfy the following criteria

$$\int \psi_j(t)\psi_l(t)dt = 0, j = 1, ..., l-1$$

and

$$\int \psi_l^2(t)dt = 1$$

The principal component scores c_{ij} are given by the following:

$$c_{ij} = \int \psi_j(t)[x_{ij}(t) - \overline{x}_i(t)]dt$$

The control of level of fit in data in functional PCA, is done by controlling the roughness of the estimated eigenfunctions (which means by modifying the definition of orthogonality). We penalize excessive curvature in principal components, by using this generalized form of orthogonality (Silverman (1996)):

$$\int \psi_j(t)\psi_l(t) + \lambda \int D^2\psi_j(t)D^2\psi_l(t) = 0, l = 1, ..., j-1$$

where λ controls the relative emphasis on orthogonality of second derivatives. We choose the smoothing parameter λ via cross-validation (see Ramsay and Silverman (2005); Green and Silverman (1994)).

Figure 6.29 shows the two principal component functions by displaying the mean curve along green line and red line indicating the consequences of adding and subtracting a small amount of each principal component. We do this because a principal component represents variation around the mean, and therefore is naturally plotted as such. We observe that these two harmonics account for 94% of the variation around the mean temperature curve. We see that the first harmonic, accounting for 86% of the variation, represents a relative constant vertical shift in the mean, and that the second shows essentially a contrast between winter and summer temperature levels.

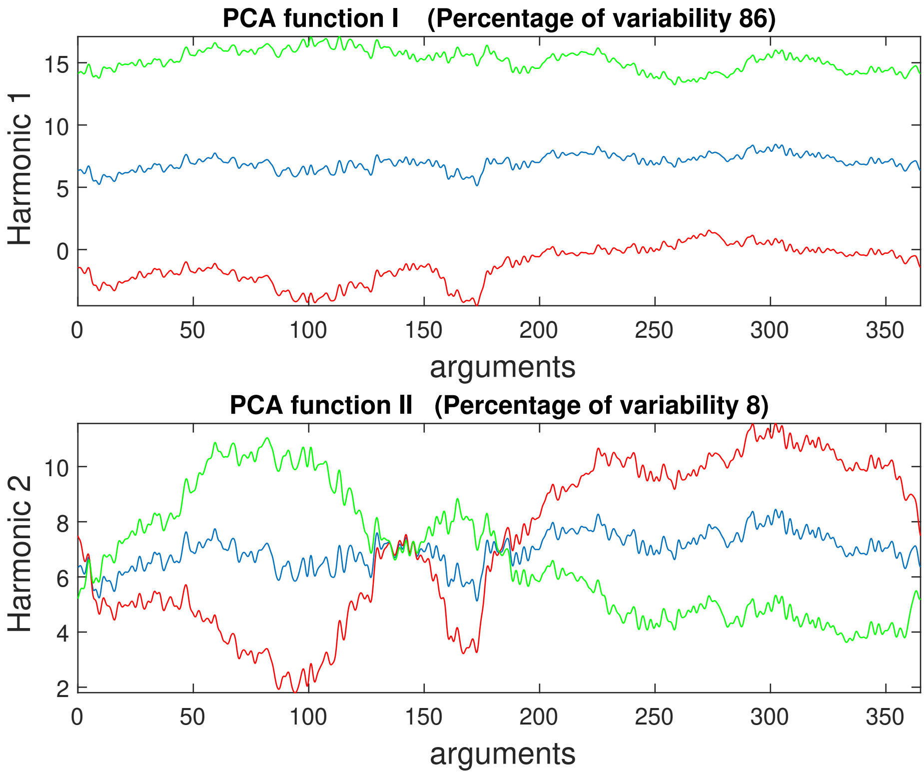


Figure 6.29: Principal components analysis of temperature: The two principal component functions or harmonics are shown as perturbations of the mean, which is the blue line. The green line show what happens when a small amount of a principal component is added to the mean, and the red line show the effect of subtracting the component. The top panel contains the strongest component.

The fact that unrotated functional principal components are so predictable emphasizes the need for looking for a rotation of them that can reveal more meaningful components of variation. The VARIMAX rotation algorithm is often used for this purpose. The results are plotted in Figure 6.30. The first component portrays variation that is strongest in first part of the year and the second captures primarily variation from the second part of the year.

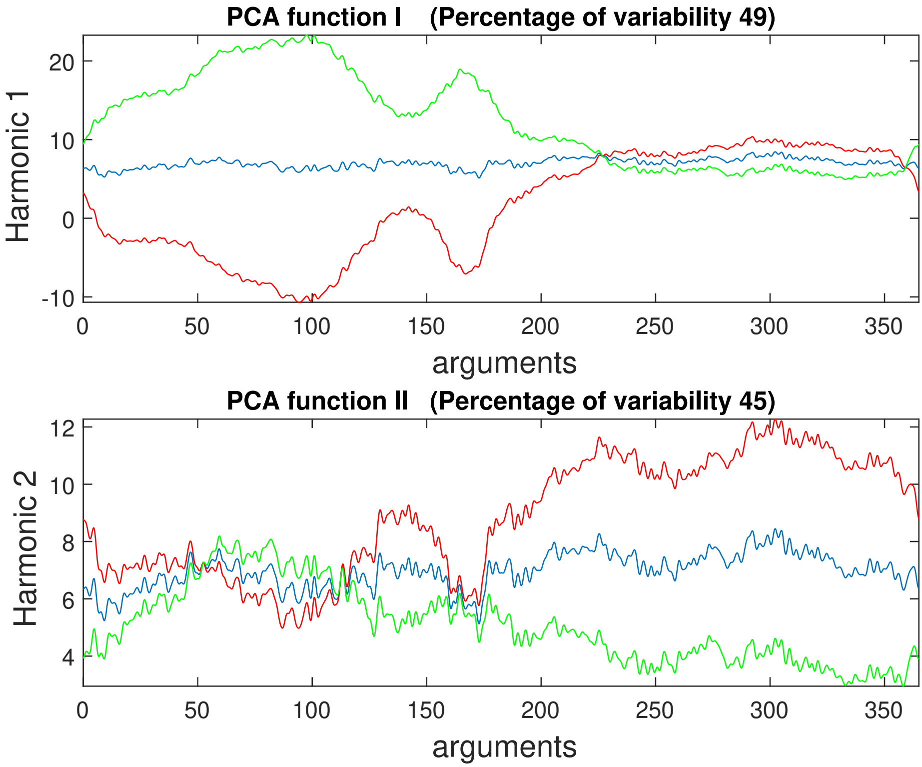


Figure 6.30: Rotated harmonics (Varimax rotation): The two rotated principal component functions are shown as perturbations of the mean, which is the blue line. The top panel shows variation primarily in the first part of the year. The bottom panel shows variation primarily in the second part of the year.

It can be profitable to plot the principal component scores for pairs of harmonics to see how curves cluster and otherwise distribute themselves within the K dimensional subspace spanned by the eigenfunctions. Figure 6.31 reveals some fascinating structure. All Lithuanian stations are contained within two clusters: the upper left with the east stations and the lower right with west stations.

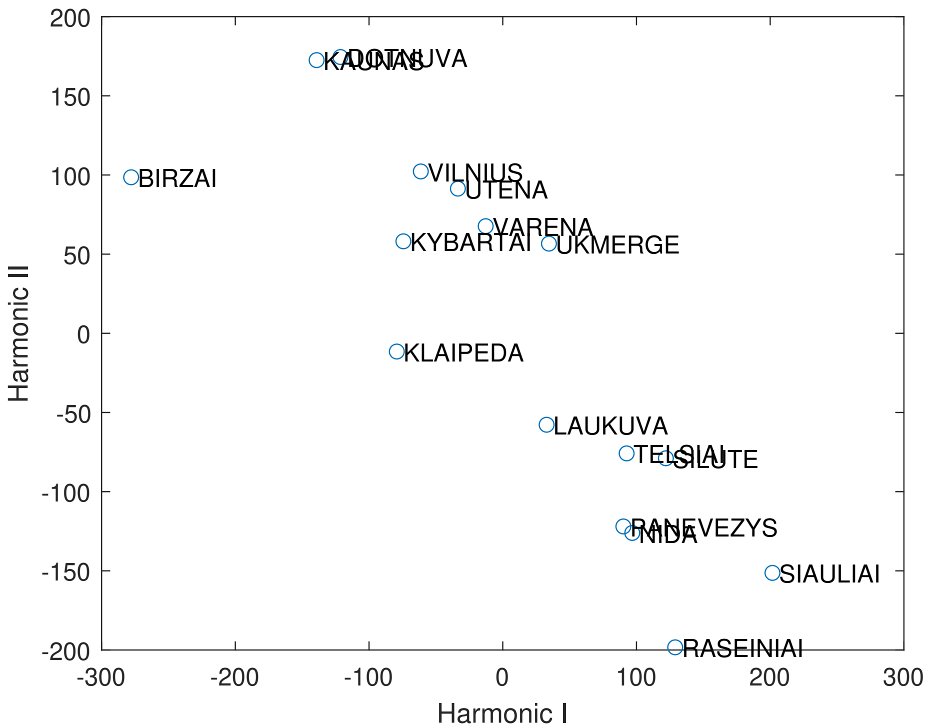


Figure 6.31: The scores for the two rotated principal component functions are shown as circles. Selected stations are labeled in order to identify the two central clusters and the outlying stations.

6.5 Functional linear model

6.5.1 Predicting temperature from climate region

In the Lithuanian weather data, we can divide the weather stations into two distinct groups: West and East Lithuania as we have seen from the section above dedicated to PCA analysis on temperature. It may be interesting to know the effect of geographic location on the shape of the temperature curves. That is, we have a model of the form:

$$T_i(t) = \beta_0(t) + \sum_{j=1}^{2} x_{ij}\beta_j(t) + \epsilon_i(t)$$

where $T_i(t)$ is a functional response. In this case, the values of x_{ij} are either 0 or 1 and ϵ_i are considered to be independently and identically distributed. If the 16 by 3 matrix Z contains these values, then the first column has all entries equal to 1, which codes the contribution of the Lithuanian mean temperature; the remaining two columns contain 1 if that weather station is in the corresponding climate zone and 0 otherwise.

In order to identify the specific effects of the two climate zones, we have to add the constraint:

$$\sum_{j=1}^{2} \beta_j(t) = 0 \; for \; all \; t$$

In our case we will impose this constraint by adding the above equation as an additional $17th$ "observation" for which $T_{17}(t) = 0$.

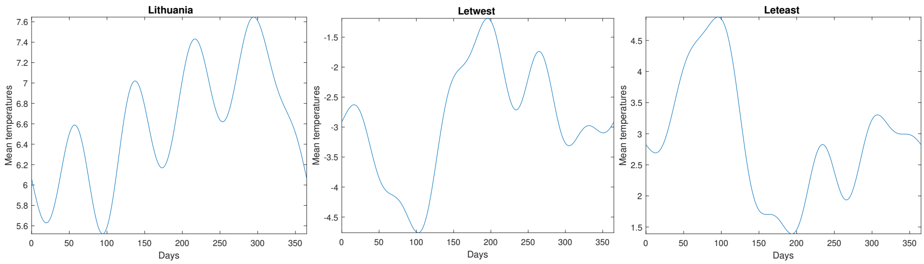


Figure 6.32: The regression coefficients estimated for predicting temperature from climate region. The first panel is the intercept coefficient, corresponding to the Lithuania mean temperature.

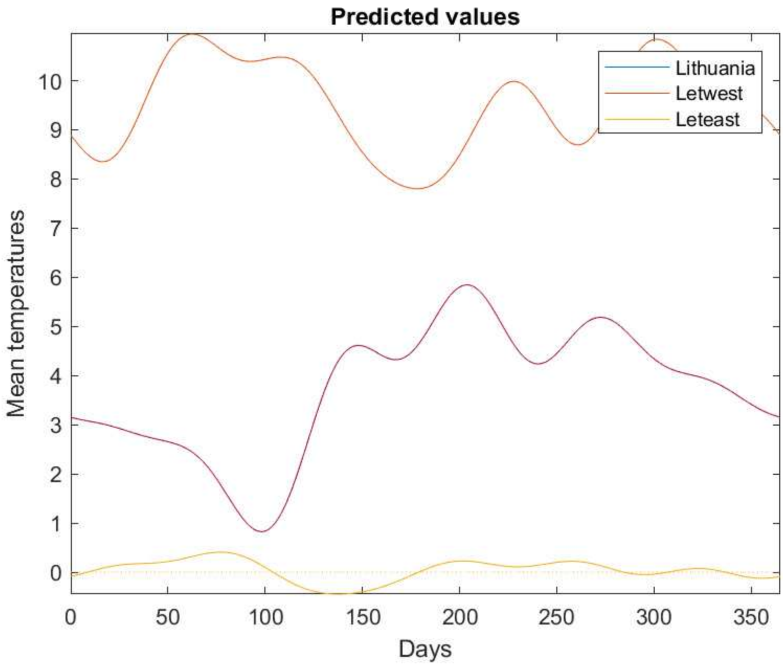


Figure 6.33: The predicted mean temperatures for the three regions.

We first create a list containing three indicator variables for the intercept term and each of the regions. In this setup, the intercept term is effectively the Lithuanian mean temperature curve, and each of the remaining regression coefficients is the perturbation of the Lithuanian mean required to fit a region's mean temperature. The three regression coefficients are shown in Figure 6.32.

Figure 6.33 shows the predicted mean temperature curves for each of the three regions. Note that we have used a fourier basis with 13 basis functions.

Let define y as a vector of 16 dependent variable observations, and an 16 by 3 design functional matrix Z available as a basis for a linear model for y, which contains x_{ij} functions. The vector coefficient function β of length 3 contains each of the regression functions. The concurrent functional linear model in matrix notation is then written as:

$$y(t) = Z(t)\beta(t) + \epsilon(t)$$

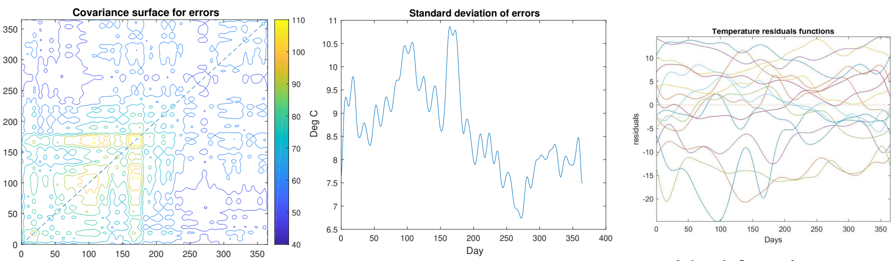
where y is a functional vector of length 16 containing the response functions. Then the vector $r(t)$ will be the corresponding 16-vector of residual functions represented as follows:

$$r(t) = y(t) - Z(t)\beta(t)$$

Figure 6.34 estimate the covariances among the residuals. It is clear that the highest values correspond to period March-June. While the standard deviation of errors is expressed as:

$$\sigma_i(r) = \sqrt{\sum_{i,j=1,i=j}^{N} r_{ij}}$$

where r_{ij} is the residual for ith observation of the jth curve expressed as $r_{ij} = y_{ij} + Z_j(t_i)\beta(t_i)$. Figure 6.34(a) displays the covariance surface of residuals where the highest values correspond to the first part of the year. In figure 6.34(b) is shown the standard deviation of errors for the mean temperature, while in figure 6.34(c) all residual function of mean temperature for each station are shown.



a. covariance surface b. standard deviation c. residual functions

Figure 6.34: A countour plot of variance-covariance surface for errors, standard deviation of errors, and temperature residuals functions for all stations.

6.5.2 Predict wind speed from climate region

We start by smoothing the logarithm of average wind speed directly. We first use 365 Fourier basis functions, and the same harmonic acceleration roughness penalty that we have been using for the temperature data.

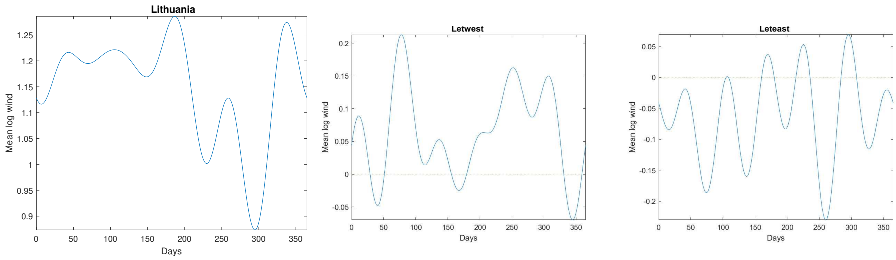


Figure 6.35: The regression coefficients estimated for predicting log wind speed from climate region. The first panel is the intercept coefficient, corresponding to the Lithuania mean log wind speed.

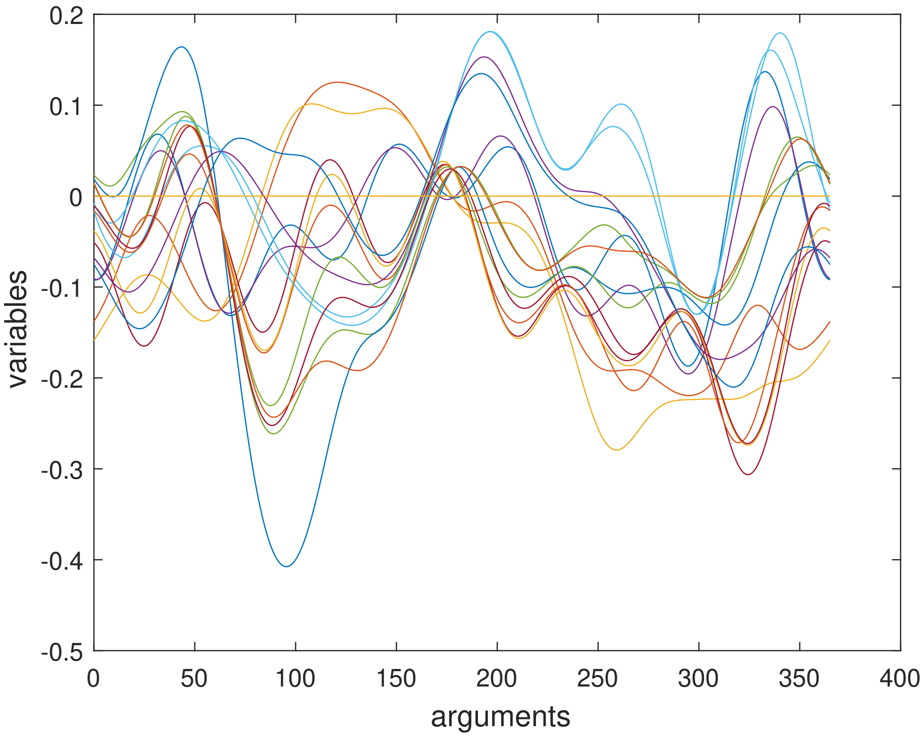


Figure 6.36: The predicted mean log wind speed for all stations.

The generalized cross-validation or GCV criterion was minimized for $\lambda = 10^4$, a level of smoothing that is equivalent to about 27 degrees of freedom. In order to speed up computation, we then opted for a simple Fourier basis expansion with eleven basis functions and no roughness penalization. For this analysis, we used an expansion of the daily average temperature residual in terms of 13 Fourier basis functions.

The three regression coefficients are shown in Figure 6.35, while figure 6.36 shows the predicted mean log wind speed curves for each of all stations.

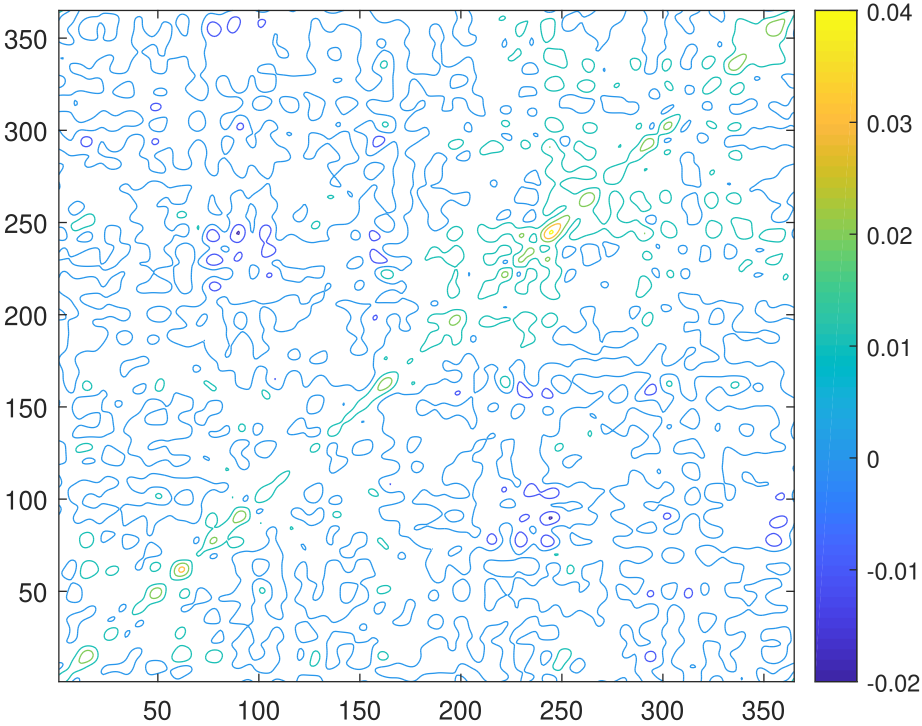


Figure 6.37: Residual matrix and get covariance of residuals: Contour plot of estimated β function .

6.5.3 Annual wind speed predicted by temperature profile

Let now predict total annual wind for Lithuanian weather stations from the pattern of temperature variation through years. Let ALW_i be the logarithm of total annual wind at weather station i, and T_i the daily temperature function then the model takes the form

$$ALW_i = \alpha + \int T_i(s)\beta(s)ds + \epsilon_i$$

where the basis coefficient expansion of β is given by:

$$\beta(s) = \sum_i c_i\phi_i(s) = c^{'}\phi(s)$$

Considering the periodic nature of the temperature and wind data, it seems natural to call for the use of a Fourier series basis. Our first strategy is therefore to represent the regularized fitting problem in terms of a basis function expansion as above, and then to apply the concept of regularization to this representation.

We have chosen to work with 35 Fourier basis functions for the regression coefficient β multiplying the temperature profiles and a constant function for α. In figure 6.38 we plot the estimate of the regression function for the temperature profiles; the intercept value is 1.0687.

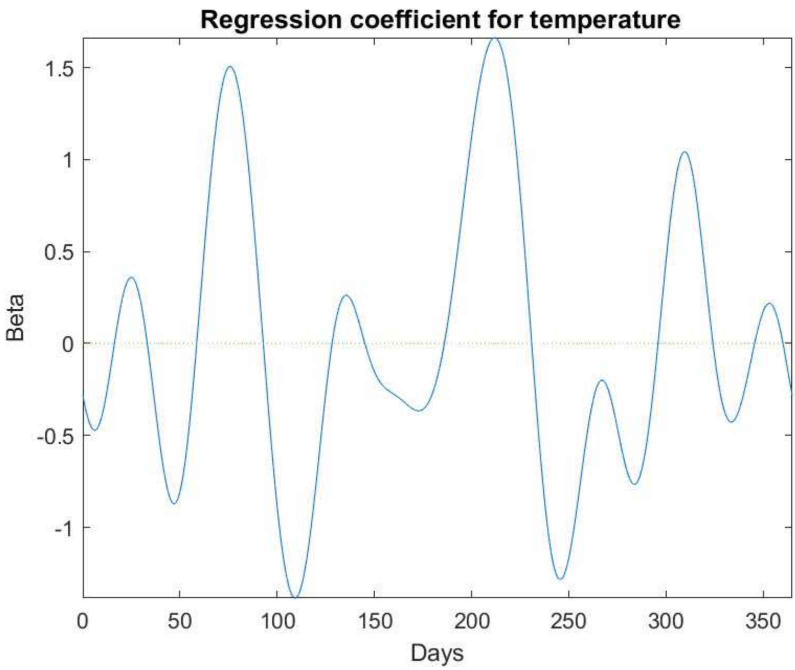


Figure 6.38: Estimated $\beta(s)$ for predicting log annual wind from average daily temperature using 35 Fourier basis functions.

In order to assess the quality of this fit we first extract the fitted values defined by this model and compute the residual sum of squares by the following formula:

$$SSE(\alpha, \beta) = \sum_{i=1}^{N} [ALW_i - \alpha - \int T_i(s)\beta(s)ds]^2$$

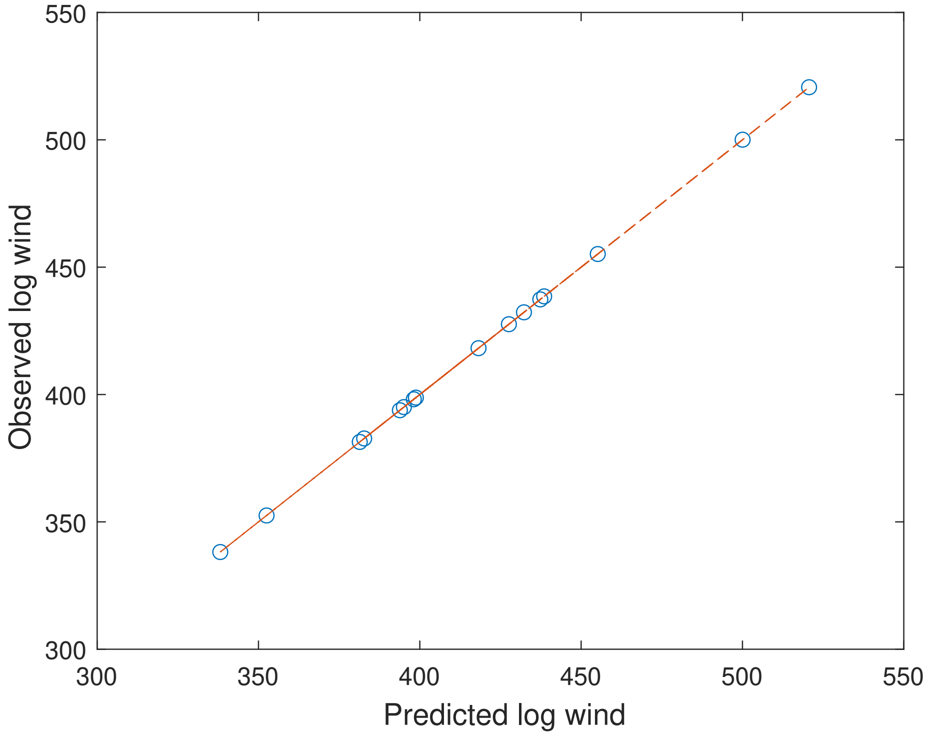


Figure 6.39: Observed log annual wind values plotted against values predicted by functional linear regression on temperature curves.

The squared multiple correlation is 1, and the corresponding F-ratio with 5 and 11 degrees of freedom is -0.65, suggesting a fit to the data that is far better than we would expect by chance.

The estimated function $\beta(s)$ in Figure 6.38 illustrates that fidelity to the observed data, as measured by the residual sum of squares, is not the only aim of the estimation. The roughness penalty approach aims of avoiding excessive local fluctuation in the estimated function (see Cardot (2002)). To this end, the penalized residual sum of squares is defined as:

$$PENSSE_{\lambda}(\alpha, \beta) = \sum_{i=1}^{N} [ALW_i - \alpha - \int T_i(s)\beta(s)ds]^2 + \lambda \int [L\beta(s)]^2 ds$$

where L is a linear differential operator that is suitable for the problem and $L\beta = (\frac{2\pi}{365})^2 D\beta + D^3\beta$. We choose the smoothing parameter λ by cross-validation method as follows:

$$CV(\lambda) = \sum_{i=1}^{N}[ALW_i - \alpha_\lambda^{-i} - \int T_i(s)\beta_\lambda^{-i}(s)]^2 ds$$

where α_λ^{-i} and β_λ^{-i} are the estimates of α and β obtained by minimizing the penalized residual sum of squares based on all the data except (T_i, ALW_i).

We use 65 basis functions to represent the temperature curves and 35 Fourier basis functions to represent β. With this number of basis functions for β, it would be possible to exactly fit the data from the 16 weather stations. However, we want to see how well cross-validation would help us in arriving at a reasonable fit by penalizing harmonic acceleration. Figure 6.40 plots the cross-validation score against the logarithms of various values of λ. The plot shows the minimum point over the range of values plotted. We choose $\lambda = 10^{9.5}$ for the final fit, corresponding to the lower minimum in the plot.

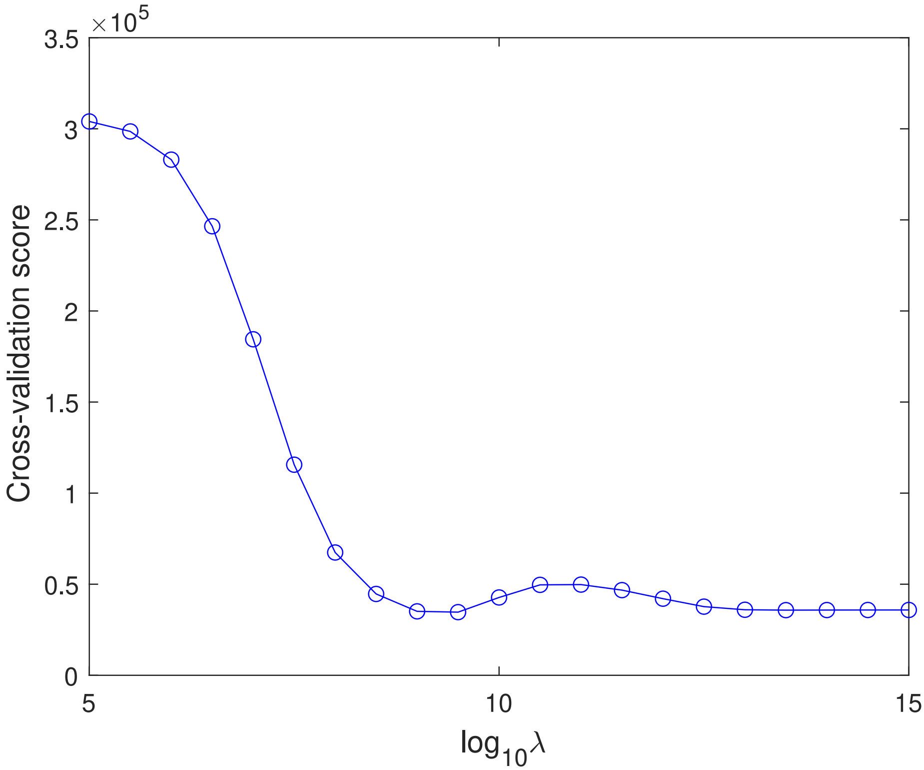


Figure 6.40: The cross-validation score function $CV(\lambda)$ for fitting log annual wind by daily temperature variation, with a penalty on the size of harmonic acceleration. The logarithm of the smoothing parameter is taken to base 10.

Figure 6.41 shows the estimated regression function after the application of the roughness penalty approach along with point-wise 95% confidence limits; the intercept value is 1.1851. The confidence intervals in the beginning of the year contain zero, suggesting that the influence of temperature on wind in that period is not important. However, we see a strong peak in the middle of the summer followed by a valley in the early fall. This pattern is, in effect, computing a contrast between summer and early fall temperatures, with more emphasis on the autumn. This pattern is repeated by another peak (weaker than the first one) in the middle of fall followed by a valley in late fall. This pattern favors weather stations that are comparatively warm during summer and some variation of warm and cool during fall.

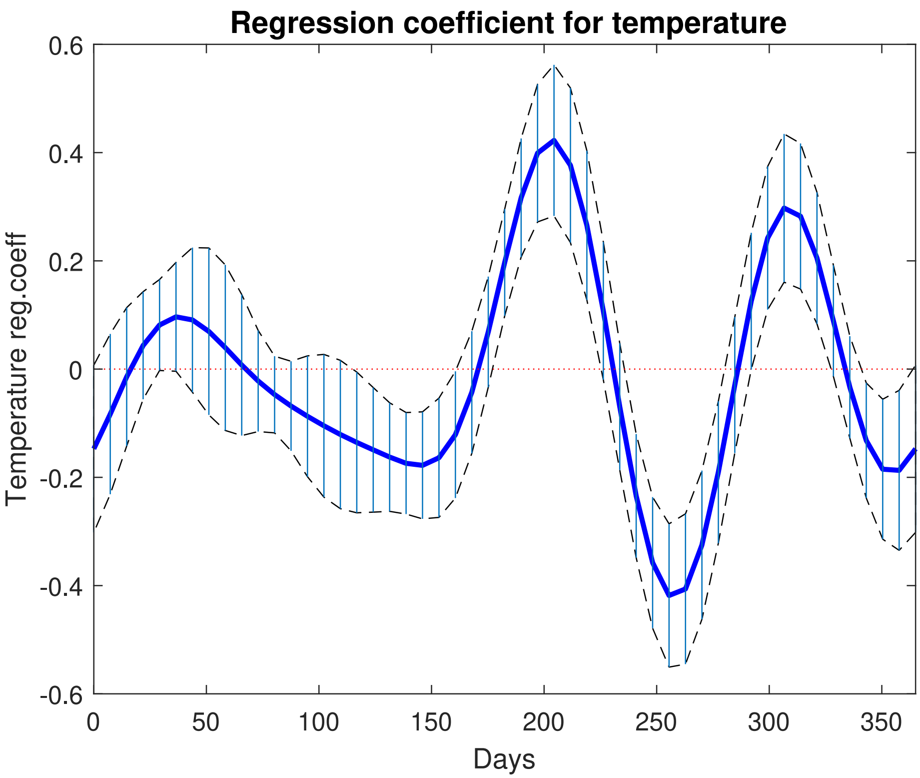


Figure 6.41: Estimate $\beta(t)$ for predicting log annual wind from average daily temperature with a harmonic acceleration penalty and smoothing parameter set to $10^{9.5}$. The dashed lines indicate point-wise 95% confidence limits for values of $\beta(s)$.

In Figure 6.42, we have plotted the observed values ALW_i against the fitted values $\widehat{ALW}_i$ obtained using this functional regression. The squared correlation between the predicted and actual values in the plot is 0.88 and F-ratio is -0.65 for 5 and 11 degrees of freedom. This simple regression diagnostic seems to confirm the model assumptions.

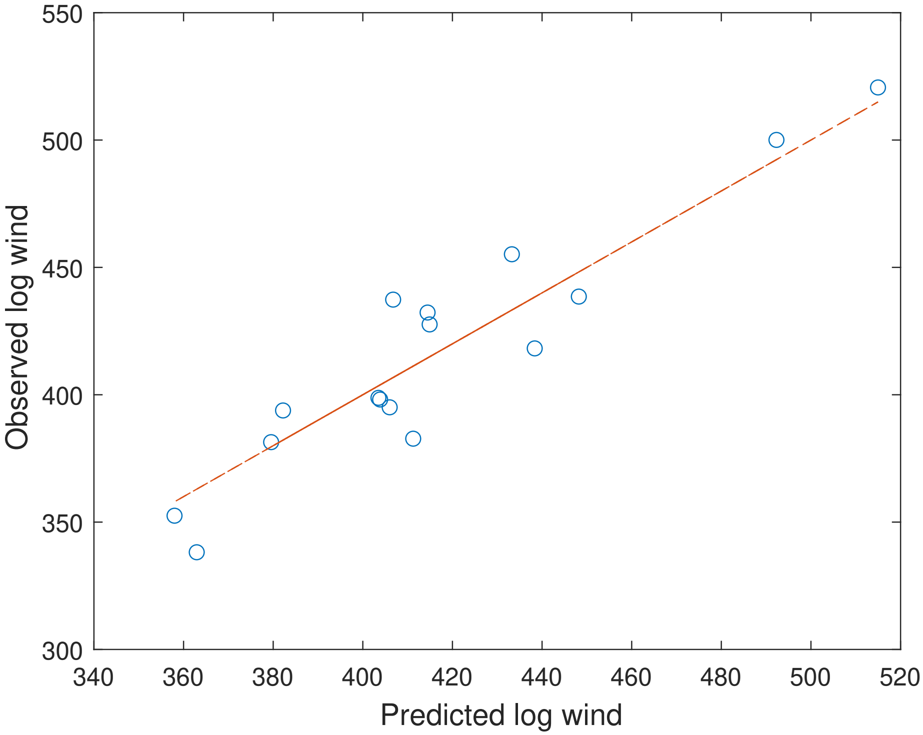


Figure 6.42: Observed values ALW_i of log annual wind plotted against the values $\widehat{ALW}_i$ predicted by the functional regression model with the smoothing parameter chosen by cross-validation. The straight line corresponds to zero residuals.

6.6 Functional linear model: Prediction of wind speed from temperature

We now consider a fully functional linear model in which both the response (log wind speed) and the covariate (temperature) variables are functions.

Predicting temperature is relatively easy, but predicting wind is quite another thing. Certainly there are important wind effects due to climate zones, but can we get additional predictability from the behavior of temperature? In Lithuania, it seems likely, that the behaviour of wind is in line with temperature behaviour, it seems that when temperature is increasing the wind speed is increasing as well and vice versa. In this section we want to investigate to what extent we can predict the complete log daily wind profile LW of a weather station from information in its complete daily temperature profile T.

Because all the functions in this study are intrinsically periodic, we expand both the log wind speed and temperatures in Fourier series. We pre-processed the data by fitting a Fourier series with 65 terms, applying a roughness penalty smoother in order to eliminate very local variation. In order to predict log wind speed from temperature we recall the formula presented in equation (6.1) where log wind speed LW_i is defined as in the following form:

$$LW_i(t) = \beta_0(t) + \int_{\Omega_t} \beta_1(t,s)T_i(s)ds + \epsilon_i(t) \tag{6.9}$$

In contrast of the concurrent model used in case of predicting log wind speed from climate region, the regression function β_1 in this case is a function of both s and t. The regression function $\beta_1(s,t)$ for a fixed value of t is interpreted as the relative weight placed on the temperature at day s that is required to predict log wind speed on day t. The unweighted fitting criterion is the integrated residual sum of squares:

$$LMSSE^3(\beta_0, \beta_1) = \int \sum_{i=1}^{N} [LW_i(t) - \beta_0(t) - \int T_i(s)\beta_1(s,t)ds]^2 dt$$

The smoothed log wind speed curves for all 16 weather stations are shown in Figure 6.43.

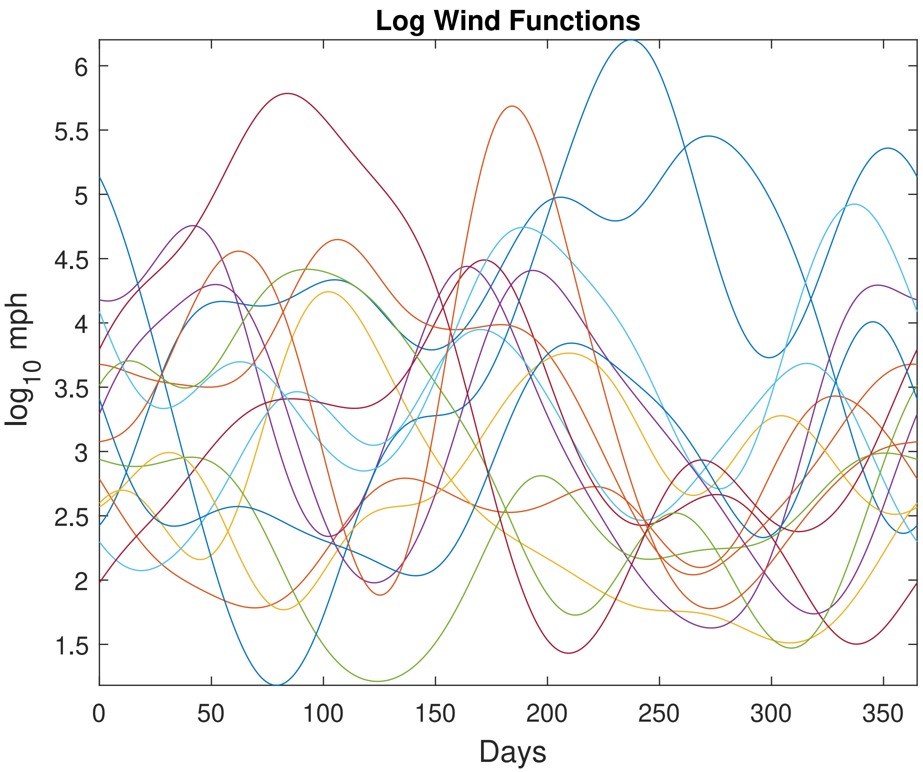


Figure 6.43: Log Wind speed functions for all stations of Lithuania.

In figure 6.44 it is shown the intercept function β_0 (see equation 6.3).

[3] LMSSE is a squared error fitting criterion for a linear model

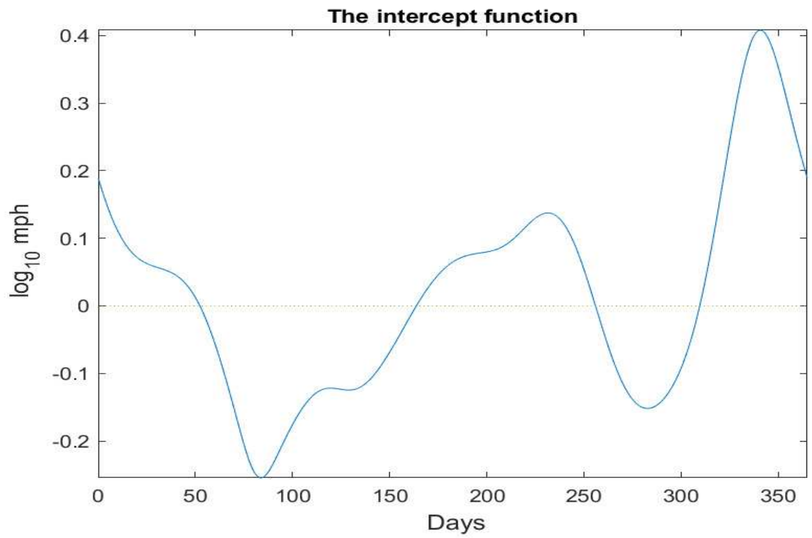


Figure 6.44: The intercept function

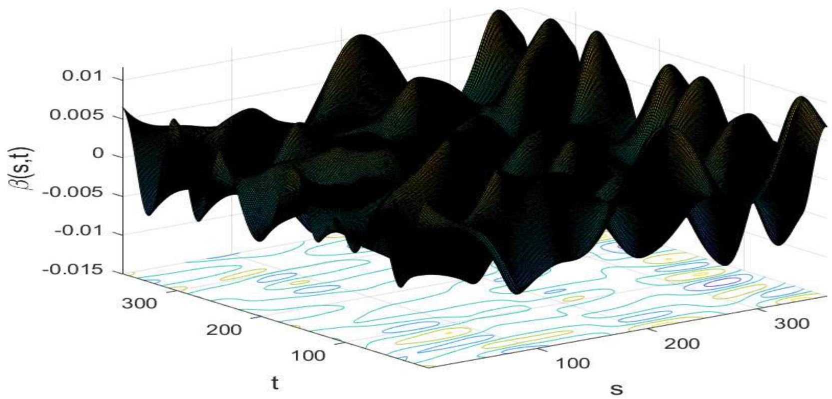


Figure 6.45: The functional parameter function $\beta_1(s,t)$ for the prediction of log wind speed from temperature, estimated direct from the data. The value $\beta_1(s,t)$ shows the influence of temperature at time s on log speed at time t.

Figure 6.45 displays the estimated regression surface $\beta_1(s,t)$ defined in equation (6.2). The estimated intercept function β_0 ranged over values smaller than the response functions, and can therefore be considered to be essentially zero. The height of the surface declines to near zero for large differences between s and t.

We see that the function β_1 estimated by this method is extremely variable and it also turns out that this β_1 gives perfect prediction of the given data in log wind speed.

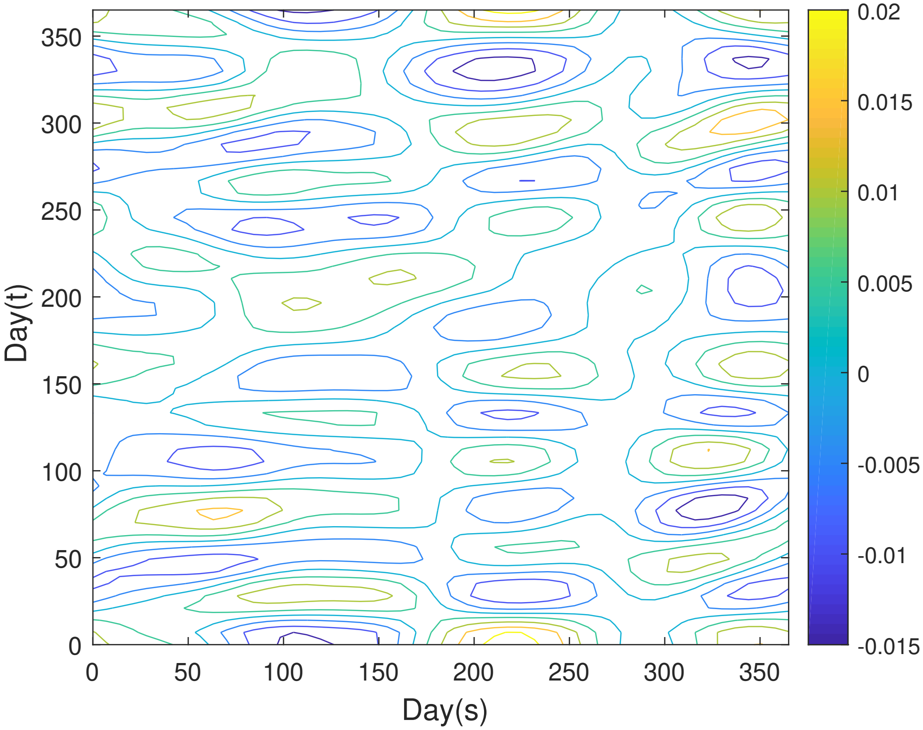


Figure 6.46: Contour plot of estimated β_1 function truncating both bases to 65 terms.

From figure 6.46 we can discern several aspects of the effect of temperature on wind. For instance, temperature in February is negatively associated with wind in the same period and it is positively associated for the most time of the year. Temperature around May is positively associated with wind in the summer months. Temperature in September has a positive association with wind throughout the year.

In order to assess the fit of the functional linear model we have chosen to consider the squared correlation function defined as follow:

$$R^2(t) = 1 - \sum_i (\widehat{y}_i(t) - y_i(t))^2 / \sum_i (y_i(t) - \overline{y}(t))$$

For all stations plotted in Figures 6.47-6.50, for instance, the values of R^2 are more than 50%, illustrating that all stations are places whose wind fit closely to those predicted by the model on the basis of their observed temperature profiles. However, Figures 6.47-6.50 demonstrates that the pattern of wind, judged by comparing the predictions with the original data after subtracting the annual mean for the individual places, is predicted very well for all stations.

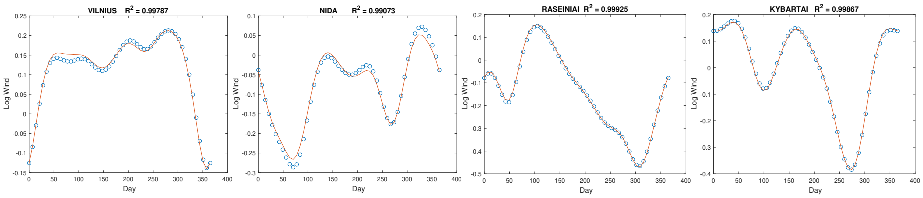


Figure 6.47: Original data (dots) and predictions (solid) of log wind speed relative to mean for each of weather stations. The prediction is carried out using an estimated β_1 function with the both bases truncated to 65 terms.

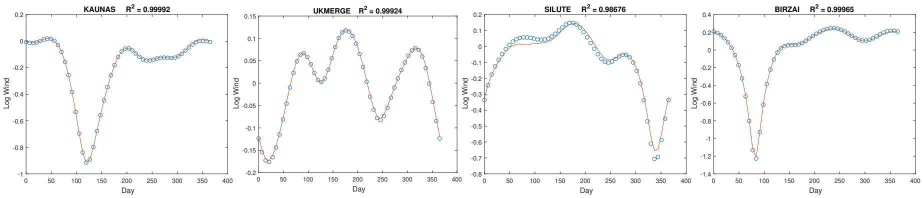


Figure 6.48: Original data (dots) and predictions (solid) of log wind speed relative to mean for each of weather stations. The prediction is carried out using an estimated β_1 function with the both bases truncated to 65 terms.

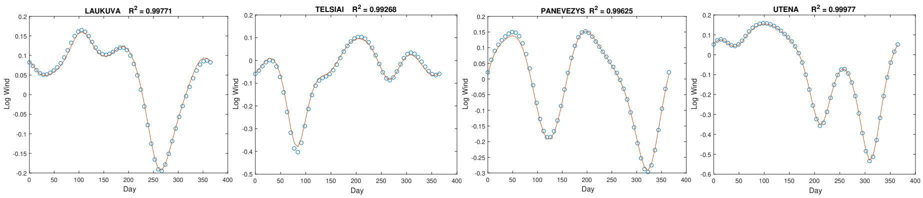


Figure 6.49: Original data (dots) and predictions (solid) of log wind speed relative to mean for each of weather stations. The prediction is carried out using an estimated β_1 function with the both bases truncated to 65 terms.

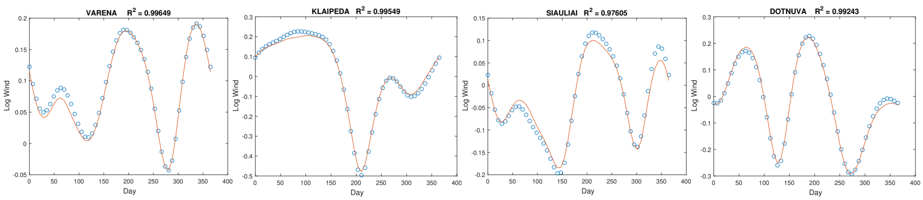


Figure 6.50: Original data (dots) and predictions (solid) of log wind speed relative to mean for each of weather stations. The prediction is carried out using an estimated β_1 function with the both bases truncated to 65 terms.

6.6.1 PCA of wind speed

After the PCA analysis applyied on temperature data, it is natural to apply PCA analysis also on the wind speed data predicted by the functional linear model.

Figure 6.51 shows the two principal component functions by displaying the mean curve along green line and red line indicating the consequences of adding and subtracting a small amount of each principal component. We observe that these two harmonics account for 63% of the variation around the mean log-wind speed curve. We see that the first harmonic, accounting for 39% of the variation, shows essentially a contrast between winter and summer log-wind speed levels and that the second harmonic represents a relative constant vertical shift in the mean.

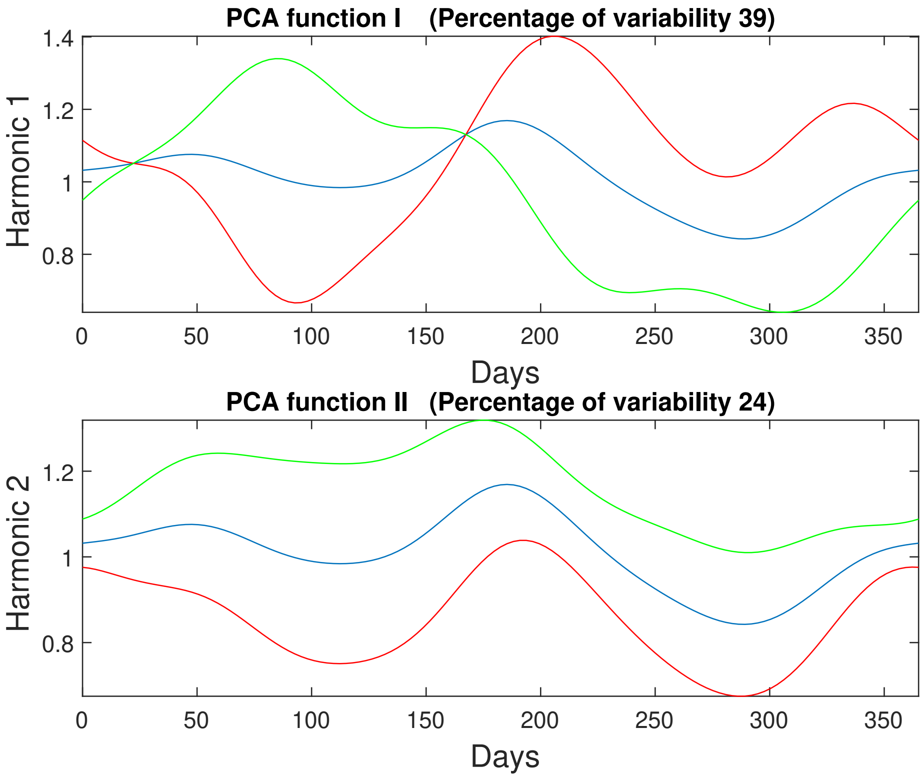


Figure 6.51: Principal components analysis of log-wind speed: The two principal component functions or harmonics are shown as perturbations of the mean, which is the blue line. The green line show what happens when a small amount of a principal component is added to the mean, and the red line show the effect of subtracting the component. The top panel contains the strongest component.

Applying the VARIMAX rotation algorithm as in case of temperature we obtain the results plotted in Figure 6.52. The first component portrays variation that is strongest in second part of the year and the second captures primarily variation from the first part of the year.

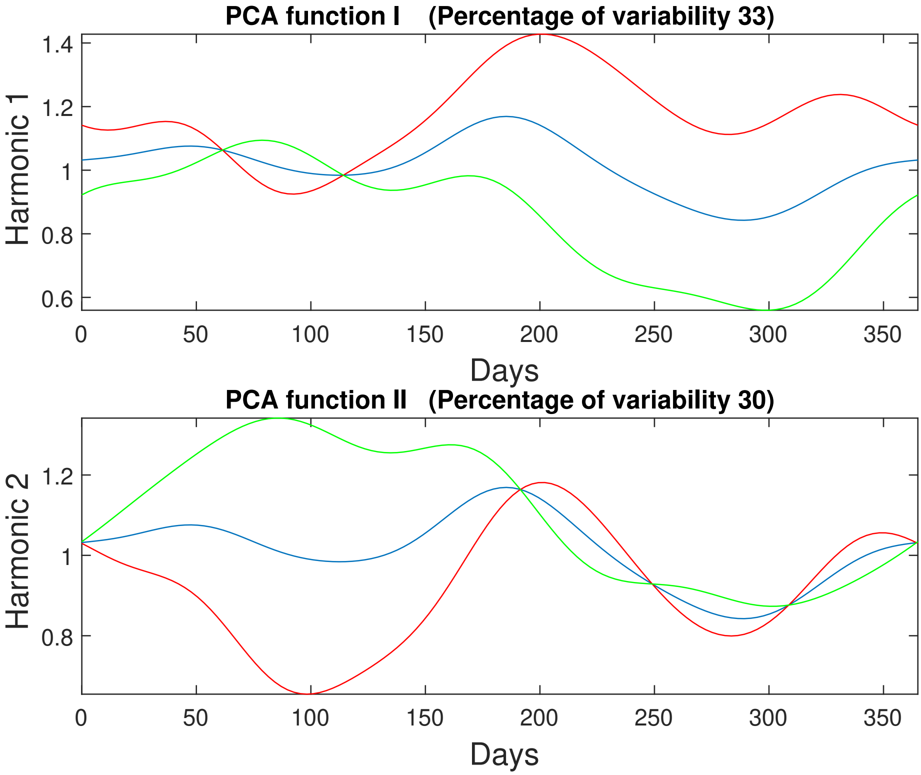


Figure 6.52: Rotated harmonics for log-wind speed (Varimax rotation): The two rotated principal component functions are shown as perturbations of the mean, which is the blue line. The top panel shows variation primarily in the second part of the year. The bottom panel shows variation primarily in the first part of the year.

Also for wind speed data it is clear two see that all meteorological stations can be represented by two clusters, as it is concluded in temperature case even the general percentage variation both groups explain is less than the one obtained in case of temperature data.

6.7 Conclusions

This paper has described the purpose, concepts, and some of the methods of functional data analysis in the context of analyzing temperature and wind speed data. These methods were applied to 27 years of weather data in 16 meteorological stations in Lithuania. Instead of working with discrete data, functional data analysis allowed for statistical tests to be performed on the data represented as functional objects. The process of transforming the data into a functional form included generalized cross validation tests and minimization of sum of squared residuals.

The main goal of this paper was to provide a model for wind speed which is easily understandable and give a reliable prediction in time. We have proposed the functional linear model observed in different patterns which is powerful enough to describe the dynamics of wind speed in time.

The functional linear model firstly is used to predict wind speed from climate zone using functional principal component analysis. Secondly, annual wind speed is fitted by using temperature as a functional covariate, where the harmonic acceleration roughness in the regression coefficient function is penalized. Thirdly the full wind speed profile is fitted by the regressing on the full temperature profile, using a level of smoothing by applying generalized cross-validation criterion. In order to compare the predictions, we calculated the squared correlation coeficient (R^2). The forecasting results from all three approaches using the functional techniques are promising.

Further analysis can be done to study the rate of change of the wind speed function as a dependent variable and the usage of other dimension reduction techniques.

Acknowledgment

The authors of this paper gratefully thank the staff at the *Lithuanian Hydrometeorological Service* (LHS) for its excellent management of data source.

Conflict of interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Bibliography

[1] Alexandridis, A., & Zapranis, A. (2013). Wind derivatives: Modeling and pricing. *Computational Economics*, **41**(3), 299–326.

[2] Benth, J. Š., & Benth, F. E. (2010). Analysis and modelling of wind speed in New York. *Journal of Applied Statistics*, **37**(6), 893–909.

[3] Benth, F. E., & Pircalabu, A. (2018). A non-Gaussian Ornstein–Uhlenbeck model for pricing wind power futures. *Applied Mathematical Finance*, **25**(1), 36–65.

[4] Boente, G., & Fraiman, R. (2000). Kernel-based functional principal components. *Statistics and probability letters*, **48**(4), 335–345.

[5] Boente, G., Salibian-Barrera, M., & Vena, P. (2020). Robust estimation for semi-functional linear regression models. *Computational Statistics Data Analysis*, 107041.

[6] Bosq, D. (1991). Modelization, nonparametric estimation and prediction for continuous time processes. In Nonparametric functional estimation and related topics (pp. 509–529). *Springer, Dordrecht*.

[7] Brockhaus, S., Scheipl, F., Hothorn, T., & Greven, S. (2015). The functional linear array model. *Statistical Modelling*, **15**(3), 279–300.

[8] Brumback, B. A., & Rice, J. A. (1998). Smoothing spline models for the analysis of nested and crossed samples of curves. *Journal of the American Statistical Association*, **93**(443), 961–976.

[9] Cardot, H. (2002). Local roughness penalties for regression splines. *Computational Statistics*, bf 17(1), 89–102.

[10] Cardot, H., Ferraty, F., & Sarda, P. (1999). Functional linear model. *Statistics and Probability Letters*, **45**(1), 11–22.

[11] Craven, P., & Wahba, G. (1979). Estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematik*, 31, 377–403.

[12] Cuevas, A., Febrero, M., & Fraiman, R. (2002). Linear functional regression: the case of fixed design and functional response.*Canadian Journal of Statistics*, **30**(2), 285–300.

[13] Dauxois, J., Pousse, A., & Romain, Y. (1982). Asymptotic theory for the principal component analysis of a vector random function: some applications to statistical inference. *Journal of multivariate analysis*, **12**(1), 136–154.

[14] Fan, Y., James, G. M., & Radchenko, P. (2015). Functional additive regression. *The Annals of Statistics*, **43**(5), 2296–2325.

[15] Ferraty, F., & Vieu, P. (2002). The functional nonparametric model and application to spectrometric data. *Computational Statistics*, **17**(4), 545–564.

[16] Ferraty, F., & Vieu, P. (2003). Curves discrimination: a nonparametric functional approach. *Computational Statistics and Data Analysis*, **44**(1–2), 161–173.

[17] Foutz, N. Z., & Jank, W. (2010). Research note—prerelease demand forecasting for motion pictures using functional shape analysis of virtual stock markets. *Marketing Science*, **29**(3), 568–579.

[18] Goia, A., May, C., & Fusai, G. (2010). Functional clustering and linear regression for peak load forecasting. *International Journal of Forecasting*, **26**(4), 700–711.

[19] Green, P. J., & Silverman, B. W. (1994). Nonparametric regression and generalized linear models. Number 58 in Monographs on Statistics and Applied Probability. *Nonparametric regression and generalized linear models*, (58).

[20] Guo, M., Zhou, L., Huang, J. Z., & Härdle, W. K. (2015). Functional data analysis of generalized regression quantiles. *Statistics and Computing*, **25**(2), 189–202.

[21] Hardle, W. K., & Osipenko, M. (2012). Spatial risk premium on weather derivatives and hedging weather exposure in electricity. *The Energy Journal*, **33**(2).

[22] Heckman, N. E., & Ramsay, J. O. (2000). Penalized regression with model-based penalties. *Canadian Journal of Statistics*, **28**(2), 241–258.

[23] Kneip, A., Poß, D., & Sarda, P. (2016). Functional linear regression with points of impact. *The Annals of Statistics*, **44**(1), 1–30.

[24] Locantore, N., Marron, J. S., Simpson, D. G., Tripoli, N., Zhang, J. T., Cohen, K. L., ... & Fan, J. (1999). Robust principal component analysis for functional data. *Test*, **8**(1), 1–73.

[25] Marron, J. S., Ramsay, J. O., Sangalli, L. M., & Srivastava, A. (2015). Functional data analysis of amplitude and phase variation. *Statistical Science*, 468–484.

[26] Morris, J. S. (2015). Functional regression. *Annual Review of Statistics and Its Application*, 2, 321–359.

[27] Pezzulli, S. (1993). Some properties of smoothed principal components analysis for functional data. *Computational Statistics*, **8**(1), 1–16.

[28] Ramsay, J. O., & Dalzell, C. J. (1991). Some tools for functional data analysis. *Journal of the Royal Statistical Society: Series B (Methodological)*, **53**(3), 539–561.

[29] Ramsay, J. O., & Silverman, B. W. (1997). Principal differential analysis. In Functional data analysis. *Springer*, 239–256.

[30] Ramsay, J. O., & Silverman, B. W. (2002). *Applied Functional Data Analysis*.

[31] Ramsay, J. O., & Silverman, B. W. (2005). Springer series in statistics. *Functional data analysis*, 10–18.

[32] Rice, J. A., & Silverman, B. W. (1991). Estimating the mean and covariance structure nonparametrically when the data are curves. *Journal of the Royal Statistical Society: Series B (Methodological)*, *53*(1), 233–243.

[33] Šaltytė Benth, J., & Šaltytė, L. (2011). Spatial–temporal model for wind speed in Lithuania. *Journal of Applied Statistics*, **38**(6), 1151–1168.

[34] Silverman, B. W. (1996). Smoothed functional principal components analysis by choice of norm. *The Annals of Statistics*, **24**(1), 1–24.

[35] Wang, J. L., Chiou, J. M., & Müller, H. G. (2016). Functional data analysis. *Annual Review of Statistics and Its Application*, 3, 257–295.

[36] Yu, P., Du, J., & Zhang, Z. (2020). Single-index partially functional linear regression model. *Statistical Papers*, **61**(3), 1107–1123.