A REGIME SWITCHING MODEL FOR THE TEMPERATURE AND PRICING WEATHER DERIVATIVES

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF APPLIED MATHEMATICS
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

AYSUN TÜRKVATAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
FINANCIAL MATHEMATICS

AUGUST 2016

Approval of the thesis:

A REGIME SWITCHING MODEL FOR THE TEMPERATURE AND PRICING WEATHER DERIVATIVES

submitted by **AYSUN TÜRKVATAN** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Department of Financial Mathematics, Middle East Technical University** by,

Prof. Dr. Bülent Karasözen
Director, Graduate School of **Applied Mathematics** ______________

Assoc. Prof. Dr. Yeliz Yolcu-Okur
Head of Department, **Financial Mathematics** ______________

Assoc. Prof. Dr. Azize Hayfavi
Supervisor, **Financial Mathematics, METU** ______________

Prof. Dr. Tolga Omay
Co-supervisor, **Business Administration, UTAA** ______________

Examining Committee Members:

Prof. Dr. Gerhard-Wilhelm Weber
Institute of Applied Mathematics, METU ______________

Assoc. Prof. Dr. Azize Hayfavi
Institute of Applied Mathematics, METU ______________

Assoc. Prof. Dr. Ömür Uğur
Institute of Applied Mathematics, METU ______________

Assoc. Prof. Dr. Kasırga Yıldırak
Department of Actuarial Science, Hacettepe University ______________

Assoc. Prof. Dr. Ümit Aksoy
Department of Mathematics, Atılım University ______________

Date: ______________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: AYSUN TÜRKVATAN

Signature :

ABSTRACT

A REGIME SWITCHING MODEL FOR THE TEMPERATURE AND PRICING WEATHER DERIVATIVES

Türkvatan, Aysun
Ph.D., Department of Financial Mathematics
Supervisor : Assoc. Prof. Dr. Azize Hayfavi
Co-Supervisor : Prof. Dr. Tolga Omay

August 2016, 81 pages

Weather has an enormous impact on many institutions, for example, in energy, agriculture, or tourism sectors. For example, a gas provider faces the reduced demand in gas in case of hot winter. Weather derivatives can be used as a tool to manage the risk exposure towards adverse or unexpected weather conditions. Weather derivatives are the financial contracts with underlying depending on weather variables such as temperature, humidity, precipitation or snow. Since the temperature is the most commonly used weather variable, we consider the temperature based weather derivatives. These are the financial contracts written on several temperature indices, such as the cumulative average temperature (CAT), or the cooling degree days (CDD). We first propose a regime-switching model for the temperature dynamics, where the parameters depend on a Markov chain. Also, since the jumps in the temperature are directly related to the regime switch, we model them by the chain itself. Morever, the estimation and forecast of the proposed model is considered. It is shown that forecast performance of the proposed model is in line with the existing models considered. After modeling the temperature dynamics, to price the derivatives, the risk-neutral probability is to be specified. Since temperature (and hence the index) is not a tradeable asset, any probability measure being equivalent to the objective probability is a risk-neutral probability. We consider a generalized version of the Esscher transform to select an equivalent measure. Then we derive prices of weather derivatives written on several temperature indices.

Keywords : regime-switching, Markov chain, expectation-maximization algorithm, pricing, weather derivatives

ÖZ

HAVA SICAKLIĞI İÇİN REJİM DEĞİŞİM MODELİ VE HAVA DURUMU TÜREVLERİNİN FİYATLAMASI

Türkvatan, Aysun

Doktora, Finansal Matematik Bölümü

Tez Yöneticisi : Doç. Dr. Azize Hayfavi

Ortak Tez Yöneticisi : Prof. Dr. Tolga Omay

Ağustos 2016, 81 sayfa

Hava durumu birçok kuruluş üzerinde çok büyük bir etkiye sahiptir, örneğin, enerji, tarım, veya turizm sektörlerindeki kuruluşlar. Örneğin, bir doğal gaz sağlayıcı kış mevsiminin sıcak geçmesi durumunda doğal gaz talebinde azalış ile karşılacaktır. Hava durumu türevleri olumsuz veya beklenmeyen hava koşullarına ilişkin riski yönetmek için bir araç olarak kullanılabilir. Hava durumu türevleri sıcaklık, nem, yağış veya kar gibi hava değişkenleri üzerine yazılan finansal sözleşmelerdir. En çok kullanılan hava değişkeni sıcaklık olduğu için sıcaklığa dayalı hava türevleri ele alınmaktadır. Bunlar kümülatif ortalama sıcaklık (CAT) veya soğutma gün dereceleri (CDD) gibi çeşitli sıcaklık endeksleri üzerine yazılan finansal sözleşmelerdir. Öncelikle, hava sıcaklığı dinamikleri için parametrelerin Markov zincirine bağlı olduğu bir rejim-değişim modeli önerilmektedir. Aynı zamanda, hava sıcaklığındaki sıçramalar, rejim değişimi ile doğrudan ilişkili olduğu için, bunlar zincirin kendisi ile modellenmektedir. Ayrıca, önerilen modelin parametre tahmini ve kestrimi ele alınmıştır. Önerilen modelin kestrim performansının ele alınan var olan diğer modeller ile uyumlu olduğu gösterilmiştir. Hava sıcaklığı dinamikleri modellendikten sonra, türevleri fiyatlamak için, risk-nötr olasılık belirlenmelidir. Ancak, hava sıcaklığı (ve dolayısıyla endeks) ticarete konu bir varlık olmadığından objektif olasılığa eşdeğer olan herhangi bir olasılık ölçüsü risk-nötr olasılıktır. Eşdeğer ölçüyü seçmek için Esscher dönüşümünün genelleştirilmiş bir versiyonu ele alınmıştır. Sonra çeşitli sıcaklık endeksleri üzerine yazılan hava durumu türevlerinin fiyatları elde edilmiştir.

Anahtar Kelimeler: rejim-değişimi, Markov zinciri, maksimum-beklenti algoritması, fiyatlama, hava durumu türevleri

Dedicated to My Family

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my thesis supervisor, Assoc. Prof. Dr. Azize Hayfavi, for her patient guidance, enthusiastic encouragement and motivation throughout this study.

I would like to express my sincere gratitude to my thesis co-supervisor, Prof. Dr. Tolga Omay, for his patient guidance, enthusiastic encouragement and motivation throughout this study.

I am especially grateful to Prof. Dr. Gerhard-Wilhelm Weber for his enthusiastic encouragement, valuable advices and endless support.

I would like to thank Assoc. Prof. Dr. Ömür Uğur, Assoc. Prof. Dr. Kasırga Yıldırak and Assoc. Prof. Dr. Ümit Aksoy for their valuable advices and support.

I am grateful to Assoc. Prof. Dr. Yeliz Yolcu-Okur for her great support.

I would like to thank all the members of the Institute of Applied Mathematics for their kindness. My thanks go to my friends for their support.

I would like to express my deepest gratitude to my family for their endless love, encouragement, motivation and support throughout my life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AR	Autoregressive
CAR	Continuous-time autoregressive
CAT	Cumulative average temperature
CDD	Cooling degree days
CRS	Constant-speed regime-switching
DRS	Double regime-switching
EM algorithm	Expectation-Maximization algorithm
HDD	Heating degree days
$L^1(\mathbb{R})$	Space of integrable functions on $\mathbb{R}$
MAE	Mean absolute error
MSE	Mean square error
$\mathbb{R}$	Set of real numbers
RMSE	Root mean square error
RS	Regime-switching
TAR	Threshold autoregressive

CHAPTER 1

INTRODUCTION

Weather derivatives provide a tool to manage the weather risk. Weather has an enormous impact on many institutions, for example, in energy, agriculture, or tourism sectors. For example, a gas provider faces the reduced demand in gas in case of hot winter. Weather derivatives can be used as a tool to manage the risk exposure towards adverse or unexpected weather conditions. Weather derivatives are the financial contracts with underlying depending on weather variables such as temperature, humidity, precipitation or snow. More information on weather derivatives can be found in [7] and [2]. Since the temperature is the most commonly used weather variable, we consider the temperature based weather derivatives. These are the financial contracts written on several temperature indices. The most common ones are the cumulative average temperature (CAT), the cooling degree days (CDD) and the heating degree days (HDD) indices.

The CAT, CDD and HDD indices over a measurement period $[\tau_1, \tau_2]$ are defined as

$$CAT(\tau_1, \tau_2) = \int_{\tau_1}^{\tau_2} T_t dt,$$

$$CDD(\tau_1, \tau_2) = \int_{\tau_1}^{\tau_2} \max(T_t - c, 0) dt,$$

and

$$HDD(\tau_1, \tau_2) = \int_{\tau_1}^{\tau_2} \max(c - T_t, 0) dt,$$

respectively, where T_t is the temperature at time t and c is a constant and denotes the threshold temperature, typically 18 degrees Celsius or 65 degrees Fahrenheit.

We consider futures written on CAT, CDD and HDD indices. To derive the futures prices, we first model the temperature dynamics. In the literature, the mean-reverting Ornstein–Uhlenbeck process in different forms is commonly used for modeling the temperature. However, Elias et al.(2014) ([18]) state that abrupt changes in temperature, caused by a combination of several factors including latitude, intensity of solar circulation, land and water surface areas, ocean currents, elevation, and clear skies, induce the regime-switching behavior in temperature. Motivated by [18], we propose a regime-switching model for the temperature dynamics, where the parameters depend on a Markov chain. Also, since the jumps in the temperature are directly related to the regime switch, we model them by the chain itself. The jumps can be considered as the shifts in the level of the temperature due to the transitions of the state of the atmospheric conditions. Moreover, the estimation and forecast of the proposed model is considered. It is shown that forecast performance of the proposed model is in line with the existing models considered.

The objective of the Markov regime-switching models is to represent the observed stochastic behavior by at least two separate regimes with different underlying stochastic processes. The switching mechanism between the regimes is an unobserved (latent) Markov chain. One of the main features of Markov regime-switching models is that the regime-switching mechanism allows for temporal changes of model dynamics. Markov regime-switching models can be considered as generalizations of hidden Markov models (Cappé et al. 2005, [10]). [10] states that a hidden Markov model is a doubly stochastic process with an underlying stochastic process that is not directly observable (”hidden”) but can be observed only through another stochastic process that produces the sequence of observations. For more information on hidden Markov models, see Elliott et al. (1995)[19] and [10]. For the applications of hidden Markov models to finance, see [31] and [32]. Unlike the hidden markov models, Markov regime-switching models allow for temporary dependence within the regimes, in particular, for mean reversion, which is a characteristic feature the temperature.

There are also threshold type regime-switching models suc as threshold autoregressive (TAR) model proposed by Tong (1983) [40], see [22]. The main difference between the threshold type regime-switching models and Markov regime-switching models is that in case of the former the switching mechanism between the regimes is observable, while in case of the later it is latent. In this thesis, we focus on Markov regime-swithcing models, we call the Markov regime-switching models simply the regime-switching models.

For the regime-switching models, the type of dependence between the regimes, that is, dependent regimes or independent regimes, is also an important issue. In the former approach, depending on the state process values, only the model parameters change, see Hamilton (1989) [24] and Hamilton (1990) [25]. On the other hand, in the latter, the individual regimes are driven by independent processes. [28] states that dependent regimes lead to computationally simpler models, on the other hand, independent regimes allow for a greater flexibility and admit qualitatively different dynamics in each regime.

After modeling the temperature dynamics, to derive the futures prices, the risk-neutral probability is to be specified. Since the temperature (and hence the index) is not a tradeable asset, any probability measure being equivalent to the objective probability is a risk-neutral probability. A generalized version of the Esscher transform is considered to select an equivalent measure. Then the prices of weather derivatives written on several temperature indices are derived using the temperature model proposed.

The structure of this thesis is as follows. In Chapter 2, after giving the literature review, a new model for the temperature dynamics is proposed. In Chapter 3, estimation and forecast of the proposed model together with existing models are considered. In Chapter 4, the proposed model under the equivalent measure is considered and the prices of weather derivatives written on several temperature indices are derived. In Chapter 5, the conclusion follows. In the Appendix, an overview of Markov chains is provided.

CHAPTER 2

A REGIME SWITCHING MODEL FOR THE TEMPERATURE

In the literature, various models for the temperature dynamics are proposed. In this part, after giving the literature review, a new model for the temperature dynamics is proposed.

2.1 The Literature Review

In the literature, the Ornstein-Uhlenbeck mean-reverting process in different forms is commonly used for modeling temperature. Dornier and Querel (2000) [16] propose the temperature model

$$dT_t = dS_t + \kappa(S_t - T_t)dt + \sigma dW_t,$$

where the seasonal mean S_t is given by

$$S_t = a + bt + c\sin(\omega t + \varphi), \tag{2.1}$$

with $\omega = 2\pi/365$. Here, κ is the speed of mean reversion, σ is the volatility of temperature and W_t is a Brownian motion. The term dS_t expresses the seasonal variation and ensures that the process tends to the seasonal mean in the long run, that is $\mathbb{E}(T_t) = S_t$.

Alaton et al. (2002) [1] suggest the temperature model given by

$$dT_t = dS_t + \kappa(S_t - T_t)dt + \sigma_t dW_t, \tag{2.2}$$

where the seasonal mean S_t is given by Equation (2.1) and σ_t is a piecewise constant function, with a constant value during each month.

Benth and Benth (2007) [5] consider the model in Equation (2.2), where both S_t and σ_t^2 are expressed by a truncated Fourier series, that is,

$$S_t = a + bt + \sum_{i=1}^{I_1} a_i \sin(\omega i(t - f_i)) + \sum_{j=1}^{J_1} b_j \cos(\omega j(t - g_j))$$

and

$$\sigma_t^2 = c + \sum_{i=1}^{I_2} c_i \sin(\omega i t) + \sum_{j=1}^{J_2} d_j \cos(\omega j t).$$

The authors of [5] state that it is sufficient to set $I_1 = 0, J_1 = 1, I_2 = J_2 = 4$.

Mraoua and Bari (2007) [33] consider the model in Equation (2.2), where S_t is given by Equation (2.1) and

$$d\sigma_t = \kappa_\sigma \left(\sigma_{trend} - \sigma_t\right) dt + \gamma_\sigma dW_t,$$

where σ_{trend} is assumed to be constant.

Zapranis and Alexandridis (2008) [41] extends the model suggested by [5]. A nonlinear AR(1) model is estimated non-parametically with a neural network, which removes the constraint of a constant mean-reverting parameter. The form of S_t and σ_t^2 are determined by wavelet analysis.

Broady et al. (2002) [9] suggest the model given by

$$dT_t = \kappa(S_t - T_t)dt + \sigma_t dW_t^H,$$

where W_t^H is a fractional Brownian motion.

Benth and Benth (2005) [4] propose an Ornstein-Uhlenbeck model with seasonal mean and volatility, where the residuals are generated by a Lévy process. In particular, it is suggested to use the class of generalized hyperbolic Lévy processes. [4] suggest

$$dT_t = dS_t + \kappa(T_t - S_t)dt + \sigma_t dL_t,$$

where L_t is a pure-jump Lévy process, and

$$S_t = a_0 + a_1 \cos(\omega(t - a_2)).$$

Benth et al. (2007) [8] consider a pth order continuous-time autoregressive (CAR(p)) model. Let $\mathbf{X}_t = (X_t^1, \ldots, X_t^p)'$ be a stochastic process in $\mathbb{R}^p$ for $p \geq 1$ defined by

$$d\mathbf{X}_t = A\mathbf{X}_t dt + \mathbf{e}_p \sigma_t dW_t,$$

where $\mathbf{e}_k,\ k = 1, \ldots, p$, is the kth unit vector in $\mathbb{R}^p$ and $\sigma_t > 0$ is a real-valued and square integrable function (over any finite time interval). The $p \times p$ matrix A is defined by

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \cdots \\ 0 & 0 & 0 & \vdots & 1 \\ -\alpha_p & -\alpha_{p-1} & -\alpha_{p-2} & \cdots & -\alpha_1 \end{bmatrix},$$

where $\alpha_k,\ k = 1, \ldots, p$, are constants. [8] propose that

$$T_t = \Lambda_t + Y_t,$$

where

$$\Lambda_t = a_0 + a_1 t + a_2 \cos(\omega(t - a_3))$$

is the deterministic seasonal mean function and $Y_t = \mathbf{e}_1^{'} \mathbf{X}_t = X_t^1$ is the deseasonalized temperature. [8] suggest that

$$\sigma_t^2 = b_1 + \sum_{k=1}^{4} b_{2k} \cos(\omega k t) + \sum_{k=1}^{4} b_{2k+1} \sin(\omega k t).$$

Benth and Benth (2011) [6] generalizes the CAR model proposed in [8]. The following CAR(p) model with seasonal stochastic volatility is proposed:

$$d\mathbf{X}_t = A\mathbf{X}_t dt + \mathbf{e}_p \phi_t dW_t,$$

where $\phi_t = \zeta_t \sigma_t$. The deterministic seasonal function ζ_t^2 is given by a truncated Fourier series of order four, having a yearly seasonality. For the stochastic volatility process σ_t, the Barndorff-Nielsen and Shephard (BNS) model [3] is used.

$$\sigma_t^2 \triangleq V_t,$$

with

$$dV_t = -\lambda V_t dt + dL_t,$$

where $\lambda > 0$ and L_t is assumed to be a subordinator independent of W_t.
[6] propose that

$$T_t = \Lambda_t + Y_t,$$

where

$$\Lambda_t = a + bt + c \sin(\omega(t - d)),$$

and the deseasonalized temperature Y_t is given by $Y_t = \mathbf{e}_1^{'} \mathbf{X}_t = X_t^1$.

Swishchuk and Cui (2013) [38] extend the model proposed by [8] to the CAR model driven by a Lévy process. [38] propose that

$$d\mathbf{X}_t = A\mathbf{X}_t dt + \mathbf{e}_p \sigma_t dL_t,$$

where L_t is a Lévy process. [38] propose that

$$T_t = \Lambda_t + Y_t,$$

where

$$\Lambda_t = a_0 + a_1 \sin(\omega(t - a_2)),$$

and $Y_t = \mathbf{e}_1^{'} \mathbf{X}_t = X_t^1$. It is suggested that

$$\sigma_t^2 = b_1 + \sum_{k=1}^{N} b_{2k} \sin(\omega k t) + \sum_{k=1}^{N} b_{2k+1} \cos(\omega k t),$$

with $N = 1$.

Elias et al. (2014) [18] suggest the model given by

$$T_t = S_t + X_t,$$

where

$$S_t = a + bt + c \sin(\omega(t + \varphi)), \tag{2.3}$$

is the deterministic annual seasonality component and X_t is the deseasonalized temperature. A two-state regime-switching model for X_t is represented by

$$X_t = \begin{cases} X_{t,1}, & \text{if } X_t \text{ is in regime } 1 \text{ with probability } p_1, \\ X_{t,2}, & \text{if } X_t \text{ is in regime } 2 \text{ with probability } p_2, \end{cases}$$

where $p_1 + p_2 = 1$. For the deseasonalized temperature X_t, different forms of two-state regime-switching models are considered. [18] state that the model where one regime is governed by a mean-reverting process and the other by a Brownian motion captures the temperature dynamics more accurately than the other models considered. In fact, the model suggested is given by

$$dX_{t,1} = \kappa(\alpha - X_{t,1})dt + \sigma_1 dW_t,$$

$$dX_{t,2} = \mu_2 dt + \sigma_2 dW_t,$$

where κ is the speed and α is the mean of the mean-reverting process, W_t is a Brownian motion, μ_2 is the mean of the Brownian motion, σ_1 and σ_2 are the volatilities.

2.2 The Newly Proposed Model

In the literature, the Ornstein-Uhlenbeck mean-reverting process in different forms is commonly used for modeling temperature. However, Elias et al. (2014) [18] examines regime-switching behavior of the temperature. Abrupt changes in temperature, caused by a combination of several factors including latitude, intensity of solar circulation, land and water surface areas, ocean currents, elevation, and clear skies, induce the regime-switching behavior in temperature ([18]). In the following, motivated by [18], a new regime-switching model for the temperature dynamics will be proposed.

We consider a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let the time interval be $[0, \mathsf{T}]$, where $\mathsf{T} < \infty$. We consider a continuous-time, homogeneous, finite-state Markov chain $\zeta := (\zeta_t : t \in [0, \mathsf{T}])$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$, with a state space $\mathbb{S}_\zeta = \{s_1, \dots, s_N\}$. Without loss of generality, we adopt the canonical state space representation of the chain in [19] and identify the state space of the chain with a set of the standard unit vectors $\mathcal{E} = \{e_1, \dots, e_N\} \subset \mathbb{R}^N$, where the kth component of e_j is the Kronecker delta δ_{jk}, for each $j, k = 1, \dots, N$.

Let $\mathbf{A} = (a_{jl})_{j,l=1,\dots,N}$ be the rate matrix of the chain ζ, where a_{jl} is the transition intensity of the chain ζ from state e_l to state e_j. Note that for each $j, l = 1, \dots, N$, we have $a_{jl} \geq 0$ with $j \neq l$, and $a_{ll} = -\sum_{j=1, j\neq l}^{N} a_{jl}$. We suppose that $a_{jl} > 0$, for all $j, l = 1, \dots, N$, with $j \neq l$. With the canonical state space representation of the chain, we have the following semimartingale representation for the chain ζ, given in Elliott et al. (1994) [19]:

$$\zeta_t = \zeta_0 + \int_0^t \mathbf{A}\zeta_{s-} ds + V_t, \qquad t \in [0, \mathsf{T}], \tag{2.4}$$

where $(V_t : t \in [0, \mathsf{T}])$ is an $\mathbb{R}^N$-valued $(\mathbf{F}^\zeta, \mathbb{P})$-martingale. Here, $\mathbf{F}^\zeta := (\mathcal{F}_t^\zeta : t \in [0, \mathsf{T}])$ is the right-continuous, $\mathbb{P}$-complete natural filtration generated by the chain ζ.

In the following, $\langle X, Y \rangle = X'Y$ denotes the inner product of $X, Y \in \mathbb{R}^N$. Then, $X^j = \langle X, e_j \rangle$ is the jth element of $X \in \mathbb{R}^N$.

Now we consider a set of jump processes associated with the chain ζ. For each $l, j = 1, \dots, N$, let $\mathcal{N}^{lj} := (\mathcal{N}_t^{lj} : t \in [0, \mathsf{T}])$, where $\mathcal{N}_t^{lj}$ denotes the number of jumps of the chain ζ from state e_l to e_j in $[0, t]$, for each $t \in [0, \mathsf{T}]$ and $l, j = 1, \dots, N$. Then,

$$\begin{aligned}
\mathcal{N}_t^{lj} &:= \sum_{0<s\leq t} \langle \zeta_{s-}, e_l \rangle \langle \zeta_s, e_j \rangle \\
&= \sum_{0<s\leq t} \langle \zeta_{s-}, e_l \rangle \langle \Delta\zeta_s, e_j \rangle \\
&= \int_0^t \langle \zeta_{s-}, e_l \rangle \langle d\zeta_s, e_j \rangle \\
&= \int_0^t \langle \zeta_{s-}, e_l \rangle \langle \mathbf{A}\zeta_{s-}, e_j \rangle ds + \int_0^t \langle \zeta_{s-}, e_l \rangle \langle dV_s, e_j \rangle \\
&= \int_0^t a_{jl} \langle \zeta_{s-}, e_l \rangle ds + \mathcal{M}_t^{lj},
\end{aligned}$$

where

$$\mathcal{M}_t^{lj} := \int_0^t \langle \zeta_{s-}, e_l \rangle \langle dV_s, e_j \rangle.$$

Here, for each $l, j = 1, \dots, N$, $\mathcal{M}^{lj} := (\mathcal{M}_t^{lj} : t \in [0, \mathsf{T}])$ is an $(\mathbf{F}^\zeta, \mathbb{P})$-martingale.

For each $j = 1, \dots, N$, let $\mathcal{N}^j := (\mathcal{N}_t^j : t \in [0, \mathsf{T}])$, where $\mathcal{N}_t^j$ counts the number of jumps of the chain ζ into the state e_j from the other states in $[0, t]$. Then,

$$\begin{aligned} \mathcal{N}_t^j &:= \sum_{l=1, l \neq j}^{N} \mathcal{N}_t^{lj} \\ &= \sum_{l=1, l \neq j}^{N} \int_0^t a_{jl} \langle \zeta_{s-}, e_l \rangle ds + \mathcal{M}_t^j, \end{aligned} \tag{2.5}$$

where

$$\begin{aligned} \mathcal{M}_t^j &:= \sum_{l=1, l \neq j}^{N} \mathcal{M}_t^{lj} \\ &= \sum_{l=1, l \neq j}^{N} \int_0^t \langle \zeta_{s-}, e_l \rangle \langle dV_s, e_j \rangle. \end{aligned} \tag{2.6}$$

Here, for each $j = 1, \dots, N$, $\mathcal{M}^j := (\mathcal{M}_t^j : t \in [0, \mathsf{T}])$ is an $(\mathbf{F}^\zeta, \mathbb{P})$-martingale.

Thus, for each $j = 1, \dots, N$, the following representations follows:

$$\mathcal{N}_t^j = \int_0^t a_s^j ds + \mathcal{M}_t^j, \tag{2.7}$$

or

$$d\mathcal{N}_t^j = a_t^j dt + d\mathcal{M}_t^j, \tag{2.8}$$

where

$$a_t^j := \sum_{l=1, l \neq j}^{N} a_{jl} \langle \zeta_{t-}, e_l \rangle. \tag{2.9}$$

Let $T := (T_t : t \in [0, \mathsf{T}])$, where T_t is the temperature at time t. We suggest the temperature model given by

$$T_t = \Lambda_t + Y_t, \tag{2.10}$$

where Λ_t is the deterministic seasonal mean function and $Y := (Y_t : t \in [0, \mathsf{T}])$ is the deseasonalized temperature process. Here, Λ_t is assumed to be bounded and continuously differentiable, can be taken as

$$\Lambda_t = a_0 + a_1 t + a_2 \cos(\omega(t - a_3)), \tag{2.11}$$

where $\omega = 2\pi/365$. For the deseasonalized temperature Y_t, we propose a regime-switching model described below.

Let $\bar{S} = (S_1, \ldots, S_N)' \in \mathbb{R}^N$ and $\bar{\sigma} = (\sigma_1, \ldots, \sigma_N)' \in \mathbb{R}^N$ with $\sigma_j > 0$ for $j = 1, \ldots, N$. We define $\mathfrak{B} := (\beta_{jl})_{j,l=1,\ldots,N}$ with $\beta_{jj} = 0$, for $j = 1, \ldots, N$. Let $\bar{\beta}^j = (\beta_{j1}, \ldots, \beta_{jN})' \in \mathbb{R}^N$, for $j = 1, \ldots, N$, that is, $\bar{\beta}^j$ is the transpose of the jth row of $\mathfrak{B}$. We define

$$S_{\zeta_t} := \langle \bar{S}, \zeta_t \rangle = \sum_{j=1}^{N} S_j \langle \zeta_t, e_j \rangle,$$

$$\sigma_{\zeta_t} := \langle \bar{\sigma}, \zeta_t \rangle = \sum_{j=1}^{N} \sigma_j \langle \zeta_t, e_j \rangle,$$

$$\beta^j_{\zeta_t} := \langle \bar{\beta}^j, \zeta_{t-} \rangle = \sum_{l=1}^{N} \beta_{jl} \langle \zeta_{t-}, e_l \rangle, \quad \text{for } j = 1, \ldots, N.$$

We propose that the dynamics of the deseasonalized temperature Y_t is given by the following regime-switching model

$$dY_t = \kappa \left(Y_t - S_{\zeta_t}\right) dt + \sigma_{\zeta_t} dW_t + \sum_{j=1}^{N} \beta^j_{\zeta_t} d\mathcal{N}^j_t. \tag{2.12}$$

Here, κ is the speed of mean reversion, assumed to be constant, S_{ζ_t} and σ_{ζ_t} are defined above, and $W := (W_t : t \in [0, \mathsf{T}])$ is a standard Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$. We assume that W and ζ are independent. For each $j = 1, \ldots, N$, $\mathcal{N}^j := (\mathcal{N}^j_t : t \in [0, \mathsf{T}])$ is the counting process associated with chain, where $\mathcal{N}^j_t$ denotes the number of jumps of the chain into the state e_j from the other states in $[0, t]$, with the representation of Equation (2.7). Since jumps in the temperature are directly related to the regime switch, we model them by the chain itself. The jumps can be considered as the shifts in the level of the temperature due to the transitions of the state of the atmospheric conditions. The jump size of the temperature is determined by $\beta^j_{\zeta_t}$, $j = 1, \ldots, N$, as defined above. Notice that the jump size depends on the states of the chain before and after a state transition, that is, when the chain jumps from the state e_l to the state e_j, the jump size of the temperature is given by β_{jl}.

For notational convenience, we simply represent the model proposed in Equation (2.12) as

$$dY_t = \kappa \left(Y_t - S_t\right) dt + \sigma_t dW_t + \sum_{j=1}^{N} \beta^j_t d\mathcal{N}^j_t, \tag{2.13}$$

where $S_t = S_{\zeta_t}$, $\sigma_t = \sigma_{\zeta_t}$ and $\beta^j_t = \beta^j_{\zeta_t}$. Let $\mathbf{F} := (\mathcal{F}_t : t \in [0, \mathsf{T}])$ be the right-continuous, $\mathbb{P}$-complete natural filtration generated by Y. We define for each $t \in [0, \mathsf{T}]$, $\mathcal{G}_t := \mathcal{F}_t \vee \mathcal{F}^\zeta_t$, which represents the enlarged σ-field generated by $\mathcal{F}_t$ and $\mathcal{F}^\zeta_t$. And write $\mathbf{G} := (\mathcal{G}_t : t \in [0, \mathsf{T}])$ for the corresponding complete enlarged filtration.

Now, by Itô's Formula, the deseasonalized temperature, given by Equation (2.13), with the initial value Y_s, is of the form

$$\begin{aligned} Y_t = e^{\kappa(t-s)}Y_s - \kappa \int_s^t e^{\kappa(t-u)} S_u du + \int_s^t e^{\kappa(t-u)} \sigma_u dW_u \\ + \int_s^t e^{\kappa(t-u)} \sum_{j=1}^N \beta_u^j d\mathcal{N}_u^j. \end{aligned} \tag{2.14}$$

Notice that by Equation (2.8), we can also write

$$\begin{aligned} Y_t = e^{\kappa(t-s)}Y_s - \kappa \int_s^t e^{\kappa(t-u)} S_u du + \int_s^t e^{\kappa(t-u)} \sum_{j=1}^N \beta_u^j a_u^j du \\ + \int_s^t e^{\kappa(t-u)} \sigma_u dW_u + \int_s^t e^{\kappa(t-u)} \sum_{j=1}^N \beta_u^j d\mathcal{M}_u^j. \end{aligned} \tag{2.15}$$

In the following we represent some useful results.

For $\mathbf{A} = (a_{jl})_{j,l=1,\dots,N}$, we denote $\mathbf{A}_0 := \mathbf{A} - \text{diag}[\mathbf{a}]$, where diag[$\mathbf{a}$] is the diagonal matrix generated by $\mathbf{a} = (a_{11}, \dots, a_{NN})' \in \mathbb{R}^N$. Also write $\mathbf{I}$ for the $N \times N$ identity matrix, $\mathbf{1} = (1, \dots, 1)' \in \mathbb{R}^N$ and $\mathbf{0} = (0, \dots, 0)' \in \mathbb{R}^N$.

Remark 2.1. Consider the semimartingale representation of the chain given by Equation (2.4). Then we have the following [17]:

1. $\mathbf{1}'\zeta_s = 1$ and $\mathbf{1}'\mathbf{A} = \mathbf{0}'$.
2. $(\mathbf{I} - \text{diag}[\zeta_s])\mathbf{A}\zeta_s = \mathbf{A}_0\zeta_s$.
3. $(\mathbf{I} - \zeta_s\mathbf{1}')(\mathbf{I} - \text{diag}[\zeta_s]) = \mathbf{I} - \zeta_s\mathbf{1}'$.
4. $(\mathbf{I} - \text{diag}[\zeta_s])(\mathbf{I} - \zeta_s\mathbf{1}') = \mathbf{I} - \text{diag}[\zeta_s]$.
5. $(\mathbf{I} - \zeta_s\mathbf{1}')\mathbf{A}_0\zeta_s = \mathbf{A}\zeta_s$.

Remark 2.2. As introduced by [17], the process $\mathcal{N} := (\mathcal{N}_t : \ t \in [0, \mathsf{T}])$ given by

$$\mathcal{N}_t = \int_0^t (\mathbf{I} - \text{diag}[\zeta_{s-}]) d\zeta_s \tag{2.16}$$

is a vector of counting process with $\mathcal{N}_t = (\mathcal{N}_t^1, \dots, \mathcal{N}_t^N)' \in \mathbb{R}^N$, where $\mathcal{N}_t^j$ counts the number of times ζ jumps to state e_j, $j = 1, \dots, N$. Note that $\mathbb{E}(\mathcal{N}_t) = \int_0^t \mathbf{A}_0\zeta_{s-}ds$. Thus, the process $\mathcal{M} := (\mathcal{M}_t : \ t \in [0, \mathsf{T}])$ given by

$$\mathcal{M}_t = \mathcal{N}_t - \int_0^t \mathbf{A}_0\zeta_{s-}ds \tag{2.17}$$

is a martingale with $\mathcal{M}_t = (\mathcal{M}_t^1, \dots, \mathcal{M}_t^N)' \in \mathbb{R}^N$, see Equation (2.7).

Therefore, by Remark 2.1 and Remark 2.2, the chain given by Equation (2.4) can also be represented as

$$\begin{aligned} \zeta_t &= \zeta_0 + \int_0^t (\mathbf{I} - \zeta_{s-}\mathbf{1}')d\mathcal{N}_s, \\ &= \zeta_0 + \int_0^t \mathbf{A}\zeta_{s-}ds + \int_0^t (\mathbf{I} - \zeta_{s-}\mathbf{1}')d\mathcal{M}_t. \end{aligned} \tag{2.18}$$

CHAPTER 3

ESTIMATION AND FORECAST OF THE MODEL

In this part, the estimation and forecast of the proposed model together with the various existing models is considered.

3.1 The Data

The daily temperature data (in degrees Celsius) for Chicago O'Hare International Airport, USA, over the period from 1 January 2001 to 31 December 2013 is considered. The temperature data is provided by the National Climatic Data Center (NCDC) [34]. February 29 is removed from the sample in each leap year, resulting in 4745 observations. The data consists of the daily maximum and minimum temperatures. The temperature futures are written on several temperature indices, where the temperature is defined to be the average of the minimum and maximum temperatures over one day. Hence, at day t, the maximum temperature is denoted by $(Tmax)_t$ and the minimum temperature is denoted by $(Tmin)_t$, and the daily average temperature is defined by $(Tav)_t = \frac{1}{2}((Tmax)_t + (Tmin)_t)$. We denote the daily average temperature at time t by T_t, and call it as the temperature.

Table 3.1: Descriptive statistics of the temperature.

Minimum	-21.95
Maximum	33.6
Range	55.55
Mode	22.5
Median	11.1
Mean	10.377
Standart deviation	10.891
Skewness	-0.23503
Kurtosis	2.0922
Jarque-Bera statistic	206.61
Jarque-Bera critical value	5.9804
Number of observations	4745

The descriptive statistics of the temperature are presented in Table 3.1. The temperature ranges from -21.95 to 33.6. In Figure 3.1, the temperature for the period of 2001-2013 is depicted.

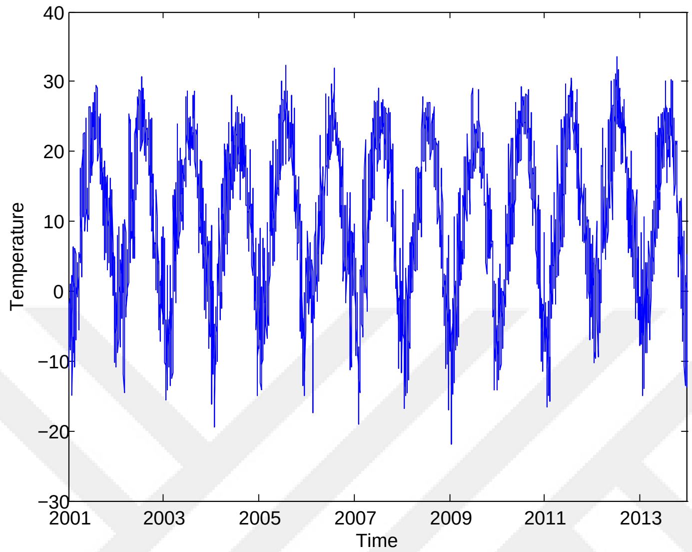


Figure 3.1: The temperature for the Chicago.

The mean, standard deviation, skewness, and kurtosis of the temperature for each day of the year are displayed in Figure 3.2 and Figure 3.3. The mean fluctuates between -6.83 and 26.02, with lowest in January and highest in August. The highest standard deviation is observed in the winter, while the lowest is observed in the summer. In Figure 3.4, the autocorrelation function (ACF) of the temperature is depicted.

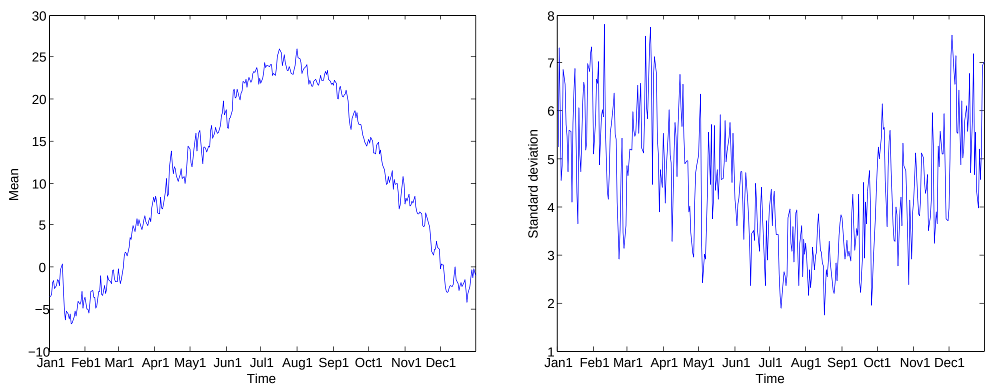


Figure 3.2: The mean and standard deviation of the temperature for each day of the year.

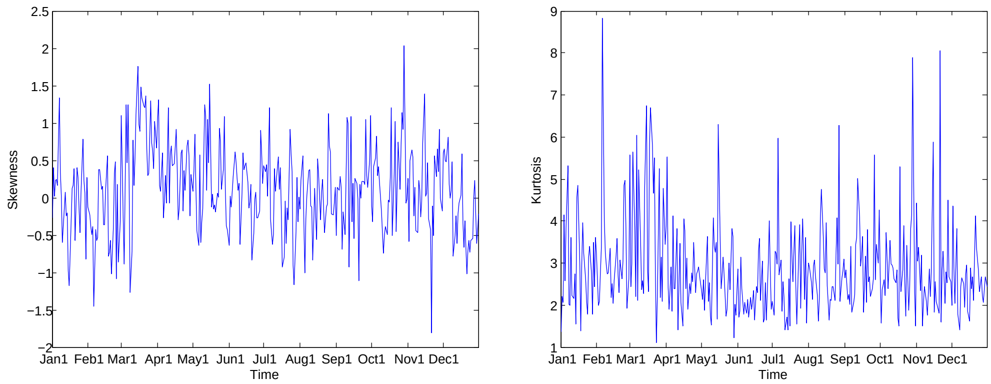


Figure 3.3: The skewness and kurtosis of the temperature for each day of the year.

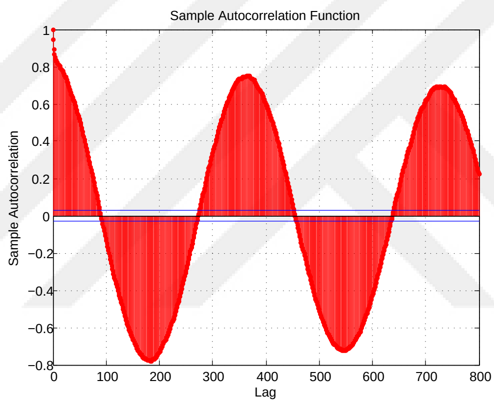


Figure 3.4: ACF of the temperature.

3.2 The Deseasonalized Temperature

We model the temperature T_t by

$$T_t = \Lambda_t + Y_t,$$

where Λ_t is the deterministic seasonal mean function and Y_t is the deseasonalized temperature.

$$\Lambda_t = a_0 + a_1 t + a_2 \cos(\omega(t - a_3)), \tag{3.1}$$

where $\omega = 2\pi/365$, is used to deseasonalize the temperature and the estimation results are presented at the Table 3.2.

Table 3.2: Estimation of Λ_t for the temperature.

	Estimate	Standard Error	t-statistic	p-value
a_0	10.113	0.13774	73.425	0
a_1	0.00011102	5.0293e-05	2.2074	0.027335
a_2	13.863	0.097268	142.52	0
a_3	-162.47	0.40817	-398.03	0
		RMSE: 4.74	Adj.R^2: 0.811	p-value: 0

In Figure 3.5 and Figure 3.6 the temperature with the fitted seasonality and deseasonalized temperature are depicted, respectively.

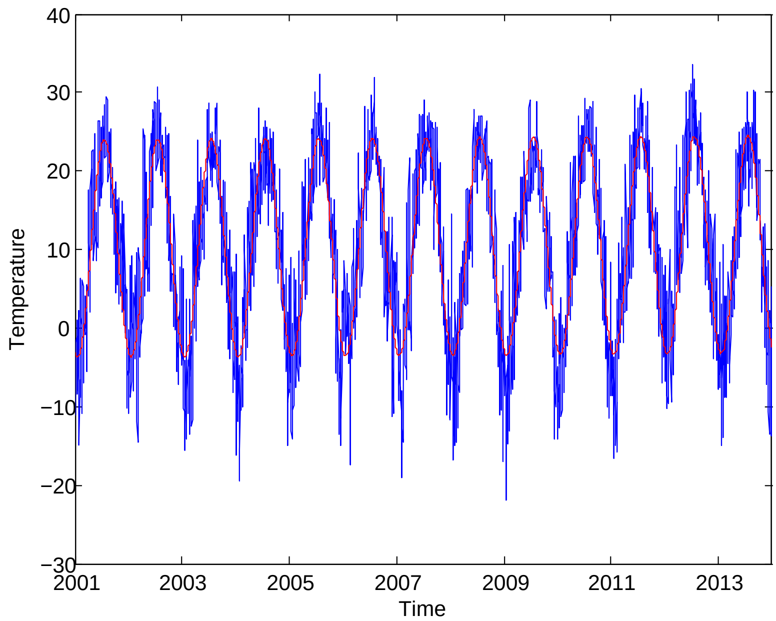


Figure 3.5: The temperature with the fitted seasonality.

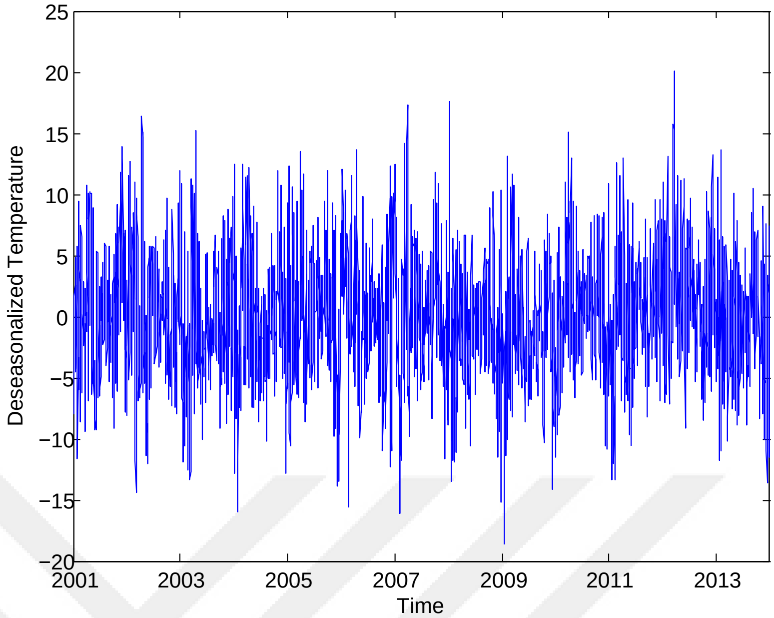


Figure 3.6: The deseasonalized temperature.

For the deseasonalized temperature Y_t, we consider various existing models and the proposed model. The last 3 years of data is reserved to be used for the forecast. Thus, the estimation of the models is based on the deseasonalized temperature between 1 January 2001 to 31 December 2010. And the deseasonalized temperature between 1 January 2011 and 31 December 2013 is used for the forecast of the models.

Before we estimate the models for the deseasonalized temperature, we proceed with the issues related with the estimation of regime-switching models.

3.3 Estimation of Regime-Switching Models

Estimation of regime-switching models necessitates inferring the model parameters and the states at the same time since the switching mechanism is unobservable. In the rest of this section, we consider an application of the Expectation-Maximization (EM) algorithm of Dempster et al. (1977) [13], which is given in Hamilton (1990) [25] and later refined by Kim (1994)[30]. We first introduce the issues related with the expectation step of the EM algorithm, then we give a detailed description of the EM algorithm.

3.3.1 Filtering

Consider an AR(1) model with first-order, N state Markov-switching mean and variance, that is,

$$y_t - \mu_{\zeta_t} = \phi_1(y_{t-1} - \mu_{\zeta_{t-1}}) + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma^2_{\zeta_t}),$$

with

$$\begin{aligned}\mu_{\zeta_t} &= \mu_1\zeta_{1t} + \cdots + \mu_N\zeta_{Nt},\\ \sigma^2_{\zeta_t} &= \sigma^2_1\zeta_{1t} + \cdots + \sigma^2_N\zeta_{Nt},\end{aligned}$$

where for $j = 1, \ldots, N$,

$$\zeta_{jt} = \begin{cases} 1, & \text{if } \zeta_t = j,\\ 0, & \text{otherwise.}\end{cases}$$

$$\mathbb{P}[\zeta_t = j|\zeta_{t-1} = i] = p_{ji}, \quad i, j = 1, \ldots, N,$$

$$\sum_{j=1}^{N} p_{ji} = 1.$$

Then, the conditional density of y_t given the past information $\mathcal{F}_{t-1}$, that is, $f(y_t|\mathcal{F}_{t-1})$ can be obtained by

$$\begin{aligned}f(y_t|\mathcal{F}_{t-1}) &= \sum_{j=1}^{N}\sum_{i=1}^{N} f(y_t, \zeta_t = j, \zeta_{t-1} = i|\mathcal{F}_{t-1})\\ &= \sum_{j=1}^{N}\sum_{i=1}^{N} f(y_t|\zeta_t = j, \zeta_{t-1} = i, \mathcal{F}_{t-1})\mathbb{P}[\zeta_t = j, \zeta_{t-1} = i|\mathcal{F}_{t-1}],\end{aligned}$$

where

$$f(y_t|\zeta_t = j, \zeta_{t-1} = i, \mathcal{F}_{t-1}) = \frac{1}{\sqrt{2\pi\sigma^2_j}} \exp\left(-\frac{(y_t - \mu_j - \phi_1(y_{t-1} - \mu_i))^2}{2\sigma^2_j}\right).$$

Thus, the log-likelihood function is given by

$$\begin{aligned}\ln L &= \sum_{t=1}^{\mathsf{T}} \ln\left(f(y_t|\mathcal{F}_{t-1})\right)\\ &= \sum_{t=1}^{\mathsf{T}} \ln\left(\sum_{j=1}^{N}\sum_{i=1}^{N} f(y_t|\zeta_t = j, \zeta_{t-1} = i, \mathcal{F}_{t-1})\mathbb{P}[\zeta_t = j, \zeta_{t-1} = i|\mathcal{F}_{t-1}]\right).\end{aligned}$$

Notice that $f(y_t|\mathcal{F}_{t-1})$ is a weighted average of N^2 conditional densities, weights being $\mathbb{P}[\zeta_t = j, \zeta_{t-1} = i|\mathcal{F}_{t-1}]$, for $i, j = 1, \ldots, N$, see the following.

The filtered probabilities refer to inferences about ζ_t conditional on information up to time t, that is, $\mathcal{F}_t$. The filtered probabilities, $\mathbb{P}[\zeta_t = j|\mathcal{F}_t]$, are obtained by the following filter, see [35]. The following two steps are iterated for $t = 1, \ldots, \mathsf{T}$, to obtain $\mathbb{P}[\zeta_t = j, \zeta_{t-1} = i|\mathcal{F}_{t-1}]$:

1. Given $\mathbb{P}[\zeta_{t-1} = i|\mathcal{F}_{t-1}],\ i = 1, \ldots, N$, at the beginning of time t, calculate

$$\mathbb{P}[\zeta_t = j, \zeta_{t-1} = i|\mathcal{F}_{t-1}] = \mathbb{P}[\zeta_t = j|\zeta_{t-1} = i]\mathbb{P}[\zeta_{t-1} = i|\mathcal{F}_{t-1}],$$

   where $\mathbb{P}[\zeta_t = j|\zeta_{t-1} = i]$ are the transition probabilities p_{ji}, for $i, j = 1, \ldots, N$.

2. Once y_t is observed at the end of time t, the probabilities are updated as follows:

$$\begin{aligned}
&\mathbb{P}[\zeta_t = j, \zeta_{t-1} = i|\mathcal{F}_t] \\
&= \mathbb{P}[\zeta_t = j, \zeta_{t-1} = i|\mathcal{F}_{t-1}, y_t] \\
&= \frac{f(y_t, \zeta_t = j, \zeta_{t-1} = i|\mathcal{F}_{t-1})}{f(y_t|\mathcal{F}_{t-1})} \\
&= \frac{f(y_t|\zeta_t = j, \zeta_{t-1} = i, \mathcal{F}_{t-1})\mathbb{P}[\zeta_t = j, \zeta_{t-1} = i|\mathcal{F}_{t-1}]}{\sum_{j=1}^{N}\sum_{i=1}^{N} f(y_t|\zeta_t = j, \zeta_{t-1} = i, \mathcal{F}_{t-1})\mathbb{P}[\zeta_t = j, \zeta_{t-1} = i|\mathcal{F}_{t-1}]}
\end{aligned}$$

   with

$$\mathbb{P}[\zeta_t = j|\mathcal{F}_t] = \sum_{i=1}^{N} \mathbb{P}[\zeta_t = j, \zeta_{t-1} = i|\mathcal{F}_t]. \tag{3.2}$$

To start the filter at time $t = 1$, the steady-state probabilities can be used. For a two-state, first-order Markov switching, the steady-state probabilities are given by

$$\pi_1 = \mathbb{P}[\zeta_0 = 1|\mathcal{F}_0] = \frac{1 - p_{22}}{2 - p_{22} - p_{11}}$$

and

$$\pi_2 = \mathbb{P}[\zeta_0 = 2|\mathcal{F}_0] = \frac{1 - p_{11}}{2 - p_{22} - p_{11}}.$$

3.3.2 Kim's Smoothing Algorithm

Remember that the filtered probabilities refer to inferences about ζ_t conditional on information up to time t, that is, $\mathcal{F}_t$. However, the smoothed probabilities refer to inferences about ζ_t conditional on all the information in the sample, that is, $\mathcal{F}_\mathsf{T}$. The smoothed probabilities, $\mathbb{P}[\zeta_t = j|\mathcal{F}_\mathsf{T}]$, are obtained as follows, see [35]:

$$\begin{aligned}
&\mathbb{P}[\zeta_t = j, \zeta_{t+1} = k|\mathcal{F}_\mathsf{T}] \\
&= \mathbb{P}[\zeta_{t+1} = k|\mathcal{F}_\mathsf{T}]\mathbb{P}[\zeta_t = j|\zeta_{t+1} = k, \mathcal{F}_\mathsf{T}] \\
&= \mathbb{P}[\zeta_{t+1} = k|\mathcal{F}_\mathsf{T}]\mathbb{P}[\zeta_t = j|\zeta_{t+1} = k, \mathcal{F}_t],
\end{aligned} \tag{3.3}$$

where for the last equality see [35]. Now, we can write

$$\begin{aligned}
&\mathbb{P}[\zeta_t = j|\zeta_{t+1} = k, \mathcal{F}_t] \\
&= \frac{\mathbb{P}[\zeta_t = j, \zeta_{t+1} = k|\mathcal{F}_t]}{\mathbb{P}[\zeta_{t+1} = k|\mathcal{F}_t]} \\
&= \frac{\mathbb{P}[\zeta_t = j|\mathcal{F}_t]\mathbb{P}[\zeta_{t+1} = k|\zeta_t = j]}{\mathbb{P}[\zeta_{t+1} = k|\mathcal{F}_t]}.
\end{aligned} \tag{3.4}$$

Then, by Equations (3.4) and (3.3), we have

$$\begin{aligned}
\mathbb{P}[\zeta_t = j|\mathcal{F}_\mathsf{T}] &= \sum_{k=1}^{N} \mathbb{P}[\zeta_t = j, \zeta_{t+1} = k|\mathcal{F}_\mathsf{T}] \\
&= \sum_{k=1}^{N} \frac{\mathbb{P}[\zeta_{t+1} = k|\mathcal{F}_\mathsf{T}]\mathbb{P}[\zeta_t = j|\mathcal{F}_t]\mathbb{P}[\zeta_{t+1} = k|\zeta_t = j]}{\mathbb{P}[\zeta_{t+1} = k|\mathcal{F}_t]}.
\end{aligned} \tag{3.5}$$

Thus, given $\mathbb{P}[\zeta_\mathsf{T} = j|\mathcal{F}_\mathsf{T}]$ at the last iteration of the filtering, Equation (3.5) is iterated for $t = \mathsf{T} - 1, \ldots, 1$, to obtain the smoothed probabilities $\mathbb{P}[\zeta_t = j|\mathcal{F}_\mathsf{T}]$, for $t = \mathsf{T} - 1, \ldots, 1$.

3.3.3 Expectation-Maximization Algorithm

The expectation-maximization (EM) algorithm, originally motivated by [13], is an alternative method for maximizing the likelihood function for models with missing observations or unobserved variables. The EM algorithm is an iterative procedure consisting of the following two steps at the $(m + 1)$th iteration [35]. Suppose that Θ is a vector of the model's unknown parameters.

1. Expectation Step (E-Step): Given the parameter estimates $\Theta^{(m)}$ obtained from the mth iteration, the expectation of the unobserved variables is formed.

2. Maximization Step (M-Step): Conditinal on the expectation of the unobserved variables, the likelihood function is maximized with respect to the parameters of the model, resulting in $\Theta^{(m+1)}$.

Each iteration results in a higher value of the likelihood function, and thus, with arbitrary initial values of the parameters $\Theta^{(0)}$, the above two steps are iterated until $\Theta^{(m+1)}$ converges.

In the following, we give the EM algorithm in detail. We consider the model

$$dY_t = (\alpha - \beta Y_t)\, dt + \sigma dW_t. \tag{3.6}$$

Discretized version of this model with swithching parameters is given by [28]

$$Y_t = \alpha_{\zeta_t} + (1 - \beta_{\zeta_t})\, Y_{t-1} + \sigma_{\zeta_t}\epsilon_t, \tag{3.7}$$

where $\epsilon_t \sim N(0,1)$. Here, ζ_t is an N state Markov chain with transition matrix $\mathbf{P} = (p_{ji})_{i,j=1,\dots,N}$, where $p_{ji} = \mathbb{P}[\zeta_t = j|\zeta_{t-1} = i]$. Let $\eta = (\alpha_j, \beta_j, \sigma_j)$, $\rho_j = \mathbb{P}[\zeta_1 = j|\mathcal{F}_0; \Theta]$ and $\Theta = (\eta, \mathbf{P}, \rho_j)$, for $j = 1, \dots, N$. The algorithm starts with an arbitrarily chosen vector of initial parameters $\Theta^{(0)} = (\eta^{(0)}, \mathbf{P}^{(0)}, \rho_j^{(0)})$, for $j = 1, \dots, N$.

3.3.3.1 The E-Step

The E-Step is composed of filtering and smoothing [30]. Suppose that the parameter vector calculated in the M-step during the previous iteration is given by $\Theta^{(m)} = (\eta^{(m)}, \mathbf{P}^{(m)}, \rho_j^{(m)})$. Then,

1. Filtering: For $t = 1, \dots, \mathsf{T}$, iterate on

$$\begin{aligned}&\mathbb{P}[\zeta_t = j|\mathcal{F}_t, \Theta^{(m)}]\\ &= \frac{\mathbb{P}[\zeta_t = j|\mathcal{F}_{t-1}, \Theta^{(m)}] f(y_t|\zeta_t = j, \mathcal{F}_{t-1}, \Theta^{(m)})}{\sum_{j=1}^{N} \mathbb{P}[\zeta_t = j|\mathcal{F}_{t-1}, \Theta^{(m)}] f(y_t|\zeta_t = j, \mathcal{F}_{t-1}, \Theta^{(m)})},\end{aligned}$$

   where

$$\begin{aligned}&f(y_t|\zeta_t = j, \mathcal{F}_{t-1}, \Theta^{(m)})\\ &= \frac{1}{\sqrt{2\pi}\sigma_j^{(m)}} \exp\left(-\frac{\left(y_t - (1-\beta_j^{(m)})y_{t-1} - \alpha_j^{(m)}\right)^2}{2(\sigma_j^{(m)})^2}\right)\end{aligned} \tag{3.8}$$

   and

$$\mathbb{P}[\zeta_{t+1} = j|\mathcal{F}_t, \Theta^{(m)}] = \sum_{i=1}^{N} p_{ji}^{(m)} \mathbb{P}[\zeta_t = i|\mathcal{F}_t, \Theta^{(m)}]$$

   until $\mathbb{P}[\zeta_{\mathsf{T}} = j|\mathcal{F}_{\mathsf{T}}, \Theta^{(m)}]$ is obtained. The starting point for the iteration is taken as $\mathbb{P}[\zeta_1 = j|\mathcal{F}_0, \Theta^{(m)}] = \rho_j^{(m)}$.

2. Smoothing: For $t = \mathsf{T} - 1, \dots, 1$, iterate on

$$\begin{aligned}&\mathbb{P}[\zeta_t = i|\mathcal{F}_{\mathsf{T}}, \Theta^{(m)}]\\ &= \sum_{j=1}^{N} \frac{\mathbb{P}[\zeta_{t+1} = j|\mathcal{F}_{\mathsf{T}}, \Theta^{(m)}]\mathbb{P}[\zeta_t = i|\mathcal{F}_t, \Theta^{(m)}] p_{ji}^{(m)}}{\mathbb{P}[\zeta_{t+1} = j|\mathcal{F}_t, \Theta^{(m)}]}.\end{aligned} \tag{3.9}$$

3.3.3.2 The M-Step

Since each observation y_t belongs to the jth regime with probability $\mathbb{P}[\zeta_t = j|\mathcal{F}_{\mathsf{T}}, \Theta^{(m)}]$, the maximum likelihood estimates are obtained maximizing the log-likelihood function of the form

$$\ln\left(L(\eta^{(m+1)})\right) = \sum_{j=1}^{N}\sum_{t=1}^{\mathsf{T}} \mathbb{P}[\zeta_t = j|\mathcal{F}_{\mathsf{T}}, \Theta^{(m)}] \ln\left(f(y_t|\zeta_t = j, \mathcal{F}_{t-1}, \eta^{(m+1)})\right). \tag{3.10}$$

The explicit formulas for the estimates $\hat{\beta}_j$, $\hat{\alpha}_j$, and $\hat{\sigma}_j$ are given by

$$\hat{\beta}_j = \frac{\sum_{t=2}^{\mathsf{T}} \mathbb{P}[\zeta_t = j|\mathcal{F}_{\mathsf{T}}, \Theta^{(m)}] y_{t-1} B_1}{\sum_{t=2}^{\mathsf{T}} \mathbb{P}[\zeta_t = j|\mathcal{F}_{\mathsf{T}}, \Theta^{(m)}] y_{t-1} B_2},$$

with

$$B_1 = y_t - y_{t-1} - \frac{\sum_{t=2}^{\mathsf{T}} \mathbb{P}[\zeta_t = j|\mathcal{F}_{\mathsf{T}}, \Theta^{(m)}](y_t - y_{t-1})}{\sum_{t=2}^{\mathsf{T}} \mathbb{P}[\zeta_t = j|\mathcal{F}_{\mathsf{T}}, \Theta^{(m)}]},$$

$$B_2 = \frac{\sum_{t=2}^{\mathsf{T}} \mathbb{P}[\zeta_t = j|\mathcal{F}_{\mathsf{T}}, \Theta^{(m)}] y_{t-1}}{\sum_{t=2}^{\mathsf{T}} \mathbb{P}[\zeta_t = j|\mathcal{F}_{\mathsf{T}}, \Theta^{(m)}]} - y_{t-1},$$

and

$$\hat{\alpha}_j = \frac{\sum_{t=2}^{\mathsf{T}} \mathbb{P}[\zeta_t = j|\mathcal{F}_{\mathsf{T}}, \Theta^{(m)}](y_t - (1-\hat{\beta}_j) y_{t-1})}{\sum_{t=2}^{\mathsf{T}} \mathbb{P}[\zeta_t = j|\mathcal{F}_{\mathsf{T}}, \Theta^{(m)}]},$$

$$\hat{\sigma}_j^2 = \frac{\sum_{t=2}^{\mathsf{T}} \mathbb{P}[\zeta_t = j|\mathcal{F}_{\mathsf{T}}, \Theta^{(m)}](y_t - \hat{\alpha}_j - (1-\hat{\beta}_j) y_{t-1})^2}{\sum_{t=2}^{\mathsf{T}} \mathbb{P}[\zeta_t = j|\mathcal{F}_{\mathsf{T}}, \Theta^{(m)}]}.$$

Moreover, $\rho_j^{(m+1)} = \mathbb{P}[\zeta_1 = j|\mathcal{F}_{\mathsf{T}}, \Theta^{(m)}]$, see [25]. Furthermore, the transition probabilities are given by [30]

$$\begin{aligned} p_{ji}^{(m+1)} &= \frac{\sum_{t=2}^{\mathsf{T}} \mathbb{P}[\zeta_t = j, \zeta_{t-1} = i|\mathcal{F}_{\mathsf{T}}, \Theta^{(m)}]}{\sum_{t=2}^{\mathsf{T}} \mathbb{P}[\zeta_{t-1} = i|\mathcal{F}_{\mathsf{T}}, \Theta^{(m)}]} \\ &= \frac{\sum_{t=2}^{\mathsf{T}} \mathbb{P}[\zeta_t = j|\mathcal{F}_{\mathsf{T}}, \Theta^{(m)}] \frac{p_{ji}^{(m)} \mathbb{P}[\zeta_{t-1}=i|\mathcal{F}_{t-1}, \Theta^{(m)}]}{\mathbb{P}[\zeta_t = j|\mathcal{F}_{t-1}, \Theta^{(m)}]}}{\sum_{t=2}^{\mathsf{T}} \mathbb{P}[\zeta_{t-1} = i|\mathcal{F}_{\mathsf{T}}, \Theta^{(m)}]}, \end{aligned} \tag{3.11}$$

where $p_{ji}^{(m)}$ is the transition probability $\mathbb{P}[\zeta_{t+1} = j|\zeta_t = i]$ obtained in the previous iteration. Then, all values obtained in the M-step are used as a new parameter vector $\Theta^{(m+1)} = (\eta^{(m+1)}, \mathbf{P}^{(m+1)}, \rho_j^{(m+1)})$, $j = 1, \ldots, N$, in the next iteration of the E-step.

3.4 Estimation of the Existing Models

In this section, we consider various existing models to model the deseasonalized temperature. For the estimation of the regime-switching models and threshold autoregressive model model, the codes in [27] and [20] are utilised, respectively.

3.4.1 The Regime-Switching Model

We consider the regime-switching (RS) model [28]

$$y_t = \alpha_{\zeta_t} + \phi_{\zeta_t} y_{t-1} + \sigma_{\zeta_t}\epsilon_t, \tag{3.12}$$

where $\epsilon_t \sim N(0,1)$. Here, ζ_t is an N state Markov chain with transition matrix $\mathbf{P} = (p_{ji})_{i,j=1,\dots,N}$, where $p_{ji} = \mathbb{P}[\zeta_t = j|\zeta_{t-1} = i]$. The estimation the RS model is conducted via the EM algorithm. The estimation results of the 2-state RS model for the deseasonalized temperature are presented at the Table 3.3.

Table 3.3: Estimation of the RS model for the deseasonalized temperature.

	α_j	ϕ_j	σ_j^2	p_{jj}
State1	-0.30393	0.60692	12.53070	0.75557
State2	0.60838	0.98778	4.09853	0.50202

After the estimation of the 2-state RS model, the data is classified as being in state 1 or state 2, by using the commonly used rule based on the smoothed probabilities obtained from the EM algorithm. Accordingly, if the smoothed probability for the state 2 at time t is bigger than 0.5, that is, $\mathbb{P}[\zeta_t = 2|\mathcal{F}_\mathsf{T}] > 0.5$, then the data is considered as being in the state 2 at time t, otherwise the data is considered as being in the state 1 at time t. The Figure 3.7 depicts classification of the data being in the state 1 or state 2, together with the smoothed probabilities for the state 2, presented by $P(State2)$. For a closer inspection, Figure 3.8 shows the last 5 years part of the Figure 3.7.

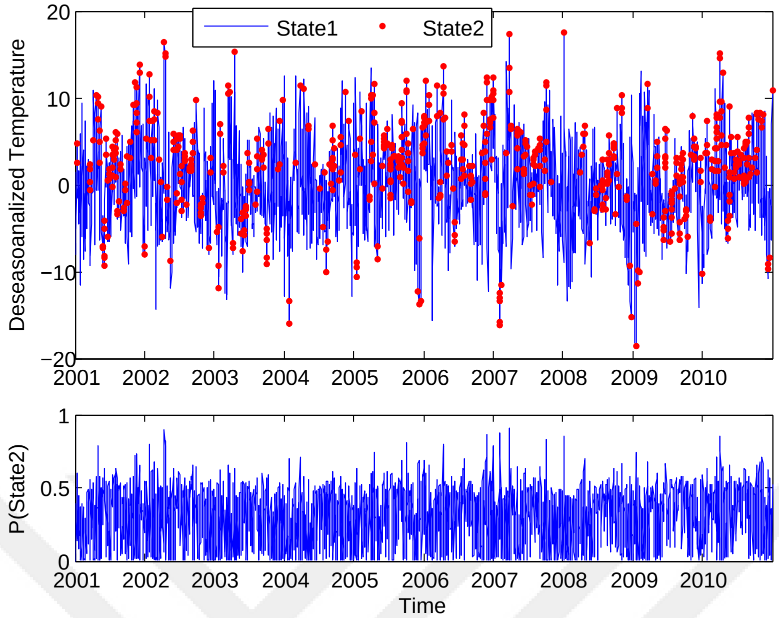


Figure 3.7: The RS model for the deseasonalized temperature.

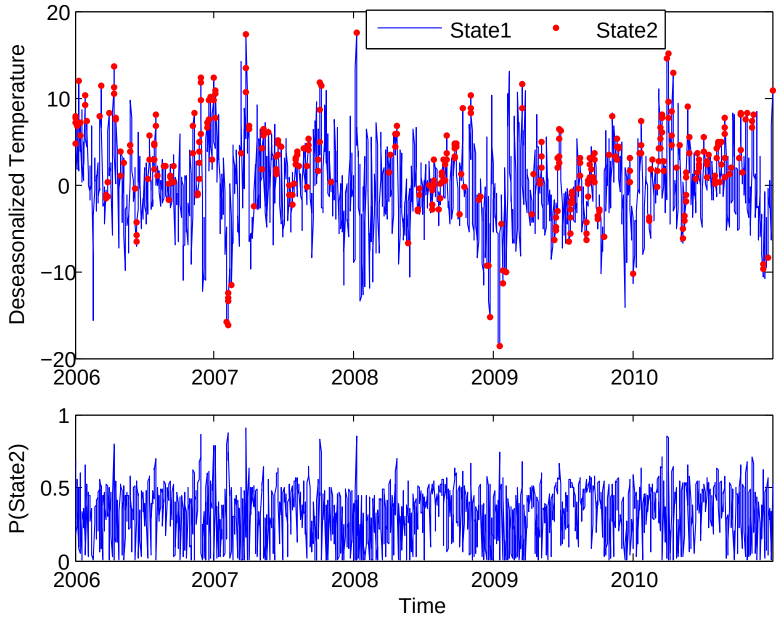


Figure 3.8: The RS model for the deseasonalized temperature (showing last 5 years).

3.4.2 The Constant-Speed Regime-Switching Model

We consider the regime-switching model given by Equation (3.12) with the parameter ϕ does not depend on the chain, that is,

$$y_t = \alpha_{\zeta_t} + \phi y_{t-1} + \sigma_{\zeta_t}\epsilon_t. \tag{3.13}$$

We call this model as the constant-speed regime-switching (CRS) model. Since the parameter ϕ does not depend on the chain, by fitting an AR(1) model to the deseasonalized temperature, $\phi = 0.72280$ is obtained. The estimation the CSR model is conducted via the EM algorithm with $\phi = 0.72280$. The estimation results of the 2-state CRS model for the deseasonalized temperature are presented at the Table 3.4.

Table 3.4: Estimation of the CRS model for the deseasonalized temperature.

	α_j	ϕ	σ_j^2	p_{jj}
State 1	-1.28239	0.72280	9.71535	0.61346
State 2	1.03188	0.72280	8.91050	0.67930

After the estimation of the 2-state CRS model, the data is classified as being in state 1 or state 2, by using the commonly used rule mentioned in the previous subsection. The Figure 3.9 depicts classification of the data being in the state 1 or state 2, together with the smoothed probabilities for the state 2, presented by $P(State2)$. For a closer inspection, Figure 3.10 shows the last 5 years part of the Figure 3.9.

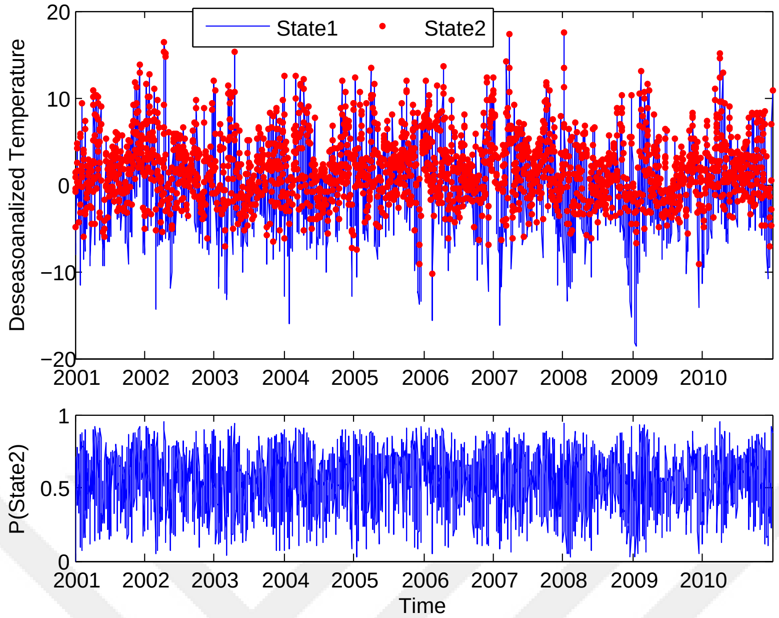


Figure 3.9: The CRS model for the deseasonalized temperature.

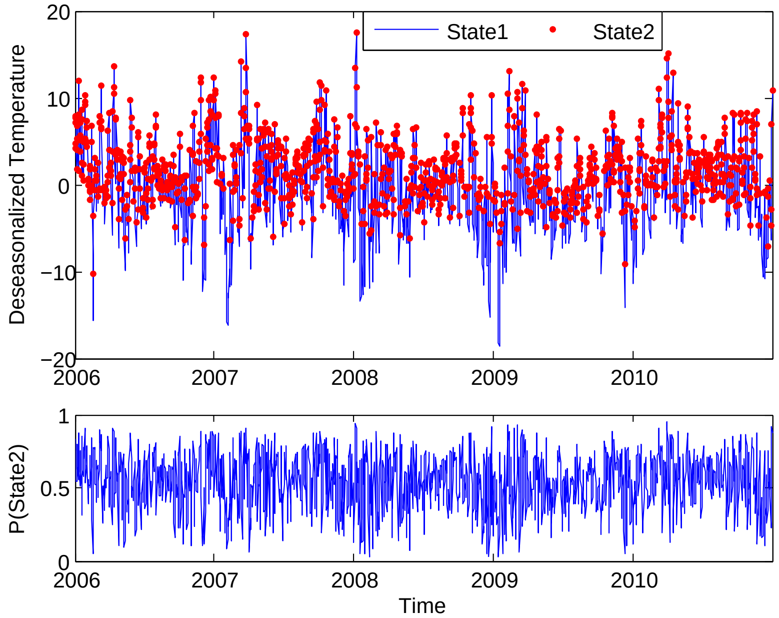


Figure 3.10: The CRS model for the deseasonalized temperature (showing last 5 years).

3.4.3 The Threshold Autoregressive Model

We consider the 2-state version of the threshold autoregressive (TAR) model proposed by Tong (1983) [40] given by

$$y_t = (1 - I_t)(\alpha_1 + \phi_1 y_{t-1}) + I_t(\alpha_2 + \phi_2 y_{t-1}) + \epsilon_t,$$

with

$$I_t = \begin{cases} 1, & \text{if } y_{t-1} \geq \tau, \\ 0, & \text{if } y_{t-1} < \tau, \end{cases}$$

where τ is the value of the threshold. The estimation results of the 2-state TAR model for the deseasonalized temperature are presented at the Table 3.5. The Figure 3.11 shows the deseasonalized temperature together with the TAR model threshold obtained from the estimation of the TAR model.

Table 3.5: Estimation of the TAR model for the deseasonalized temperature.

	α_j	ϕ_j	τ
State1	0.024384	0.74528	2.75
State2	0.71011	0.59864	2.75

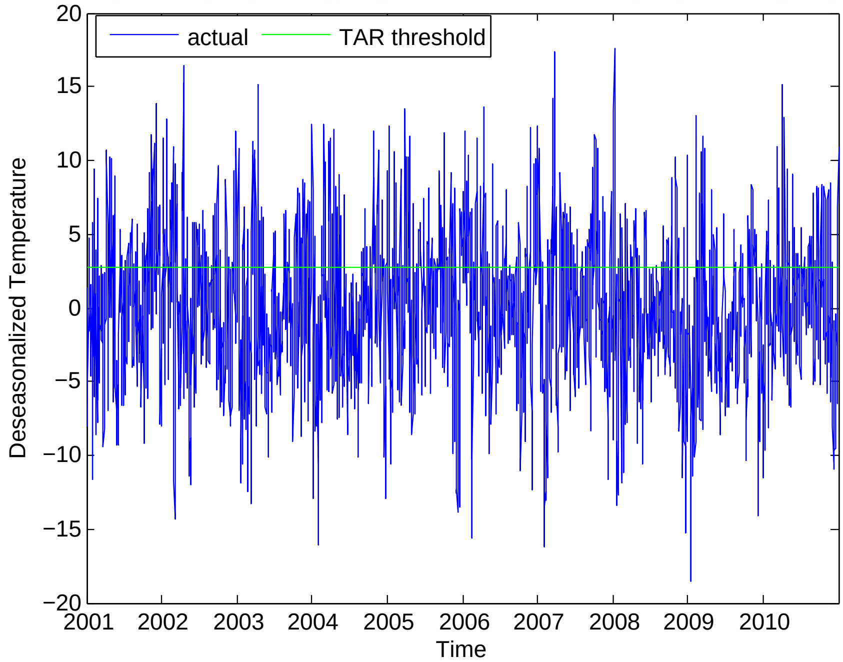


Figure 3.11: The TAR model for the deseasonalized temperature together with the TAR model threshold.

3.5 Estimation of the Newly Proposed Model

For the deseasonalized temperature Y_t, we consider discretized version of proposed model given by Equation (2.12). We consider the model given by

$$Y_t = -\kappa S_{\zeta_t} + (1+\kappa)Y_{t-1} + \sigma_{\zeta_t}\epsilon_t + \sum_{j=1}^{N} \beta^j_{\zeta_t} \Delta\mathcal{N}^j_t,$$

where $\epsilon_t \sim N(0,1)$ and $\Delta\mathcal{N}^j_t = \mathcal{N}^j_t - \mathcal{N}^j_{t-1}$. Here, ζ_t is an N state Markov chain with transition matrix $\mathbf{P} = (p_{ji})_{i,j=1,\dots,N}$, where $p_{ji} = \mathbb{P}[\zeta_t = j|\zeta_{t-1} = i]$, and $\mathcal{N}^j_t$ counts the jumps of the chain into the state j from the other states in time t. With the convention $\alpha_{\zeta_t} = -\kappa S_{\zeta_t}$ and $\phi = 1+\kappa$, we can write

$$Y_t = \alpha_{\zeta_t} + \phi Y_{t-1} + \sigma_{\zeta_t}\epsilon_t + \sum_{j=1}^{N} \beta^j_{\zeta_t} \Delta\mathcal{N}^j_t. \tag{3.14}$$

Figure 3.12 depicts daily mean deseasonalized temperature, that is, the normal of the each day of the year, and Figure 3.13 displays the difference from the daily mean deseasonalized temperature. The figures give an idea of how to define the regimes. For example, one can consider the cold and hot fronts, that is, we sometimes have hot winters due to the hot fronts.

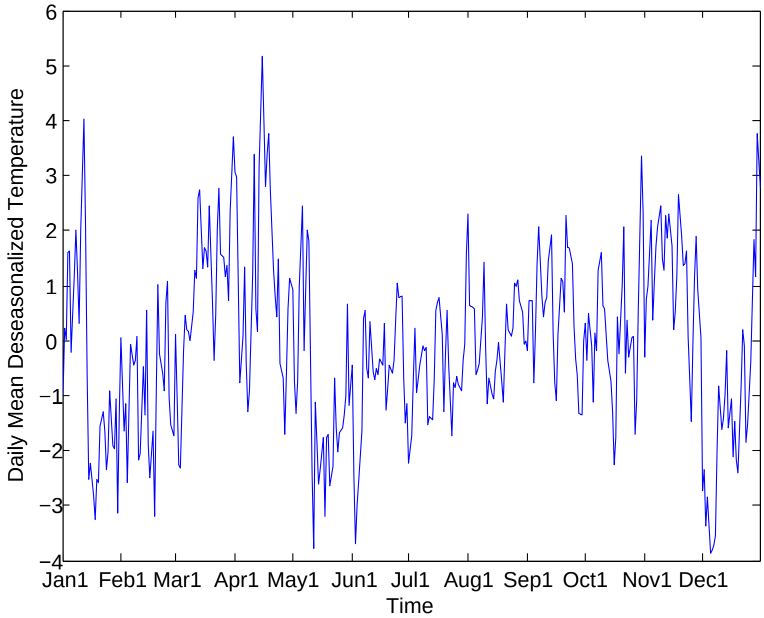


Figure 3.12: The daily mean deseasonalized temperature.

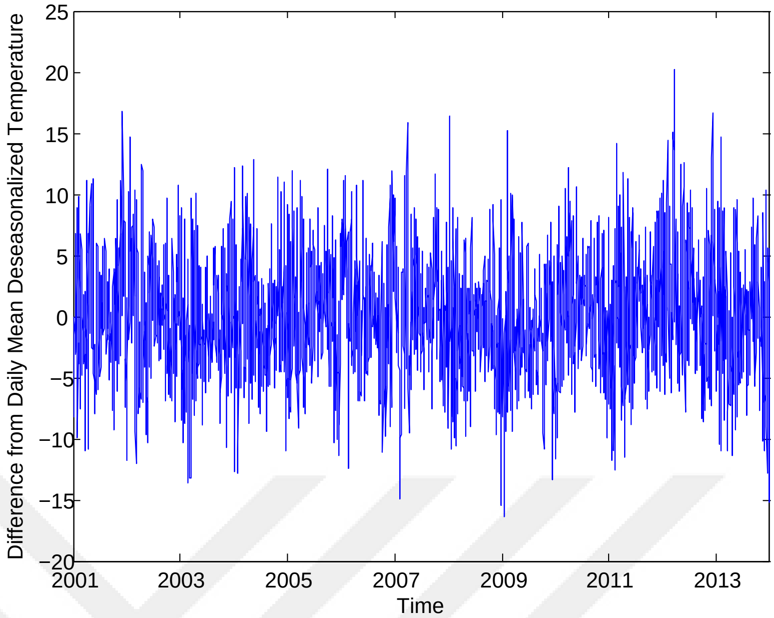


Figure 3.13: Difference from the daily mean deseasonalized temperature.

We consider the case $N = 2$. We denote the parameters by α_1, σ_1, and β_1 if we are in the state 1, and by α_2, σ_2, and β_2 if we are in the state 2. Notice that the jump part of Equation (3.14), $\sum_{j=1}^{2} \beta_{\zeta_t}^{j} \Delta \mathcal{N}_t^j$, equals β_1 if the regime changes from state 2 to state 1, equals β_2 if the regime changes from state 1 to state 2, and equals zero otherwise. It can be shown that we can write

$$\beta_2 = -\beta_1 = \frac{\phi(\alpha_2 - \alpha_1)}{1 - \phi},$$

and thus

$$\phi = \frac{\beta_2}{\alpha_2 - \alpha_1 + \beta_2}.$$

To estimate the model given by Equation (3.14), we modify the EM algorithm mentioned previously. We apply repeatedly the modified EM algorithm until the jump times and transition times coincide, see Figure 3.14. In the following we will call the model Equation (3.14) as the double regime-switching (DRS) model, with double referring to coincidence of jump and transition times. The estimation results of the 2-state DRS model for the deseasonalized temperature are presented at the Table 3.6.

Table 3.6: Estimation of the DRS model for the deseasonalized temperature.

	α_j	ϕ	σ_j^2	β_j	p_{jj}
State1	-0.66115	0.72107	8.92121	-6.55048	0.82458
State2	1.87273	0.72107	10.75911	6.55048	0.48448

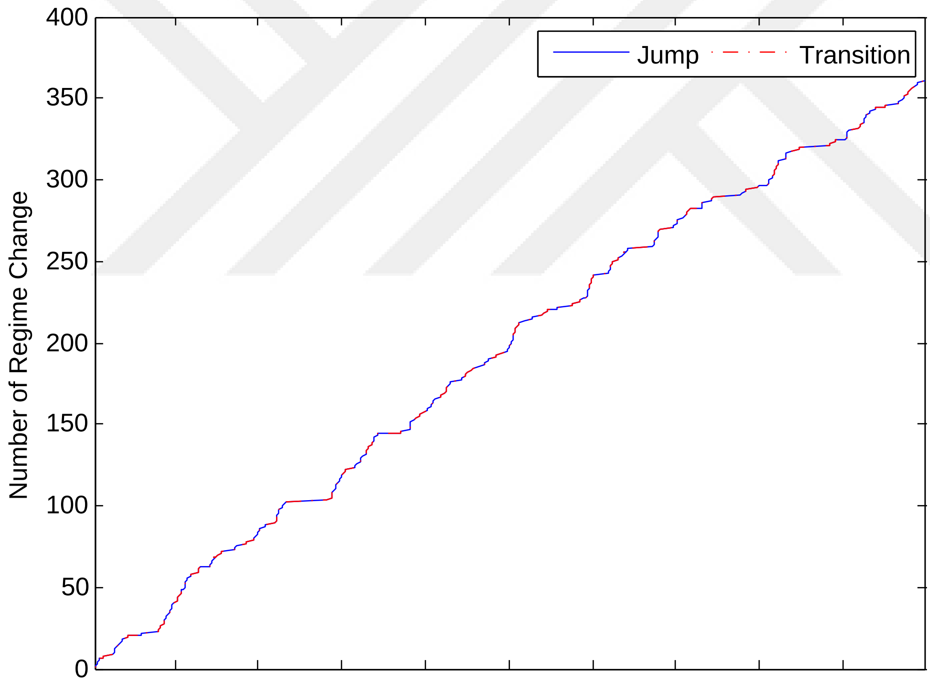


Figure 3.14: The jump time and the transition time.

Figure 3.15 depicts classification of the data being in the state 2 or state 1,, together with the smoothed probabilities for the state 2, presented by $P(State2)$. For a closer inspection, Figure 3.16 shows the last 5 years part of the Figure 3.15.

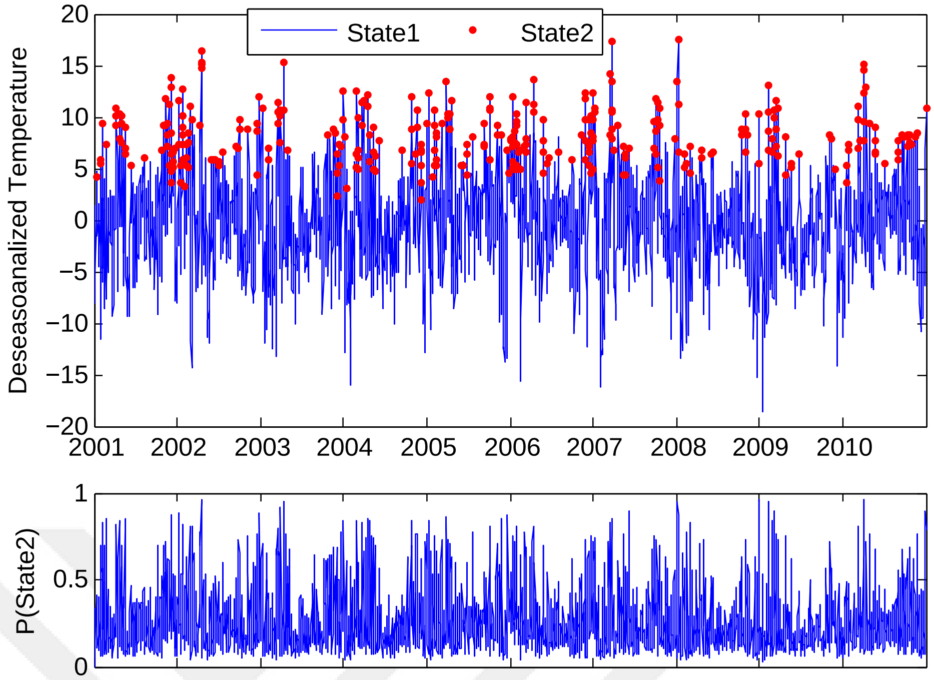


Figure 3.15: The DRS model for the deseasonalized temperature.

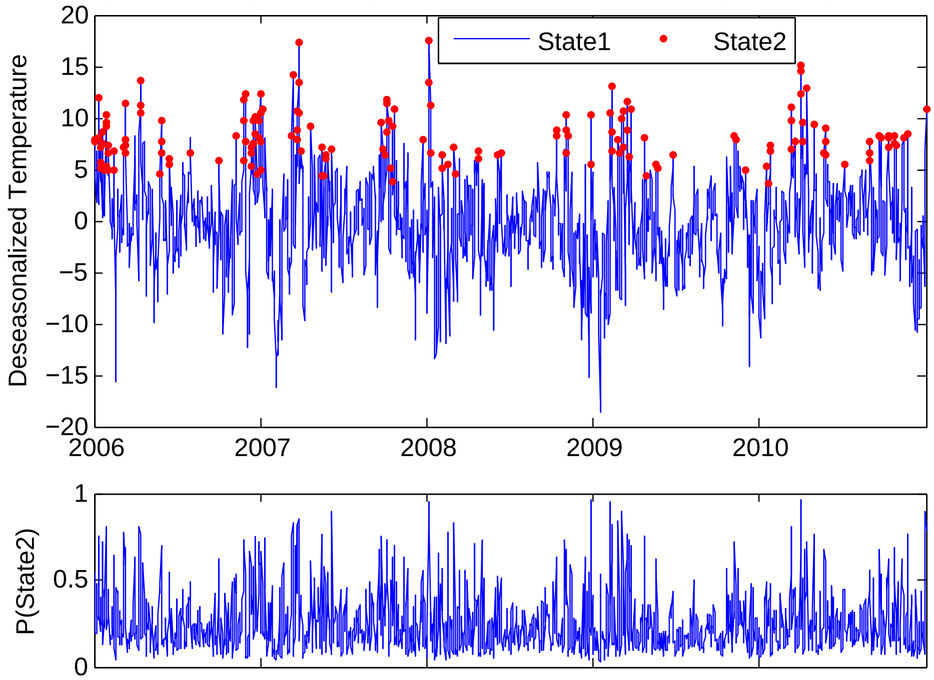


Figure 3.16: The DRS model for the deseasonalized temperature (showing last 5 years).

Notice that with the estimation of the DRS model, we are also able to obtain a threshold, called the DRS threshold, for the temperature such that we are in the state 2 at day t, if the temperature exceeds that day's normal temperature by an amount $b > 0$, otherwise we are in the state 1. We find that $b = 5.8292$. However, the DRS threshold is different from the constant threshold obtained from the TAR model. Our DRS threshold makes use of the smoothed probabilities and for each day of the year it is different. The obtained DRS threshold together with deseasoanalized temperature is presented in Figure 3.17.

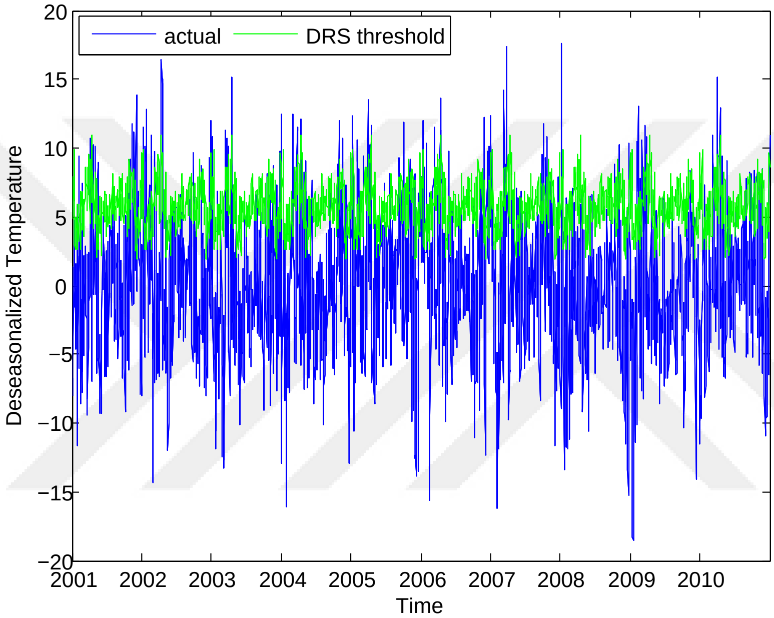


Figure 3.17: The DRS model for the deseasonalized temperature together with the DRS model threshold.

3.6 Forecast of the Newly Proposed Model

In this section, the 1-step ahead forecast of the proposed model together with the existing models will be given. Remember that for the forecast of the models, we use the deseasonalized temperature between 1 January 2011 and 31 December 2013.

The h-step ahead predicted probabilities can be calculated by

$$\mathbb{P}[\zeta_{t+h} = j|\mathcal{F}_t] = \mathbf{P}^h \mathbb{P}[\zeta_t = j|\mathcal{F}_t], \tag{3.15}$$

where $\mathbf{P} = (p_{ji})_{i,j=1,2}$, is the transition matrix of the chain with $p_{ji} = \mathbb{P}[\zeta_t = j|\zeta_{t-1} =$

$i]$, and $\mathbf{P}^h$ is the hth power of the transition matrix. For $h = 1$, we obtain

$$\mathbb{P}[\zeta_{t+1} = 1|\mathcal{F}_\mathrm{t}] = p_{11}\mathbb{P}[\zeta_t = 1|\mathcal{F}_\mathrm{t}] + p_{12}\mathbb{P}[\zeta_t = 2|\mathcal{F}_\mathrm{t}]$$

and

$$\mathbb{P}[\zeta_{t+1} = 2|\mathcal{F}_\mathrm{t}] = p_{21}\mathbb{P}[\zeta_t = 1|\mathcal{F}_\mathrm{t}] + p_{22}\mathbb{P}[\zeta_t = 2|\mathcal{F}_\mathrm{t}].$$

For the proposed DRS model, see Section 3.5, the 1-step ahead forecast of Y_t is calculated as

$$\begin{aligned}\mathbb{E}\left[Y_{t+1}|\mathcal{F}_t\right] = {} & (\alpha_1 + \phi Y_t)p_{11}\mathbb{P}[\zeta_t = 1|\mathcal{F}_\mathrm{t}] + (\alpha_2 + \phi Y_t)p_{22}\mathbb{P}[\zeta_t = 2|\mathcal{F}_\mathrm{t}] \\ & + (\alpha_2 + \beta_2 + \phi Y_t)p_{21}\mathbb{P}[\zeta_t = 1|\mathcal{F}_\mathrm{t}] + (\alpha_1 + \beta_1 + \phi Y_t)p_{12}\mathbb{P}[\zeta_t = 2|\mathcal{F}_\mathrm{t}].\end{aligned} \tag{3.16}$$

Moreover, we also consider the 1-step ahead forecasts of existing models considered in Section 3.4. 1-step ahead forecasts of RS, CRS, TAR and DRS models for the deseasonalized temperature are depicted from Figure 3.18 to Figure 3.21, respectively.

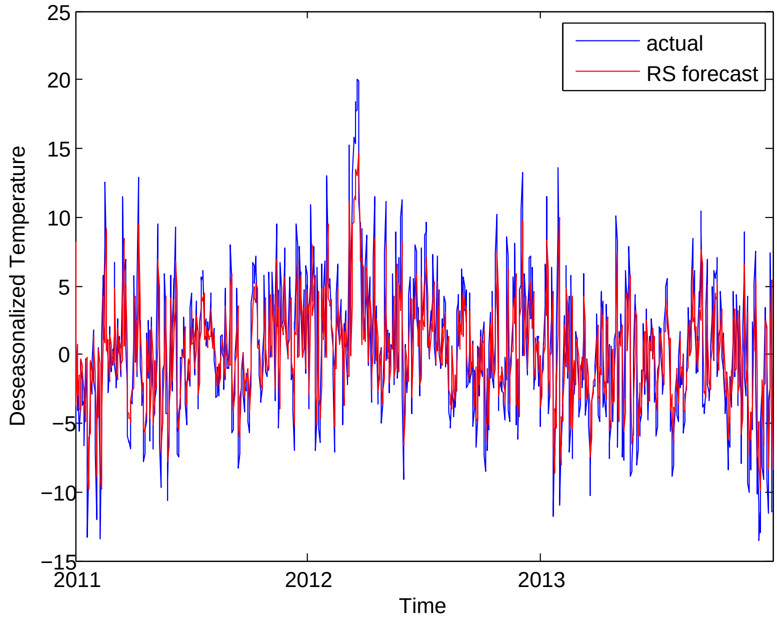


Figure 3.18: 1-step ahead forecast of the RS model for the deseasonalized temperature.

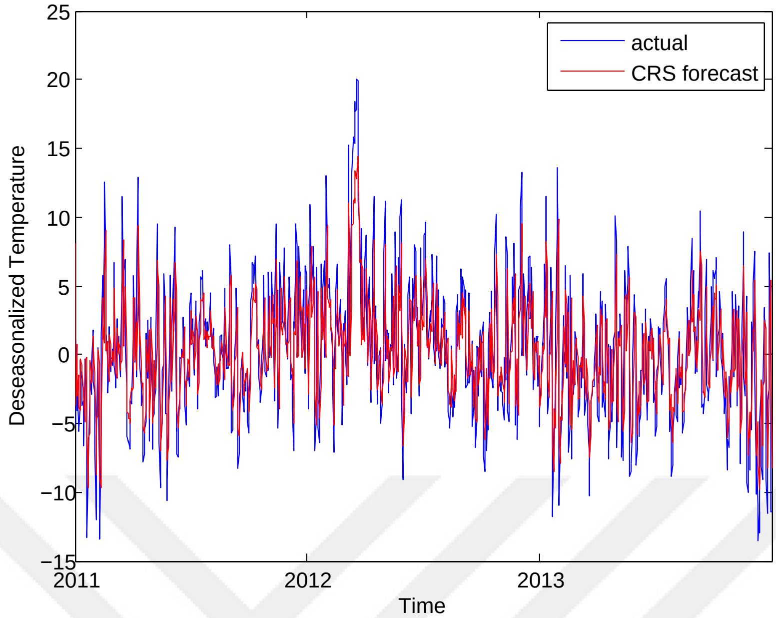


Figure 3.19: 1-step ahead forecast of the CRS model for the deseasonalized temperature.

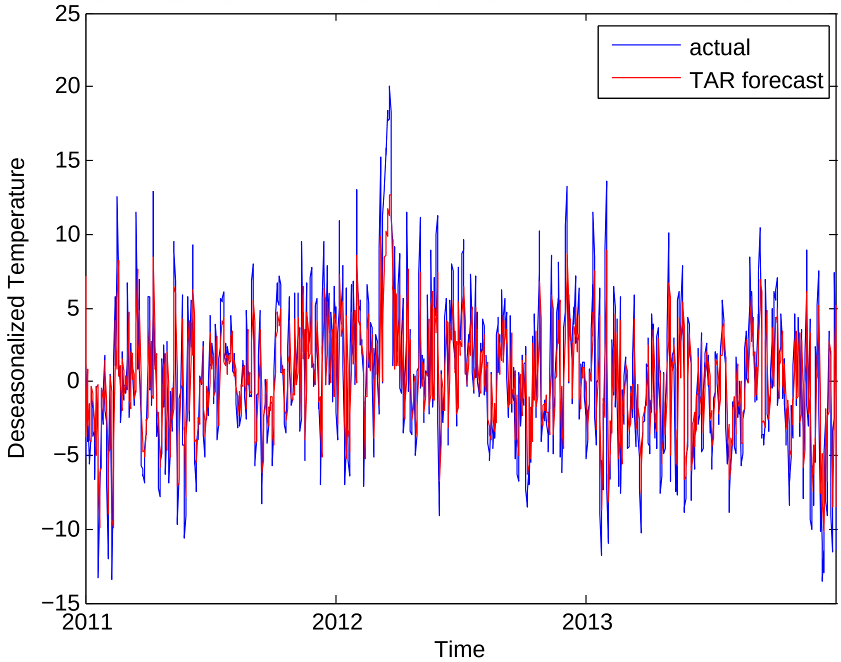


Figure 3.20: 1-step ahead forecast of the TAR model for the deseasonalized temperature.

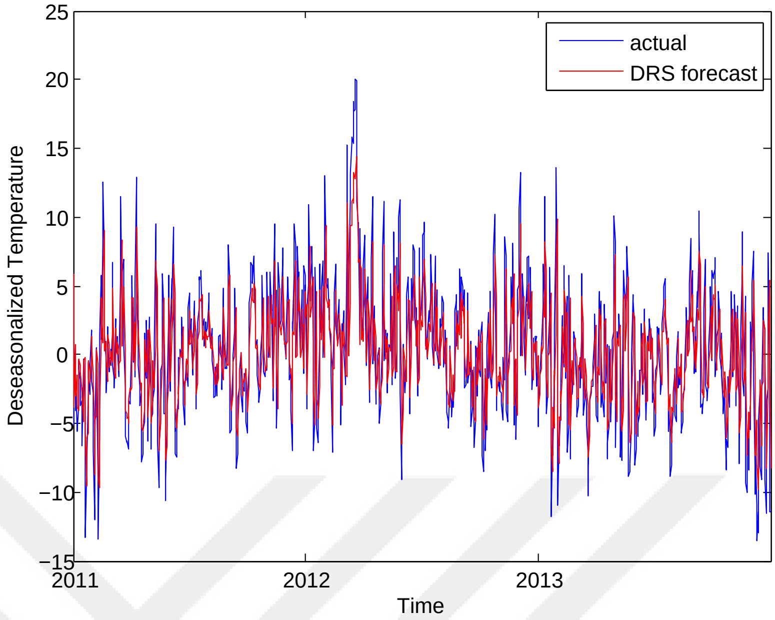


Figure 3.21: 1-step ahead forecast of the DRS model for the deseasonalized temperature.

We compare the forecast performance of the proposed DRS model with the existing models. The forecast error e_k is the difference between the actual value Y_k and the forecast value F_k, that is, $e_k = Y_k - F_k$, for $k = 1, \dots, M$. The most widely used forecast error measures are the Mean Square Error (MSE), Root Mean Square Error (RMSE) and Mean Absolute Error (MAE), given by [26]

$$MSE = \frac{1}{M}\sum_{k=1}^{M}(Y_k - F_k)^2 = \frac{1}{M}\sum_{k=1}^{M}(e_k)^2,$$

$$RMSE = \sqrt{\frac{1}{M}\sum_{k=1}^{M}(Y_k - F_k)^2} = \sqrt{\frac{1}{M}\sum_{k=1}^{M}(e_k)^2},$$

$$MAE = \frac{1}{M}\sum_{k=1}^{M}|Y_k - F_k| = \frac{1}{M}\sum_{k=1}^{M}|e_k|.$$

The forecast performance of the models are represented in Table 3.7. According to the table, DRS model has MSE of 10.896, while RS, CRS and TAR models have MSE of 10.926, 10.928 and 10.953, respectively. Moreover, DRS model has RMSE of 3.3009, while RS, CRS and TAR models have RMSE of 3.3055, 3.3057 and 3.3096, respectively. That is, the proposed DRS model has the smallest MSE and RMSE values among all the existing models considered. On the other hand, TAR model has MAE of 2.5431, while DRS, RS and CRS models have MAE of 2.5453, 2.5463 and 2.5476, respectively. The question is whether the differences in the respective MSE and MAE values are significant. In the following, we perform the Diebold-Mariano test, see [15], [14], to further assess the (relative) forecast performance of the models.

Table 3.7: Forecast performance of the models.

	RS model	CRS model	TAR model	DRS model
MSE	10.926	10.928	10.953	10.896
RMSE	3.3055	3.3057	3.3096	3.3009
MAE	2.5463	2.5476	2.5431	2.5453

The Diebold-Mariano test has the null hypothesis that two models have equal forecast accuracy and thus the difference of their forecast errors is not statistically significant. The alternative hypothesis is that one model produces better forecasts than the other. For both MSE and MAE criteria, we apply three times the Diebold-Mariano test with the alternative hypothesis that the proposed DRS model produces better forecasts than the RS, CRS and TAR models, respectively.

The results of the test are given in Table 3.8 and Table 3.9, respectively. According to the tables, we cannot reject the null hypothesis with the 5% significance level. Thus, we can say that none of the forecasting errors are significantly differrent from the other. Therefore, we can conclude that the forecast performance of the proposed DRS model is in line with the RS, CRS and TAR models.

Table 3.8: Diebold-Mariano test (MSE).

	Test Stat. (MSE)	p-value (MSE)
DRS model vs RS model	0.9013	0.18371
DRS model vs CRS model	1.1113	0.13321
DRS model vs TAR model	0.9325	0.17554

Table 3.9: Diebold-Mariano test (MAE).

	Test Stat. (MAE)	p-value (MAE)
DRS model vs RS model	0.3523	0.36229
DRS model vs CRS model	1.0671	0.14297
DRS model vs TAR model	-0.3320	0.63007

CHAPTER 4

PRICING WEATHER DERIVATIVES

At the Chicago Mercantile Exchange (CME), various weather derivatives are offered for trade. In 1999, the CME launched its first weather derivatives, which are the futures and options on temperature indices for several United States cities. Nowadays, the weather derivatives for various cities around the world are offered (see [12]). At the CME, in addition to the temperature futures written on several indices, and there are also call and put options written on these futures. In this part, the proposed model under the equivalent measure is considered and the prices of weather derivatives written on several temperature indices are derived.

4.1 The Newly Proposed Model Under the Equivalent Measure

In the following section, we will derive prices of temperature futures written on several temperature indices using the temperature model proposed in Section 2.2. To derive the futures prices, the risk-neutral probability $\mathbb{Q}$ is to be specified. Since temperature (and hence the index) is not a tradeable asset, any probability measure $\mathbb{Q}$ being equivalent to the objective probability is a risk-neutral probability. In the following, a generalized version of the Esscher transform is considered to select an equivalent measure by following [37].

Let $\mathfrak{L}(Y)$ be the space of all processes $\theta := \{\theta_t,\ t \in [0, \mathsf{T}]\}$ such that

1. For each $t \in [0, \mathsf{T}]$, $\theta_t := \langle \bar{\theta}, \zeta_t \rangle$, where $\bar{\theta} := (\theta_1, \dots, \theta_N)' \in \mathbb{R}^N$,
2. θ is integrable with respect to Y in the sense of stochastic integration.

For each $\theta \in \mathfrak{L}(Y)$, we define $(\theta \cdot Y) := ((\theta \cdot Y)_t :\ t \in [0, \mathsf{T}])$, where

$$(\theta \cdot Y)_t := \int_0^t \theta_s dY_s$$

is the stochastic integral of θ with respect to Y. In the following, θ is called as the Esscher transform parameter. For each $\theta \in \mathfrak{L}(Y)$, a $\mathbf{G}$-adapted process $D^\theta := (D_t^\theta :$

$t \in [0, \mathsf{T}])$ is defined in the following way.

$$D_t^\theta := \exp\left((\theta \cdot Y)_t\right).$$

Then, by Itô's Formula, we can write

$$\begin{aligned} D_t^\theta &= 1 + \int_0^t D_s^\theta \theta_s \kappa (Y_s - S_s) ds + \int_0^t D_s^\theta \theta_s \sigma_s dW_s \\ &+ \frac{1}{2} \int_0^t D_s^\theta \theta_s^2 \sigma_s^2 ds + \int_0^t D_{s-}^\theta \sum_{j=1}^N (e^{\theta_s \beta_s^j} - 1) d\mathcal{N}_s^j. \end{aligned}$$

Thus, we obtain

$$D_t^\theta = 1 + \int_0^t D_{s-}^\theta dU_s^\theta,$$

where $U^\theta := (U_t^\theta : \ t \in [0, \mathsf{T}])$ is given by

$$\begin{aligned} U_t^\theta &:= \int_0^t \theta_s \kappa (Y_s - S_s) ds + \frac{1}{2} \int_0^t \theta_s^2 \sigma_s^2 ds \\ &+ \int_0^t \theta_s \sigma_s dW_s + \int_0^t \sum_{j=1}^N (e^{\theta_s \beta_s^j} - 1) d\mathcal{N}_s^j \\ &= \int_0^t \theta_s \kappa (Y_s - S_s) ds + \frac{1}{2} \int_0^t \theta_s^2 \sigma_s^2 ds + \int_0^t \sum_{j=1}^N (e^{\theta_s \beta_s^j} - 1) a_s^j ds \\ &+ \int_0^t \theta_s \sigma_s dW_s + \int_0^t \sum_{j=1}^N (e^{\theta_s \beta_s^j} - 1) d\mathcal{M}_s^j. \end{aligned}$$

Thus, D^θ is the Doléans-Dade exponential of U^θ, that is,

$$D_t^\theta = \mathcal{E}(U_t^\theta), \qquad t \in [0, \mathsf{T}].$$

Consequently, the Laplace cumulant process (see [29]), $L^\theta := (L_t^\theta : \ t \in [0, \mathsf{T}])$ of the stochastic integral process $(\theta \cdot Y)$ is given by

$$L_t^\theta := \int_0^t \theta_s \kappa (Y_s - S_s) ds + \frac{1}{2} \int_0^t \theta_s^2 \sigma_s^2 ds + \int_0^t \sum_{j=1}^N (e^{\theta_s \beta_s^j} - 1) a_s^j ds,$$

which is the predictable finite variation part of U^θ. The Doléans-Dade exponential $\mathcal{E}(L_t^\theta)$ of L_t^θ is the unique solution of

$$\mathcal{E}(L_t^\theta) = 1 + \int_0^t \mathcal{E}(L_s^\theta) dL_s^\theta, \qquad t \in [0, \mathsf{T}].$$

Since L^θ is a finite variation process, we have

$$\mathcal{E}(L_t^\theta) = \exp(L_t^\theta).$$

Thus, for each $\theta \in \mathfrak{L}(Y)$, the logarithmic transform $\tilde{L}^\theta := (\tilde{L}_t^\theta : t \in [0, \mathsf{T}])$ is given by

$$\tilde{L}_t^\theta := log(\mathcal{E}(L_t^\theta)) = L_t^\theta, \qquad t \in [0, \mathsf{T}]. \tag{4.1}$$

Let $\mathcal{Z}^\theta := (\mathcal{Z}_t^\theta : t \in [0, \mathsf{T}])$ be a $\mathbf{G}$-adapted process associated with $\theta \in \mathfrak{L}(Y)$ as follows:

$$\mathcal{Z}_t^\theta := \exp\left((\theta \cdot Y)_t - \tilde{L}_t^\theta\right), \qquad t \in [0, \mathsf{T}].$$

Then by Equations (2.13) and (4.1), we get

$$\begin{aligned} \mathcal{Z}_t^\theta = \exp(&\int_0^t \theta_s \sigma_s dW_s - \frac{1}{2}\int_0^t \theta_s^2 \sigma_s^2 ds + \int_0^t \sum_{j=1}^N \theta_s \beta_s^j d\mathcal{M}_s^j \\ &- \int_0^t \sum_{j=1}^N (e^{\theta_s \beta_s^j} - 1 - \theta_s \beta_s^j) a_s^j ds). \end{aligned} \tag{4.2}$$

Then, $\mathcal{Z}^\theta$ is a $(\mathbf{G}, \mathbb{P})$-(local) martingale, see [37]. Notice that by Itô's Formula, we have

$$\mathcal{Z}_t^\theta = 1 + \int_0^t \mathcal{Z}_s^\theta \theta_s \sigma_s dW_s + \int_0^t \mathcal{Z}_{s-}^\theta \sum_{j=1}^N (e^{\theta_s \beta_s^j} - 1) d\mathcal{M}_s^j. \tag{4.3}$$

For each $\theta \in \mathfrak{L}(Y)$, we define a new probability measure $\mathbb{Q}^\theta$ equivalent to $\mathbb{P}$ on $\mathcal{G}_\mathsf{T}$ by a generalized version of the regime-switching Esscher transform as follows:

$$\frac{d\mathbb{Q}^\theta}{d\mathbb{P}}|_{\mathcal{G}_\mathsf{T}} := \mathcal{Z}_\mathsf{T}^\theta. \tag{4.4}$$

The following lemma follows from [37].

Lemma 4.1. *For each $t \in [0, \mathsf{T}]$, let*

$$W_t^\theta := W_t - \int_0^t \theta_s \sigma_s ds \tag{4.5}$$

and

$$\mathcal{M}_t^{\theta j} := \mathcal{N}_t^j - \int_0^t a_s^{\theta j} ds, \tag{4.6}$$

where

$$\begin{aligned} a_t^{\theta j} &:= e^{\theta_t \beta_t^j} a_t^j \\ &= \sum_{l=1, l \neq j}^N e^{\theta_l \beta_{jl}} a_{jl} \langle \zeta_{t-}, e_l \rangle. \end{aligned} \tag{4.7}$$

Then, $W^\theta := (W_t^\theta \,:\, t \in [0,\mathsf{T}])$ is a standard Brownian motion under $\mathbb{Q}^\theta$, and $\mathcal{M}^{\theta j} := (\mathcal{M}_t^{\theta j} \,:\, t \in [0,\mathsf{T}])$ is an $(\mathbf{F}^\zeta, \mathbb{Q}^\theta)$-martingale, for each $j = 1, \ldots, N$.

Furthermore, suppose $\mathbf{A}^\theta := (a_{jl}^\theta)_{j,l=1,\ldots,N}$, where

$$a_{jl}^\theta := \begin{cases} e^{\theta_l \beta_{jl}} a_{jl}, & \text{if} \quad l \neq j \\ \\ -\sum_{j=1, j\neq l}^N e^{\theta_l \beta_{jl}} a_{jl}, & \text{if} \quad l = j. \end{cases} \tag{4.8}$$

Then, the chain ζ has the following semimartingale decomposition under $\mathbb{Q}^\theta$.

$$\zeta_t = \zeta_0 + \int_0^t \mathbf{A}^\theta \zeta_{s-} ds + V_t^\theta, \qquad t \in [0,\mathsf{T}], \tag{4.9}$$

where $V^\theta := (V_t^\theta \,:\, t \in [0,\mathsf{T}])$ is an $\mathbb{R}^N$-valued $(\mathbf{F}^\zeta, \mathbb{Q}^\theta)$-martingale.

Note that, by the above lemma, the number of jumps of the chain ζ into the state e_j, for each $j = 1, 2, \ldots, N$, from the other states in $[0, t]$, denoted by $\mathcal{N}_t^j$, has the following representation under $\mathbb{Q}^\theta$:

$$\mathcal{N}_t^j = \sum_{l=1, l\neq j}^N \int_0^t a_{jl}^\theta \langle \zeta_{s-}, e_l \rangle ds + \mathcal{M}_t^{\theta j}, \tag{4.10}$$

where

$$\mathcal{M}_t^{\theta j} = \sum_{l=1, l\neq j}^N \int_0^t \langle \zeta_{s-}, e_l \rangle \langle dV_s^\theta, e_j \rangle. \tag{4.11}$$

Here, for each $j = 1, 2, \ldots, N$, $\mathcal{M}^{\theta j} := (\mathcal{M}_t^{\theta j} \,:\, t \in [0,\mathsf{T}])$ is an $(\mathbf{F}^\zeta, \mathbb{Q}^\theta)$-martingale. Thus, for each $j = 1, 2, \ldots, N$, under $\mathbb{Q}^\theta$ we have

$$d\mathcal{M}_t^{\theta j} = d\mathcal{N}_t^j - a_t^{\theta j} dt, \tag{4.12}$$

where

$$\begin{aligned} a_t^{\theta j} &= \sum_{l=1, l\neq j}^N a_{jl}^\theta \langle \zeta_{t-}, e_l \rangle \\ &= \sum_{l=1, l\neq j}^N e^{\theta_l \beta_{jl}} a_{jl} \langle \zeta_{t-}, e_l \rangle. \end{aligned} \tag{4.13}$$

Now, we give the dynamics of the deseasonalized temperature under $\mathbb{Q}^\theta$.

Lemma 4.2. *Let $\bar{R} = (R_1, \ldots, R_N)' \in \mathbb{R}^N$, where*

$$R_l = \theta_l \sigma_l^2 + \sum_{j=1, j\neq l}^{N} e^{\theta_l \beta_{jl}} \beta_{jl} a_{jl}, \tag{4.14}$$

for $l = 1, \ldots, N$. We define

$$R_t := \langle \bar{R}, \zeta_t \rangle = \sum_{l=1}^{N} R_l \langle \zeta_t, e_l \rangle.$$

Then, under $\mathbb{Q}^\theta$ the deseasonalized temperature Y_t, given by Equation (2.13)*, has the dynamics*

$$dY_t = \left(\kappa \left(Y_t - S_t\right) + R_t\right) dt + \sigma_t dW_t^\theta + \sum_{j=1}^{N} \beta_t^j d\mathcal{M}_t^{\theta j}. \tag{4.15}$$

Moreover, the solution of Equation (4.15) *with the initial value Y_s is given by*

$$\begin{aligned} Y_t = e^{\kappa(t-s)} Y_s - \kappa \int_s^t e^{\kappa(t-u)} S_u du + \int_s^t e^{\kappa(t-u)} R_u du + \int_s^t e^{\kappa(t-u)} \sigma_u dW_u^\theta \\ + \int_s^t e^{\kappa(t-u)} \sum_{j=1}^{N} \beta_u^j d\mathcal{M}_u^{\theta j}. \end{aligned} \tag{4.16}$$

Proof. Remember that, by Equation (2.13), under $\mathbb{P}$,

$$dY_t = \kappa \left(Y_t - S_t\right) dt + \sigma_t dW_t + \sum_{j=1}^{N} \beta_t^j d\mathcal{N}_t^j.$$

By Lemma 4.1,

$$dW_t^\theta = dW_t - \theta_t \sigma_t dt \tag{4.17}$$

and

$$d\mathcal{M}_t^{\theta j} = d\mathcal{N}_t^j - e^{\theta_t \beta_t^j} a_t^j dt. \tag{4.18}$$

Hence, under $\mathbb{Q}^\theta$ we can write

$$\begin{aligned} dY_t &= \kappa \left(Y_t - S_t\right) dt + \left(\theta_t \sigma_t^2 + \sum_{j=1}^{N} e^{\theta_t \beta_t^j} \beta_t^j a_t^j\right) dt + \sigma_t dW_t^\theta + \sum_{j=1}^{N} \beta_t^j d\mathcal{M}_t^{\theta j} \\ &= \kappa \left(Y_t - S_t\right) dt + R_t dt + \sigma_t dW_t^\theta + \sum_{j=1}^{N} \beta_t^j d\mathcal{M}_t^{\theta j}, \end{aligned}$$

and by applying Itô's Formula, the result follows. □

Lemma 4.3. *Under* $\mathbb{Q}^\theta$,

$$\begin{aligned}\int_{\tau_1}^{\tau_2} T_t dt = {} & \kappa^{-1}\left(e^{\kappa(\tau_2-s)} - e^{\kappa(\tau_1-s)}\right) Y_s + \int_{\tau_1}^{\tau_2} \Lambda_u du \\ & - \int_s^{\tau_2} \left(e^{\kappa(\tau_2-u)} - 1\right) S_u du + \int_s^{\tau_1} \left(e^{\kappa(\tau_1-u)} - 1\right) S_u du \\ & + \kappa^{-1} \int_s^{\tau_2} \left(e^{\kappa(\tau_2-u)} - 1\right) R_u du - \kappa^{-1} \int_s^{\tau_1} \left(e^{\kappa(\tau_1-u)} - 1\right) R_u du \\ & + \kappa^{-1} \int_s^{\tau_2} \left(e^{\kappa(\tau_2-u)} - 1\right) \sigma_u dW_u^\theta - \kappa^{-1} \int_s^{\tau_1} \left(e^{\kappa(\tau_1-u)} - 1\right) \sigma_u dW_u^\theta \\ & + \kappa^{-1} \int_s^{\tau_2} \left(e^{\kappa(\tau_2-u)} - 1\right) \sum_{j=1}^N \beta_u^j d\mathcal{M}_u^{\theta j} \\ & - \kappa^{-1} \int_s^{\tau_1} \left(e^{\kappa(\tau_1-u)} - 1\right) \sum_{j=1}^N \beta_u^j d\mathcal{M}_u^{\theta j}.\end{aligned} \tag{4.19}$$

Proof. By Equation (2.10), we have

$$\int_{\tau_1}^{\tau_2} T_t dt = \int_{\tau_1}^{\tau_2} \Lambda_u du + \int_{\tau_1}^{\tau_2} Y_u du. \tag{4.20}$$

From Equation (4.15), we can write

$$\begin{aligned}Y_{\tau_2} = {} & Y_{\tau_1} + \kappa \int_{\tau_1}^{\tau_2} Y_u du - \kappa \int_{\tau_1}^{\tau_2} S_u du + \int_{\tau_1}^{\tau_2} R_u du \\ & + \int_{\tau_1}^{\tau_2} \sigma_u dW_u^\theta + \int_{\tau_1}^{\tau_2} \sum_{j=1}^N \beta_u^j d\mathcal{M}_u^{\theta j}\end{aligned}$$

and, thus,

$$\begin{aligned}\int_{\tau_1}^{\tau_2} Y_u du = {} & \kappa^{-1}\left(Y_{\tau_2} - Y_{\tau_1}\right) + \int_{\tau_1}^{\tau_2} S_u du - \kappa^{-1} \int_{\tau_1}^{\tau_2} R_u du \\ & - \kappa^{-1} \int_{\tau_1}^{\tau_2} \sigma_u dW_u^\theta - \kappa^{-1} \int_{\tau_1}^{\tau_2} \sum_{j=1}^N \beta_u^j d\mathcal{M}_u^{\theta j}.\end{aligned} \tag{4.21}$$

Now, by Equation (4.16), we have

$$\begin{aligned}\kappa^{-1}\left(Y_{\tau_2}-Y_{\tau_1}\right)=&\ \kappa^{-1}\left(e^{\kappa(\tau_2-s)}-e^{\kappa(\tau_1-s)}\right)Y_s\\&-\int_s^{\tau_2}e^{\kappa(\tau_2-u)}S_u du+\int_s^{\tau_1}e^{\kappa(\tau_1-u)}S_u du\\&+\kappa^{-1}\int_s^{\tau_2}e^{\kappa(\tau_2-u)}R_u du-\kappa^{-1}\int_s^{\tau_1}e^{\kappa(\tau_1-u)}R_u du\\&+\kappa^{-1}\int_s^{\tau_2}e^{\kappa(\tau_2-u)}\sigma_u dW_u^\theta-\kappa^{-1}\int_s^{\tau_1}e^{\kappa(\tau_1-u)}\sigma_u dW_u^\theta\\&+\kappa^{-1}\int_s^{\tau_2}e^{\kappa(\tau_2-u)}\sum_{j=1}^N\beta_u^j d\mathcal{M}_u^{\theta j}-\kappa^{-1}\int_s^{\tau_1}e^{\kappa(\tau_1-u)}\sum_{j=1}^N\beta_u^j d\mathcal{M}_u^{\theta j}.\end{aligned}\tag{4.22}$$

Hence, by Equation (4.22), we can write Equation (4.21) as

$$\begin{aligned}\int_{\tau_1}^{\tau_2}Y_u du=&\ \kappa^{-1}\left(e^{\kappa(\tau_2-s)}-e^{\kappa(\tau_1-s)}\right)Y_s\\&-\int_s^{\tau_2}\left(e^{\kappa(\tau_2-u)}-1\right)S_u du+\int_s^{\tau_1}\left(e^{\kappa(\tau_1-u)}-1\right)S_u du\\&+\kappa^{-1}\int_s^{\tau_2}\left(e^{\kappa(\tau_2-u)}-1\right)R_u du-\kappa^{-1}\int_s^{\tau_1}\left(e^{\kappa(\tau_1-u)}-1\right)R_u du\\&+\kappa^{-1}\int_s^{\tau_2}\left(e^{\kappa(\tau_2-u)}-1\right)\sigma_u dW_u^\theta-\kappa^{-1}\int_s^{\tau_1}\left(e^{\kappa(\tau_1-u)}-1\right)\sigma_u dW_u^\theta\\&+\kappa^{-1}\int_s^{\tau_2}\left(e^{\kappa(\tau_2-u)}-1\right)\sum_{j=1}^N\beta_u^j d\mathcal{M}_u^{\theta j}\\&-\kappa^{-1}\int_s^{\tau_1}\left(e^{\kappa(\tau_1-u)}-1\right)\sum_{j=1}^N\beta_u^j d\mathcal{M}_u^{\theta j}.\end{aligned}\tag{4.23}$$

Therefore, by putting Equation (4.23) into Equation (4.20), the result follows. □

We denote the expectation under $\mathbb{Q}^\theta$ by $\mathbb{E}^\theta[\cdot]$. We now give a useful result.

Lemma 4.4. *Consider the chain ζ of which the semimartingale decomposition under $\mathbb{Q}^\theta$ is given by Equation* (4.9)*. Then we have*

$$\mathbb{E}^\theta\left[\zeta_t|\mathcal{G}_s\right]=e^{\mathbf{A}^\theta(t-s)}\zeta_s.\tag{4.24}$$

Proof. By Itô's Formula we have

$$\begin{aligned}d\left(e^{-\mathbf{A}^\theta t}\zeta_t\right)&=-\mathbf{A}^\theta e^{-\mathbf{A}^\theta t}\zeta_t dt+e^{-\mathbf{A}^\theta t}d\zeta_t\\&=e^{-\mathbf{A}^\theta t}dV_t^\theta,\end{aligned}$$

since by Equation (4.9)

$$d\zeta_t = \mathbf{A}^\theta \zeta_{t-} dt + dV_t^\theta.$$

Hence, we can write

$$e^{-\mathbf{A}^\theta t}\zeta_t = e^{-\mathbf{A}^\theta s}\zeta_s + \int_s^t e^{-\mathbf{A}^\theta u} dV_u^\theta.$$

Thus,

$$\mathbb{E}^\theta \left[e^{-\mathbf{A}^\theta t}\zeta_t | \mathcal{G}_s \right] = e^{-\mathbf{A}^\theta s}\zeta_s,$$

since by the martingale property

$$\mathbb{E}^\theta \left[\int_s^t e^{-\mathbf{A}^\theta u} dV_u^\theta | \mathcal{G}_s \right] = 0.$$

Therefore, the result follows. □

Remark 4.1. Notice that

1. By Lemma 4.4, we have

$$\begin{aligned} \mathbb{Q}^\theta [\zeta_u = e_j | \zeta_s] &= \mathbb{E}^\theta \left[\mathbb{I}_{\{\zeta_u = e_j\}}(u) | \zeta_s \right] \\ &= \mathbb{E}^\theta [\langle \zeta_u, e_j \rangle | \zeta_s] \\ &= \langle \mathbb{E}^\theta [\zeta_u | \zeta_s], e_j \rangle \\ &= \langle e^{\mathbf{A}^\theta (u-s)} \zeta_s, e_j \rangle. \end{aligned}$$

2. For any square matrix $\mathbf{Q}$,

$$e^{\mathbf{Q}} := \sum_{k=0}^\infty \frac{\mathbf{Q}^k}{k!}.$$

In the following, we give another representation of the chain under $\mathbb{Q}^\theta$, which will be useful later on.

Remark 4.2. Remember that $\mathcal{N}_t = (\mathcal{N}_t^1, \dots, \mathcal{N}_t^N)' \in \mathbb{R}^N$, given by Equation (2.16), is a vector of counting proces, see Remark 2.2. Notice that we have $\mathbb{E}^\theta(\mathcal{N}_t) = \int_0^t \mathbf{A}_0^\theta \zeta_{s-} ds$. Thus, the process $\mathcal{M}^\theta := (\mathcal{M}_t^\theta : \ t \in [0, \mathsf{T}])$ given by

$$\mathcal{M}_t^\theta = \mathcal{N}_t - \int_0^t \mathbf{A}_0^\theta \zeta_{s-} ds \tag{4.25}$$

is a martingale with $\mathcal{M}_t^\theta = (\mathcal{M}_t^{\theta 1}, \dots, \mathcal{M}_t^{\theta N})' \in \mathbb{R}^N$ (see Equation (4.6)).

By Remarks 2.1 and 4.2, under $\mathbb{Q}^\theta$ the chain, given by Equation (4.9), can also be represented as

$$\begin{aligned} \zeta_t &= \zeta_0 + \int_0^t (\mathbf{I} - \zeta_{s-}\mathbf{1}') d\mathcal{N}_s, \\ &= \zeta_0 + \int_0^t \mathbf{A}^\theta \zeta_{s-} ds + \int_0^t (\mathbf{I} - \zeta_{s-}\mathbf{1}') d\mathcal{M}_t^\theta. \end{aligned} \tag{4.26}$$

4.2 Pricing Temperature Futures

In this section, we derive prices of temperature futures written on CAT, CDD and HDD indices using the temperature model proposed in Section 2.2. Remember that the cumulative average temperature (CAT), the cooling degree days (CDD) and the heating degree days (HDD) indices over a measurement period $[\tau_1, \tau_2]$ are defined as

$$CAT(\tau_1, \tau_2) = \int_{\tau_1}^{\tau_2} T_t dt. \tag{4.27}$$

$$CDD(\tau_1, \tau_2) = \int_{\tau_1}^{\tau_2} \max(T_t - c, 0) dt \tag{4.28}$$

and

$$HDD(\tau_1, \tau_2) = \int_{\tau_1}^{\tau_2} \max(c - T_t, 0) dt, \tag{4.29}$$

respectively, where c is a constant and denotes the threshold temperature, typically, 18 degrees Celsius or 65 degrees Fahrenheit.

Consider a futures contract written on a CAT index over the measurement period $[\tau_1, \tau_2]$. By definition of a futures contract, it is costless to enter. At the end of the measurement period, the buyer of the contract receives the amount in Equation (4.27) and pays the CAT futures price $F_{\text{CAT}}(s, \tau_1, \tau_2; T)$ if the contract was entered time $s \leq \tau_1$. Therefore, from arbitrage theory, we must have

$$0 = e^{-r(\tau_2 - s)} \mathbb{E}^\theta \left[\int_{\tau_1}^{\tau_2} T_t dt - F_{\text{CAT}}(s, \tau_1, \tau_2; T) | \mathcal{G}_s \right],$$

where $r > 0$ is a constant risk-free rate of return. Assuming that the futures prices is adapted, the CAT futures price is defined by

$$F_{\text{CAT}}(s, \tau_1, \tau_2; T) = \mathbb{E}^\theta \left[\int_{\tau_1}^{\tau_2} T_t dt | \mathcal{G}_s \right]. \tag{4.30}$$

Similarly, the CDD and HDD futures prices are defined by

$$F_{\text{CDD}}(s, \tau_1, \tau_2; T) = \mathbb{E}^\theta \left[\int_{\tau_1}^{\tau_2} \max(T_t - c, 0) dt | \mathcal{G}_s \right], \tag{4.31}$$

and

$$F_{\text{HDD}}(s, \tau_1, \tau_2; T) = \mathbb{E}^\theta \left[\int_{\tau_1}^{\tau_2} \max(c - T_t, 0) dt | \mathcal{G}_s \right], \tag{4.32}$$

respectively.

Moreover, since

$$\max(c - x, 0) = \max(x - c, 0) + c - x,$$

the following CDD-HDD parity follows:

$$F_{\text{HDD}}(s, \tau_1, \tau_2; T) = F_{\text{CDD}}(s, \tau_1, \tau_2; T) + c(\tau_2 - \tau_1) - F_{\text{CAT}}(s, \tau_1, \tau_2; T). \tag{4.33}$$

4.2.1 The CAT Futures

The CAT futures price is given by the following Theorem.

Theorem 4.5. *The futures price $F_{CAT}(s, \tau_1, \tau_2; T)$ at time $s \leq \tau_1$ written on a CAT index over the interval $[\tau_1, \tau_2]$ is*

$$\begin{aligned} F_{CAT}(s, \tau_1, \tau_2; T) = &\int_{\tau_1}^{\tau_2} \Lambda_u du + \kappa^{-1}\left(e^{\kappa(\tau_2 - s)} - e^{\kappa(\tau_1 - s)}\right) Y_s \\ &+ \mathcal{S}(s, \tau_1, \tau_2) + \Theta(s, \tau_1, \tau_2), \end{aligned} \tag{4.34}$$

where

$$\begin{aligned} \mathcal{S}(s, \tau_1, \tau_2) = &-\int_s^{\tau_2} \left(e^{\kappa(\tau_2 - u)} - 1\right) \langle \bar{S}, e^{\mathbf{A}^\theta(u-s)}\zeta_s \rangle du \\ &+ \int_s^{\tau_1} \left(e^{\kappa(\tau_1 - u)} - 1\right) \langle \bar{S}, e^{\mathbf{A}^\theta(u-s)}\zeta_s \rangle du \end{aligned} \tag{4.35}$$

and

$$\begin{aligned} \Theta(s, \tau_1, \tau_2) = &\kappa^{-1}\int_s^{\tau_2} \left(e^{\kappa(\tau_2 - u)} - 1\right) \langle \bar{R}, e^{\mathbf{A}^\theta(u-s)}\zeta_s \rangle du \\ &- \kappa^{-1}\int_s^{\tau_1} \left(e^{\kappa(\tau_1 - u)} - 1\right) \langle \bar{R}, e^{\mathbf{A}^\theta(u-s)}\zeta_s \rangle du. \end{aligned} \tag{4.36}$$

Proof. Remember that by Equation (4.30)

$$F_{\text{CAT}}(s, \tau_1, \tau_2; T) = \mathbb{E}^\theta\left[\int_{\tau_1}^{\tau_2} T_t dt | \mathcal{G}_s\right]. \tag{4.37}$$

By Equation (4.19) and Fubini's Theorem, we have

$$\begin{aligned} \mathbb{E}^\theta\left[\int_{\tau_1}^{\tau_2} T_t dt | \mathcal{G}_s\right] = &\int_{\tau_1}^{\tau_2} \Lambda_u du + \kappa^{-1}\left(e^{\kappa(\tau_2 - s)} - e^{\kappa(\tau_1 - s)}\right) Y_s \\ &- \int_s^{\tau_2} \left(e^{\kappa(\tau_2 - u)} - 1\right) \mathbb{E}^\theta\left[S_u | \mathcal{G}_s\right] du \\ &+ \int_s^{\tau_1} \left(e^{\kappa(\tau_1 - u)} - 1\right) \mathbb{E}^\theta\left[S_u | \mathcal{G}_s\right] du \\ &+ \kappa^{-1}\int_s^{\tau_2} \left(e^{\kappa(\tau_2 - u)} - 1\right) \mathbb{E}^\theta\left[R_u | \mathcal{G}_s\right] du \\ &- \kappa^{-1}\int_s^{\tau_1} \left(e^{\kappa(\tau_1 - u)} - 1\right) \mathbb{E}^\theta\left[R_u | \mathcal{G}_s\right] du, \end{aligned} \tag{4.38}$$

since by the martingale property

$$\mathbb{E}^\theta\left[\int_s^{\tau_2} \left(e^{\kappa(\tau_2 - u)} - 1\right) \sigma_u dW_u^\theta | \mathcal{G}_s\right] = 0,$$

$$\mathbb{E}^\theta\left[\int_s^{\tau_1}\left(e^{\kappa(\tau_1-u)}-1\right)\sigma_u dW_u^\theta|\mathcal{G}_s\right]=0,$$

$$\mathbb{E}^\theta\left[\int_s^{\tau_2}\left(e^{\kappa(\tau_2-u)}-1\right)\sum_{j=1}^N \beta_u^j d\mathcal{M}_u^{\theta j}|\mathcal{G}_s\right]=0,$$

$$\mathbb{E}^\theta\left[\int_s^{\tau_1}\left(e^{\kappa(\tau_1-u)}-1\right)\sum_{j=1}^N \beta_u^j d\mathcal{M}_u^{\theta j}|\mathcal{G}_s\right]=0.$$

Remember that by Lemma 4.4

$$\mathbb{E}^\theta\left[\zeta_u|\mathcal{G}_s\right]=e^{\mathbf{A}^\theta(u-s)}\zeta_s.$$

Thus, we have

$$\begin{aligned}\mathbb{E}^\theta\left[S_u|\mathcal{G}_s\right]&=\mathbb{E}^\theta\left[\langle\bar{S},\zeta_u\rangle|\mathcal{G}_s\right]\\&=\langle\bar{S},\mathbb{E}^\theta\left[\zeta_u|\mathcal{G}_s\right]\rangle\\&=\langle\bar{S},e^{\mathbf{A}^\theta(u-s)}\zeta_s\rangle\end{aligned}\tag{4.39}$$

and

$$\begin{aligned}\mathbb{E}^\theta\left[R_u|\mathcal{G}_s\right]&=\mathbb{E}^\theta\left[\langle\bar{R},\zeta_u\rangle|\mathcal{G}_s\right]\\&=\langle\bar{R},\mathbb{E}^\theta\left[\zeta_u|\mathcal{G}_s\right]\rangle\\&=\langle\bar{R},e^{\mathbf{A}^\theta(u-s)}\zeta_s\rangle.\end{aligned}\tag{4.40}$$

Therefore, by inserting Equations (4.39) and (4.40) into Equation (4.38), the result follows. □

Proposition 4.6. *The futures price $F_{CAT}(s,\tau_1,\tau_2;T)$ at time s, with $\tau_1\le s\le\tau_2$, written on a CAT index over the interval $[\tau_1,\tau_2]$ is*

$$F_{CAT}(s,\tau_1,\tau_2;T)=\int_{\tau_1}^s T_u du+F_{CAT}(s,s,\tau_2;T).$$

Proof. By Equation (4.30),

$$\begin{aligned}F_{\text{CAT}}(s,\tau_1,\tau_2;T)&=\mathbb{E}^\theta\left[\int_{\tau_1}^{\tau_2}T_u du|\mathcal{G}_s\right]\\&=\int_{\tau_1}^s T_u du+\mathbb{E}^\theta\left[\int_s^{\tau_2}T_u du|\mathcal{G}_s\right].\end{aligned}$$

Therefore, the result follows. □

4.2.2 The CDD and HDD Futures

The following lemma is necessary to price the CDD futures.

Lemma 4.7. *Let $\Phi_{Y_t|\mathcal{G}_s}(u)$ be the characteristic function of Y_t conditional on $\mathcal{G}_s$, where $s \le t$. Then, with $i = \sqrt{-1}$,*

$$\begin{aligned}\Phi_{Y_t|\mathcal{G}_s}(u) &:= \mathbb{E}^\theta\left[e^{iuY_t}|\mathcal{G}_s\right] \\ &= \langle \zeta_s \exp\left(iue^{\kappa(t-s)}Y_s + \int_s^t \left(diag[g(r, ue^{kt})] + \mathbf{B}^\theta(r)\right) dr \right), \mathbf{1}\rangle,\end{aligned} \tag{4.41}$$

where

$$\begin{aligned} g(t,u) &:= (g_1(t,u), \ldots, g_N(t,u))', \\ g_l(t,u) &:= iue^{-\kappa t}(-\kappa S_l + R_l) - \frac{1}{2}u^2 e^{-2\kappa t}\sigma_l^2 \\ &\quad + \sum_{j=1, j\neq l}^N e^{\theta_l \beta_{jl}}(e^{iue^{-\kappa t}\beta_{jl}} - 1 - iue^{-\kappa t}\beta_{jl})a_{jl}\end{aligned}$$

for each $l = 1, \ldots, N$, and $\mathbf{B}^\theta(t) := (b^\theta_{jl}(t))_{j,l=1,\ldots,N}$ with

$$b^\theta_{jl}(t) = \begin{cases} e^{iue^{-\kappa t}\beta_{jl}}a^\theta_{jl}, & \textit{for } l \neq j, \\ -\sum_{j=1,j\neq l}^N e^{iue^{-\kappa t}\beta_{jl}}a^\theta_{jl}, & \textit{for } l = j.\end{cases}$$

Proof. Let be

$$Z_t := e^{-\kappa t}Y_t$$

and apply Itô's Formula to e^{iuZ_t}. Then,

$$\begin{aligned} e^{iuZ_t} &= e^{iuZ_s} + \int_s^t e^{iuZ_r}\left(iue^{-\kappa r}(-\kappa S_r + R_r) - \frac{1}{2}u^2 e^{-2\kappa r}\sigma_r^2\right) dr \\ &+ \int_s^t e^{iuZ_r}\sum_{j=1}^N e^{\theta_r \beta_r^j}(e^{iue^{-\kappa r}\beta_r^j} - 1 - iue^{-\kappa r}\beta_r^j)a_r^j dr \\ &+ \int_s^t e^{iuZ_r} iue^{-\kappa r}\sigma_r dW_r^\theta + \int_s^t e^{iuZ_{r-}}\sum_{j=1}^N (e^{iue^{-\kappa r}\beta_r^j} - 1)d\mathcal{M}_r^{\theta j} \\ &= e^{iuZ_s} + \int_s^t e^{iuZ_r}\langle g(r,u), \zeta_r\rangle dr \\ &+ \int_s^t e^{iuZ_r} iue^{-\kappa r}\sigma_r dW_r^\theta + \int_s^t e^{iuZ_{r-}}\sum_{j=1}^N (e^{iue^{-\kappa r}\beta_r^j} - 1)d\mathcal{M}_r^{\theta j}.\end{aligned}$$

We define for each $s \le t \in [0, \mathsf{T}]$, $\bar{\mathcal{G}}_{s,t} := \mathcal{F}_s \vee \mathcal{F}_t^\zeta$, which represents the enlarged σ-field generated by $\mathcal{F}_s$ and $\mathcal{F}_t^\zeta$. Moreover, write $\bar{\mathbf{G}} := (\bar{\mathcal{G}}_{s,t} :\ s,t \in [0,\mathsf{T}])$ for the corresponding complete enlarged filtration. Let $\Phi_{Z_t|\bar{\mathcal{G}}_{s,t}}(u)$ denote the characteristic function of Z_t conditional on $\bar{\mathcal{G}}_{s,t}$, that is,

$$\Phi_{Z_t|\bar{\mathcal{G}}_{s,t}}(u) := \mathbb{E}^\theta \left[e^{iuZ_t} | \bar{\mathcal{G}}_{s,t} \right]. \tag{4.42}$$

Then, from above we obtain

$$d\Phi_{Z_t|\bar{\mathcal{G}}_{s,t}}(u) = \Phi_{Z_t|\bar{\mathcal{G}}_{s,t}}(u) \left(\langle g(t,u), \zeta_t \rangle dt + \sum_{j=1}^{N} (e^{iue^{-\kappa t}\beta_t^j} - 1) d\mathcal{M}_t^{\theta j} \right). \tag{4.43}$$

Let $\mathbf{D}^\theta(t) := (d_{jl}^\theta(t))_{j,l=1,\dots,N}$, where

$$d_{jl}^\theta(t) = \begin{cases} e^{iue^{-\kappa t}\beta_{jl}}, & \text{for } l \neq j, \\ \dfrac{\sum_{j=1, j\neq l}^{N} e^{iue^{-\kappa t}\beta_{jl}} a_{jl}^\theta}{\sum_{j=1, j\neq l}^{N} a_{jl}^\theta}, & \text{for } l = j. \end{cases}$$

Notice that $d_{jl}^\theta(t) = b_{jl}^\theta(t)/a_{jl}^\theta$, for each $j, l = 1, \dots, N$. Define $\mathbf{D}_0^\theta(t) := \mathbf{D}^\theta(t) - diag[d^\theta(t)]$, where $d^\theta(t) = (d_{11}^\theta(t), \dots, d_{NN}^\theta(t))' \in \mathbb{R}^N$. Then, we can write

$$\sum_{j=1}^{N} (e^{iue^{-\kappa t}\beta_t^j} - 1) d\mathcal{M}_t^{\theta j} = \left(\mathbf{D}_0^\theta(t)\zeta_{t-} + \zeta_{t-} - \mathbf{1} \right)' d\mathcal{M}_t^\theta, \tag{4.44}$$

where $\mathcal{M}_t^\theta = (\mathcal{M}_t^{\theta 1}, \dots, \mathcal{M}_t^{\theta N})' \in \mathbb{R}^N$. Remember that by Remark 4.2, we have

$$\zeta_t = \zeta_0 + \int_0^t (\mathbf{I} - \zeta_{s-}\mathbf{1}') d\mathcal{N}_s$$

with

$$\mathcal{M}_t^\theta = \mathcal{N}_t - \int_0^t \mathbf{A}_0^\theta \zeta_{s-} ds.$$

Therefore, Equation (4.43) can be written as

$$d\Phi_{Z_t|\bar{\mathcal{G}}_{s,t}}(u) = \Phi_{Z_t|\bar{\mathcal{G}}_{s,t}}(u) \left(\langle g(t,u), \zeta_t \rangle dt + \left(\mathbf{D}_0^\theta(t)\zeta_{t-} + \zeta_{t-} - \mathbf{1} \right)' d\mathcal{M}_t^\theta \right). \tag{4.45}$$

We define

$$h(t,u) := \zeta_t \Phi_{Z_t|\bar{\mathcal{G}}_{s,t}}(u), \qquad t \in [0, \mathsf{T}]. \tag{4.46}$$

By Itô's Formula we obtain

$$
\begin{aligned}
h(t,u) &= h(s,u) + \int_s^t \Phi_{Z_r|\bar{\mathcal{G}}_{s,r}}(u)\left(\mathbf{A}^\theta \zeta_{r-}dr + dV_r^\theta\right) \\
&+ \int_s^t \zeta_{r-} d\Phi_{Z_r|\bar{\mathcal{G}}_{s,r}}(u) + \sum_{s<r\le t} \Delta\zeta_r \Delta\Phi_{Z_r|\bar{\mathcal{G}}_{s,r}}(u) \\
&= h(s,u) + \int_s^t \left(diag[g(r,u)] + \mathbf{A}^\theta\right) h(r,u)dr \\
&+ \int_s^t \Phi_{Z_r|\bar{\mathcal{G}}_{s,r}}(u) dV_r^\theta \\
&+ \int_s^t h(r-,u)\left(\mathbf{D}_0^\theta(r)\zeta_{r-} + \zeta_{r-} - \mathbf{1}\right)' d\mathcal{M}_r^\theta \\
&+ \sum_{s<r\le t} \Delta\zeta_r \Delta\Phi_{Z_r|\bar{\mathcal{G}}_{s,r}}(u).
\end{aligned}
\tag{4.47}
$$

Here, we used $\Phi_{Z_t|\bar{\mathcal{G}}_{s,t}}(u)\zeta_t\langle g(t,u),\zeta_t\rangle = diag[g(t,u)]\zeta_t\Phi_{Z_t|\bar{\mathcal{G}}_{s,t}}(u)$.

Now, by using

$$(\mathbf{I} - \zeta_t\mathbf{1}')diag[\Delta\mathcal{N}_t]\zeta_t = \mathbf{0},$$

we can write

$$
\begin{aligned}
&\sum_{s<r\le t} \Delta\zeta_r \Delta\Phi_{Z_r|\bar{\mathcal{G}}_{s,r}}(u) \\
&= \sum_{s<r\le t} (\mathbf{I} - \zeta_{r-}\mathbf{1}')\Delta\mathcal{N}_r \Phi_{Z_r|\bar{\mathcal{G}}_{s,r}}(u)\left(\mathbf{D}_0^\theta(r)\zeta_{r-} + \zeta_{r-} - \mathbf{1}\right)' \Delta\mathcal{N}_r \\
&= \sum_{s<r\le t} \Phi_{Z_r|\bar{\mathcal{G}}_{s,r}}(u)(\mathbf{I} - \zeta_{r-}\mathbf{1}')diag[\Delta\mathcal{N}_r]\left(\mathbf{D}_0^\theta(r)\zeta_{r-} + \zeta_{r-} - \mathbf{1}\right) \\
&= \sum_{s<r\le t} \Phi_{Z_r|\bar{\mathcal{G}}_{s,r}}(u)(\mathbf{I} - \zeta_{r-}\mathbf{1}')diag[\Delta\mathcal{N}_r]\left(\mathbf{D}_0^\theta(r)\zeta_{r-} - \mathbf{1}\right) \\
&= \int_s^t \Phi_{Z_r|\bar{\mathcal{G}}_{s,r}}(u)(\mathbf{I} - \zeta_{r-}\mathbf{1}')diag[\mathbf{A}_0^\theta\zeta_{r-}]\left(\mathbf{D}_0^\theta(r)\zeta_{r-} - \mathbf{1}\right) dr \\
&+ \int_s^t \Phi_{Z_r|\bar{\mathcal{G}}_{s,r}}(u)(\mathbf{I} - \zeta_{r-}\mathbf{1}')diag[d\mathcal{M}_r^\theta]\left(\mathbf{D}_0^\theta(r)\zeta_{r-} - \mathbf{1}\right).
\end{aligned}
\tag{4.48}
$$

It can be easily shown that

$$diag[\mathbf{A}_0^\theta\zeta_t]\left(\mathbf{D}_0^\theta(t)\zeta_t - \mathbf{1}\right) = \left(\mathbf{B}_0^\theta(t) - \mathbf{A}_0^\theta\right)\zeta_t.$$

Hence, we can write

$$\begin{aligned}
&\int_s^t \Phi_{Z_r|\bar{\mathcal{G}}_{s,r}}(u)(\mathbf{I}-\zeta_{r-}\mathbf{1}')diag[\mathbf{A}_0^\theta\zeta_{r-}]\left(\mathbf{D}_0^\theta(r)\zeta_{r-}-\mathbf{1}\right)dr \\
&=\int_s^t \Phi_{Z_r|\bar{\mathcal{G}}_{s,r}}(u)(\mathbf{I}-\zeta_{r-}\mathbf{1}')\left(\mathbf{B}_0^\theta(r)-\mathbf{A}_0^\theta\right)\zeta_{r-}dr \\
&=\int_s^t \left(\mathbf{B}^\theta(r)-\mathbf{A}^\theta\right)h(r,u)dr,
\end{aligned} \tag{4.49}$$

since

$$(\mathbf{I}-\zeta_t\mathbf{1}')\mathbf{B}_0^\theta(t)\zeta_t=\mathbf{B}^\theta(t)\zeta_t$$

and

$$(\mathbf{I}-\zeta_t\mathbf{1}')\mathbf{A}_0^\theta\zeta_t=\mathbf{A}^\theta\zeta_t.$$

Thus by combining Equations (4.48) and (4.49), we get

$$\begin{aligned}
&\sum_{s<r\leq t}\Delta\zeta_r\Delta\Phi_{Z_r|\bar{\mathcal{G}}_{s,r}}(u) \\
&=\int_s^t \left(\mathbf{B}^\theta(r)-\mathbf{A}^\theta\right)h(r,u)dr \\
&+\int_s^t \Phi_{Z_r|\bar{\mathcal{G}}_{s,r}}(u)(\mathbf{I}-\zeta_{r-}\mathbf{1}')diag[d\mathcal{M}_r^\theta]\left(\mathbf{D}_0^\theta(r)\zeta_{r-}-\mathbf{1}\right).
\end{aligned} \tag{4.50}$$

Thus, by Equation (4.50), Equation (4.47) becomes

$$\begin{aligned}
h(t,u)=h(s,u)&+\int_s^t \left(diag[g(r,u)]+\mathbf{B}^\theta(r)\right)h(r,u)dr \\
&+\int_s^t \Phi_{Z_r|\bar{\mathcal{G}}_{s,r}}(u)dV_r^\theta \\
&+\int_s^t h(r-,u)\left(\mathbf{D}_0^\theta(r)\zeta_{r-}+\zeta_{r-}-\mathbf{1}\right)'d\mathcal{M}_r^\theta \\
&+\int_s^t \Phi_{Z_r|\bar{\mathcal{G}}_{s,r}}(u)(\mathbf{I}-\zeta_{r-}\mathbf{1}')diag[d\mathcal{M}_r^\theta]\left(\mathbf{D}_0^\theta(r)\zeta_{r-}-\mathbf{1}\right).
\end{aligned} \tag{4.51}$$

Therefore, by Fubini's Theorem and using the martingale property, we have

$$\mathbb{E}^\theta\left[h(t,u)|\mathcal{G}_s\right]=h(s,u)+\int_s^t \left(diag[g(r,u)]+\mathbf{B}^\theta(r)\right)\mathbb{E}^\theta\left[h(r,u)|\mathcal{G}_s\right]dr.$$

Thus, we get

$$d\mathbb{E}^\theta\left[h(t,u)|\mathcal{G}_s\right]=\left(diag[g(t,u)]+\mathbf{B}^\theta(t)\right)\mathbb{E}^\theta\left[h(t,u)|\mathcal{G}_s\right]dt$$

and after solving we have

$$\mathbb{E}^{\theta}\left[h(t,u)|\mathcal{G}_s\right] = e^{iuZ_s}\zeta_s \exp\left(\int_s^t \left(diag[g(r,u)] + \mathbf{B}^{\theta}(r)\right) dr\right). \tag{4.52}$$

Notice that

$$\begin{aligned}\Phi_{Z_t|\bar{\mathcal{G}}_{s,t}}(u) &= \langle \zeta_t \Phi_{Z_t|\bar{\mathcal{G}}_{s,t}}(u), \mathbf{1}\rangle \\ &= \langle h(t,u), \mathbf{1}\rangle,\end{aligned} \tag{4.53}$$

since $\langle \zeta_t, \mathbf{1}\rangle = 1$.

Now, by the tower property and Equation (4.53), we have

$$\begin{aligned}\mathbb{E}^{\theta}\left[e^{iuY_t}|\mathcal{G}_s\right] &= \mathbb{E}^{\theta}\left[\mathbb{E}^{\theta}\left[e^{iuY_t}|\bar{\mathcal{G}}_{s,t}\right]|\mathcal{G}_s\right] \\ &= \mathbb{E}^{\theta}\left[\mathbb{E}^{\theta}\left[e^{iue^{\kappa t}Z_t}|\bar{\mathcal{G}}_{s,t}\right]|\mathcal{G}_s\right] \\ &= \mathbb{E}^{\theta}\left[\Phi_{Z_t|\bar{\mathcal{G}}_{s,t}}(ue^{\kappa t})|\mathcal{G}_s\right] \\ &= \mathbb{E}^{\theta}\left[\langle h(t,ue^{\kappa t}), \mathbf{1}\rangle|\mathcal{G}_s\right] \\ &= \langle \mathbb{E}^{\theta}\left[h(t,ue^{\kappa t})|\mathcal{G}_s\right], \mathbf{1}\rangle.\end{aligned}$$

Therefore the result follows by Equation (4.52). □

To price the CDD futures, we apply Fourier transform techniques. We denote the space of integrable functions on $\mathbb{R}$ by $L^1(\mathbb{R})$. For $f \in L^1(\mathbb{R})$, the Fourier transform of f is defined by

$$\hat{f}(u) = \int_{\mathbb{R}} f(x)e^{-ixu}dx.$$

And if, also $\hat{f} \in L^1(\mathbb{R})$, then the inverse Fourier transform is given by [21]

$$f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}(u)e^{ixu}du.$$

We have the following lemma from [7].

Lemma 4.8. *For the function f given by*

$$f(x) = max(x-c, 0),$$

with $c > 0$, define

$$f_\epsilon(x) := e^{-\epsilon x}f(x).$$

Then, $f_\epsilon \in L^1(\mathbb{R})$ for all $\epsilon > 0$.
Moreover, for any $\epsilon > 0$, the Fourier transform of f_ϵ is given by

$$\hat{f}_\epsilon(u) = \frac{1}{(\epsilon + iu)^2}e^{-(\epsilon+iu)c}$$

and $\hat{f}_\epsilon \in L^1(\mathbb{R})$.

Now, we are ready to give the CDD futures price.

Theorem 4.9. *The futures price $F_{CDD}(s,\tau_1,\tau_2;T)$ at time $s \leq \tau_1$ written on a CDD index over the interval $[\tau_1,\tau_2]$ is*

$$F_{CDD}(s,\tau_1,\tau_2;T) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}_\epsilon(u)\int_{\tau_1}^{\tau_2} \Psi(s,t,u)dtdu, \tag{4.54}$$

where for $\epsilon > 0$,

$$\Psi(s,t,u) = \varpi(s,t,u)\langle \zeta_s \exp\left(\int_s^t \left(diag[g(r,(u-i\epsilon)e^{\kappa t})] + \mathbf{B}^\theta(r)\right) dr\right), \mathbf{1}\rangle,$$

with

$$\varpi(s,t,u) = \exp\left((\epsilon+iu)(\Lambda_t + e^{\kappa(t-s)}Y_s)\right)$$

and

$$\hat{f}_\epsilon(u) = \frac{1}{(\epsilon+iu)^2}e^{-(\epsilon+iu)c}.$$

Proof. Remember that by Equation (4.31)

$$F_{\text{CDD}}(s,\tau_1,\tau_2;T) = \mathbb{E}^\theta\left[\int_{\tau_1}^{\tau_2} \max(T_t - c, 0)dt|\mathcal{G}_s\right].$$

Now by Lemma 4.8, we can write

$$\max(T_t - c, 0) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}_\epsilon(u)\exp\left((\epsilon+iu)T_t\right)du. \tag{4.55}$$

Thus, by Equation (4.55) and Fubini's Theorem,

$$\begin{aligned}
&\mathbb{E}^\theta\left[\max(T_t - c,0)|\mathcal{G}_s\right] \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}_\epsilon(u)\mathbb{E}^\theta\left[\exp\left((\epsilon+iu)T_t\right)|\mathcal{G}_s\right]du \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}_\epsilon(u)\exp\left((\epsilon+iu)\Lambda_t\right)\mathbb{E}^\theta\left[\exp\left((\epsilon+iu)Y_t\right)|\mathcal{G}_s\right]du \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}_\epsilon(u)\exp\left((\epsilon+iu)\Lambda_t\right)\mathbb{E}^\theta\left[\exp\left(i(u-i\epsilon)Y_t\right)|\mathcal{G}_s\right]du \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}_\epsilon(u)\exp\left((\epsilon+iu)\Lambda_t\right)\Phi_{Y_t|\mathcal{G}_s}(u-i\epsilon)du,
\end{aligned} \tag{4.56}$$

where $\Phi_{Y_t|\mathcal{G}_s}(u) = \mathbb{E}^\theta\left[e^{iuY_t}|\mathcal{G}_s\right]$ is given by Equation (4.41).

Therefore, by Fubini's Theorem and Equation (4.56), we obtain

$$\begin{aligned}
&\mathbb{E}^{\theta}\left[\int_{\tau_1}^{\tau_2} \max(T_t - c, 0)dt|\mathcal{G}_s\right] \\
&= \int_{\tau_1}^{\tau_2} \mathbb{E}^{\theta}\left[\max(T_t - c, 0)|\mathcal{G}_s\right] dt \\
&= \int_{\tau_1}^{\tau_2} \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}_{\epsilon}(u) \exp\left((\epsilon + iu)\Lambda_t\right) \Phi_{Y_t|\mathcal{G}_s}(u - i\epsilon) du dt \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}_{\epsilon}(u) \int_{\tau_1}^{\tau_2} \exp\left((\epsilon + iu)\Lambda_t\right) \Phi_{Y_t|\mathcal{G}_s}(u - i\epsilon) dt du.
\end{aligned}$$

We notice that, by Lemma 4.7,

$$\begin{aligned}
&\Phi_{Y_t|\mathcal{G}_s}(u - i\epsilon) \\
&= \langle \zeta_s \exp\left((\epsilon + iu)e^{\kappa(t-s)}Y_s + \int_s^t \left(diag[g(r, (u - i\epsilon)e^{kt})] + \mathbf{B}^{\theta}(r)\right) dr\right), \mathbf{1}\rangle.
\end{aligned}$$

Therefore, the result follows. □

The CDD futures price can be computed by the fast Fourier transform method [11]. Remember that by Equation (4.32), the futures price $F_{\text{HDD}}(s, \tau_1, \tau_2; T)$ at time $s \leq \tau_1$ written on a HDD index over the interval $[\tau_1, \tau_2]$ is given by

$$F_{\text{HDD}}(s, \tau_1, \tau_2; T) = \mathbb{E}^{\theta}\left[\int_{\tau_1}^{\tau_2} \max(c - T_t, 0)dt|\mathcal{G}_s\right].$$

Since we have already found the CDD and CAT futures prices, the HDD futures price can be easily derived by the CDD-HDD parity, given by Equation (4.33). That is,

$$F_{\text{HDD}}(s, \tau_1, \tau_2; T) = F_{\text{CDD}}(s, \tau_1, \tau_2; T) + c(\tau_2 - \tau_1) - F_{\text{CAT}}(s, \tau_1, \tau_2; T),$$

where $F_{\text{CDD}}(s, \tau_1, \tau_2; T)$ and $F_{\text{CAT}}(s, \tau_1, \tau_2; T)$ are given by Equations (4.54) and (4.34), respectively.

4.3 Monte Carlo Simulation

In this section, the Monte Carlo simulation, see [23], will be considered to price temperature futures written on CDD and HDD indices using the proposed model. We begin with some information about the temperature futures trading at CME for Chicago, of which we have the temperature data.

For Chicago, there are monthly CDD index futures and HDD index futures. While the monthly CDD contract periods are May, June, July, August and September, monthly

HDD contract periods are November, December, January, February and March. In addition to the monthly contracts, CDD May-September and CDD July-August seasonal strip futures, and HDD November-March and HDD February-March seasonal strip futures are also available.

The trading schedule of the contracts are determined by the CME. The temperature futures are not settled on the index value for a particular day, but the aggregated index value over an measurement period, typically a month or several consecutive months, referred as seasonal strips. The accumulation period of each contract begins with the first calendar day of the contract period and ends with the last calendar day of the contract period. The CDD and HDD indices over a measurement period $[\tau_1, \tau_2]$ are calculated as

$$CDD(\tau_1, \tau_2) = \sum_{t=\tau_1}^{\tau_2} \max(T_t - c, 0) \tag{4.57}$$

and

$$HDD(\tau_1, \tau_2) = \sum_{t=\tau_1}^{\tau_2} \max(c - T_t, 0), \tag{4.58}$$

respectively. Here, we take the threshold temperature c as 18 degrees Celsius.

All futures are settled immediately after the measurement period has terminated. A futures contract on an index over a given measurement period is settled against the index value times a cash amount. For the United States cities, this cash amount is USD 20 per index point. We assume that the cash amount is USD 1 per index point. Remember that by definition of a futures contract, it is costless to enter. At the end of the measurement period, the buyer of the CDD futures contract receives the amount in Equation (4.57) and pays the CDD futures price. Similarly, the buyer of the HDD futures contract receives the amount in Equation (4.58) and pays the HDD futures price.

Now, the Monte Carlo simulation is done as follows: For the daily average deseasonalized temperature Y_t, we consider the proposed DRS model given in (3.14). First, 10000 trajectories for the daily average deseasonalized temperature Y_t are simulated by using the parameter estimates in Table 3.6. The simulated trajectories begins at 1 January 2011 and ends at 31 March 2013, that is, they are of length 820. Remember that the data from 1 January 2001 to 31 December 2010 is used for estimation. Then, the daily average temperature T_t is obtained by adding the seasonality component to Y_t. That is, $T_t = \Lambda_t + Y_t$, where Λ_t is given by (3.1). Then, the corresponding payoff, see Section 4.2, is calculated for each trajectory and averaging over all trajectories, the corresponding futures prices are found. The results together with a comparison of RS and CRS models, see Section 3.4, are given below.

Figure 4.1 to Figure 4.3 display the expected CDD index value obtained via RS, CRS and DRS models together with the actual CDD index from May to September, respectively. Moreover, Figure 4.4 to Figure 4.6 depict the expected HDD index value obtained via RS, CRS and DRS models together with the actual HDD index from November to March, respectively.

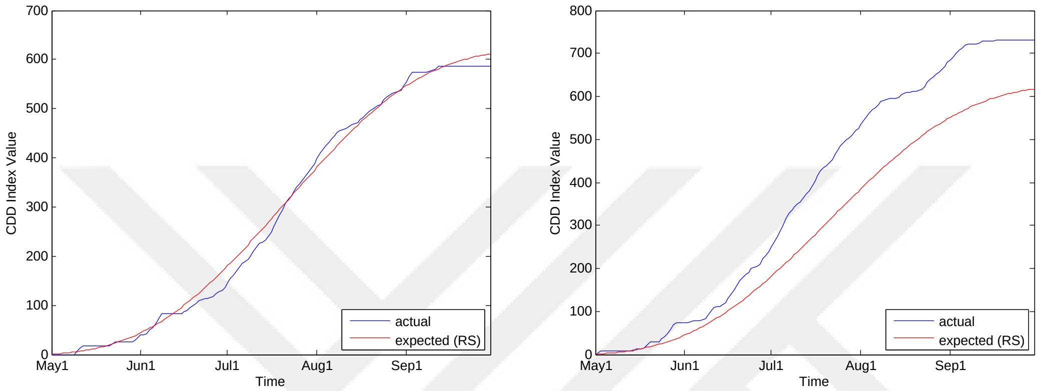


Figure 4.1: RS model expected CDD index value, 2011(left) and 2012(right).

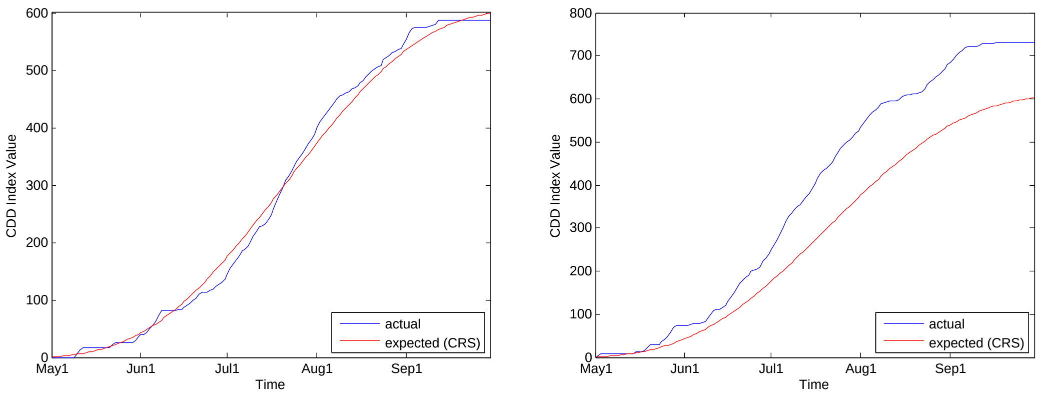


Figure 4.2: CRS model expected CDD index value, 2011(left) and 2012(right).

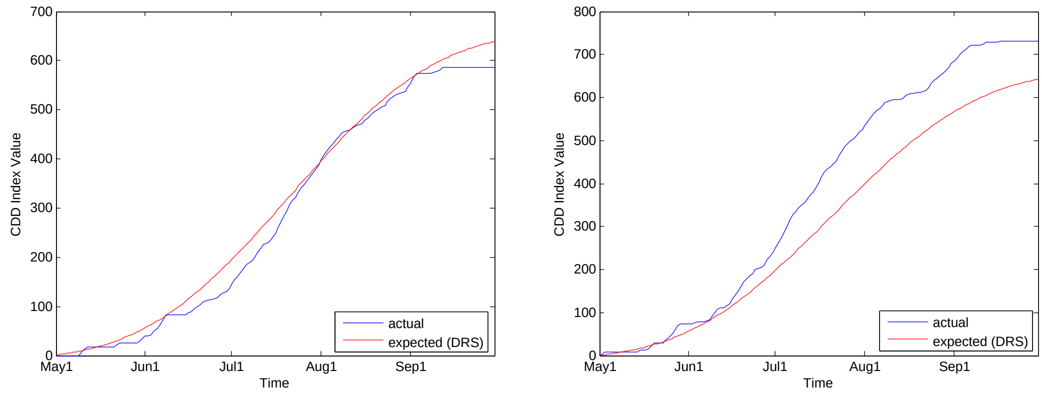


Figure 4.3: DRS model expected CDD index value, 2011(left) and 2012(right).

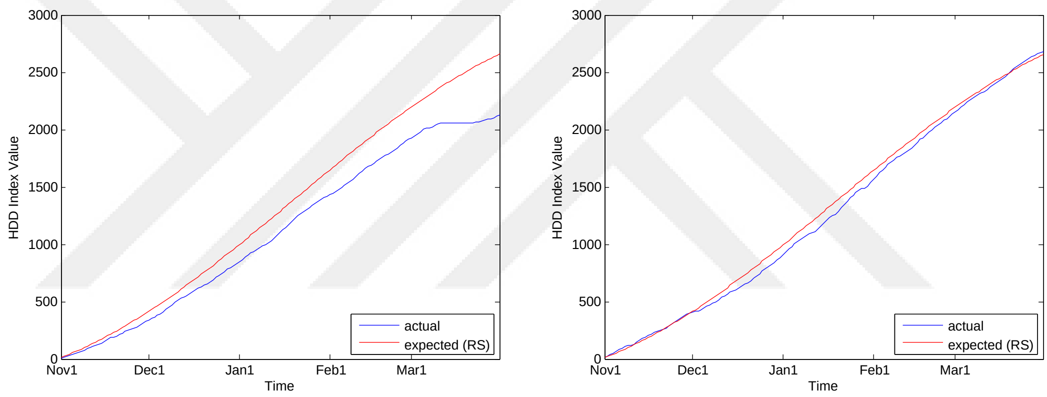


Figure 4.4: RS model expected HDD index value, 2011(left) and 2012(right).

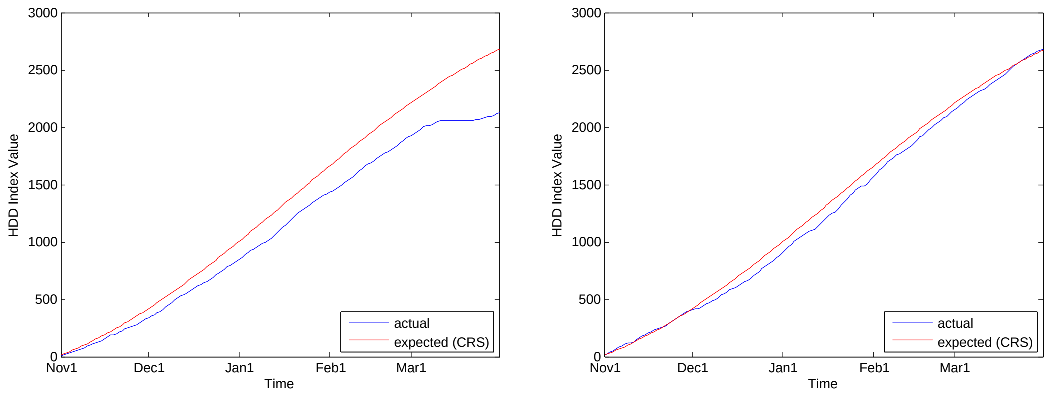


Figure 4.5: CRS model expected HDD index value, 2011(left) and 2012(right).

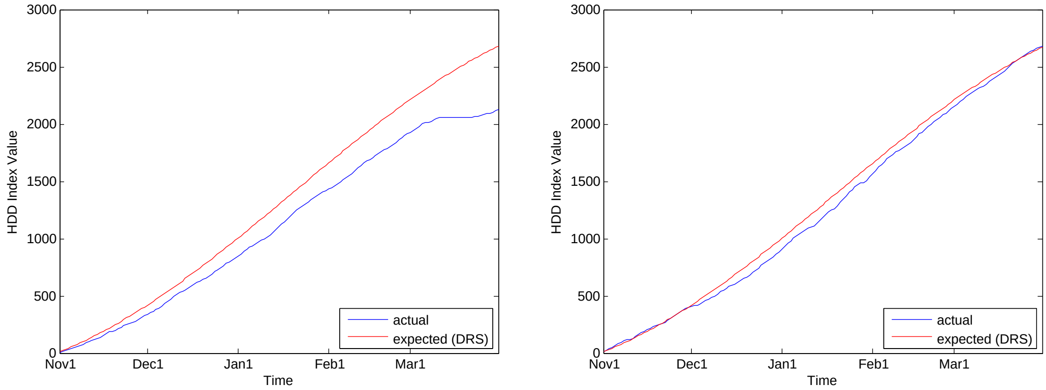


Figure 4.6: DRS model expected HDD index value, 2011(left) and 2012(right).

Table 4.1 to Table 4.3 and Table 4.4 to Table 4.6, display the futures prices obtained and actual index values together with difference. The difference is defined as the value obtained from the model minus the actual value. It can be said that while in 2011, the obtained CDD prices are very close to the actual values, in 2012, the obtained HDD prices are very close to the actual values.

As it can be seen from Table 4.1 to Table 4.3, in 2011 and 2012, for each CDD month and July-August seasonal strip, all the models considered give very close values. However, for the 2011 May-September seasonal strip, the actual value is 586.25, while the CRS, RS and DRS models produce a value of 599.2, 611.5 and 639.27, respectively. On the other hand, for the 2012 May-September seasonal strip, the actual value is 730.7, while the DRS, RS and CRS models produce a value of 642.88, 617.4, and 602.52, respectively.

Table 4.1: CDD futures prices based on the RS model.

	RS model (2011)	Actual (2011)	Diff. (2011)	RS model (2012)	Actual (2012)	Diff. (2012)
May	41.523	35.95	5.5725	42.323	74.25	-31.927
Jun	132.73	99.9	32.833	133.71	165.8	-32.088
Jul	200.29	252.45	-52.157	202.87	286.15	-83.279
Aug	168.48	156.25	12.226	168.8	153	15.796
Sep	68.471	41.7	26.771	69.695	51.5	18.195
Jul-Aug	368.77	408.7	-39.931	371.67	439.15	-67.483
May-Sep	611.5	586.25	25.245	617.4	730.7	-113.3

Table 4.2: CDD futures prices based on the CRS model.

	CRS model (2011)	Actual (2011)	Diff. (2011)	CRS model (2012)	Actual (2012)	Diff. (2012)
May	39.817	35.95	3.8669	40.346	74.25	-33.904
Jun	130.05	99.9	30.151	130.56	165.8	-35.238
Jul	197.19	252.45	-55.256	199.51	286.15	-86.638
Aug	164.77	156.25	8.5242	165.25	153	12.253
Sep	67.36	41.7	25.66	66.845	51.5	15.345
Jul-Aug	361.97	408.7	-46.732	364.76	439.15	-74.385
May-Sep	599.2	586.25	12.946	602.52	730.7	-128.18

Table 4.3: CDD futures prices based on the DRS model.

	DRS model (2011)	Actual (2011)	Diff. (2011)	DRS model (2012)	Actual (2012)	Diff. (2012)
May	53.394	35.95	17.444	53.826	74.25	-20.424
Jun	136.87	99.9	36.968	138.72	165.8	-27.082
Jul	199.92	252.45	-52.525	200.77	286.15	-85.382
Aug	170.25	156.25	14.001	170.13	153	17.134
Sep	78.837	41.7	37.137	79.435	51.5	27.935
Jul-Aug	370.18	408.7	-38.524	370.9	439.15	-68.248
May-Sep	639.27	586.25	53.024	642.88	730.7	-87.819

As it can be seen from Table 4.4 to Table 4.6, in 2011 and 2012, for each HDD month and December-February seasonal strip, all the models considered give very close values. However, for the 2011 November-March seasonal strip, the actual value is 2126.1, while the RS, DRS and CRS models produce a value of 2659.4, 2679.6 and 2680.1, respectively. On the other hand, for the 2012 November-March seasonal strip, the actual value is 2679, while the CRS, DRS and RS models produce a value of 2671.9, 2671.8 and 2653.1, respectively.

Table 4.4: HDD futures prices based on the RS model.

	RS model (2011)	Actual (2011)	Diff. (2011)	RS model (2012)	Actual (2012)	Diff. (2012)
Nov	397.85	325.05	72.795	396.08	398.1	-2.0239
Dec	578.09	502.65	75.442	576.49	487.15	89.344
Jan	651.63	588.9	62.733	651.3	651.15	0.15055
Feb	554.44	498.25	56.193	552.43	595.7	-43.27
Mar	477.36	211.2	266.16	476.78	546.95	-70.169
Dec-Feb	1784.2	1589.8	194.37	1780.2	1734	46.225
Nov-Mar	2659.4	2126.1	533.32	2653.1	2679	-25.968

Table 4.5: HDD futures prices based on the CRS model.

	CRS model (2011)	Actual (2011)	Diff. (2011)	CRS model (2012)	Actual (2012)	Diff. (2012)
Nov	401.47	325.05	76.421	399.33	398.1	1.2285
Dec	582.41	502.65	79.764	581.83	487.15	94.68
Jan	657.53	588.9	68.63	654.48	651.15	3.3254
Feb	557.44	498.25	59.189	557.09	595.7	-38.611
Mar	481.26	211.2	270.06	479.16	546.95	-67.789
Dec-Feb	1797.4	1589.8	207.58	1793.4	1734	59.394
Nov-Mar	2680.1	2126.1	554.06	2671.9	2679	-7.1662

Table 4.6: HDD futures prices based on the DRS model.

	DRS model (2011)	Actual (2011)	Diff. (2011)	DRS model (2012)	Actual (2012)	Diff. (2012)
Nov	402.9	325.05	77.848	401.44	398.1	3.3392
Dec	580.88	502.65	78.227	580.8	487.15	93.647
Jan	656.06	588.9	67.16	653.77	651.15	2.6189
Feb	558.06	498.25	59.809	556.68	595.7	-39.02
Mar	481.72	211.2	270.52	479.1	546.95	-67.851
Dec-Feb	1795	1589.8	205.2	1791.2	1734	57.246
Nov-Mar	2679.6	2126.1	553.56	2671.8	2679	-7.2656

CHAPTER 5

CONCLUSION AND OUTLOOK

Weather derivatives can be used as a tool to manage the risk exposure towards adverse or unexpected weather conditions. We consider the temperature based weather derivatives. These are the financial contracts written on several temperature indices such as CAT, CDD and HDD.

To derive derivatives prices, we first model the temperature dynamics. In the literature, the mean-reverting Ornstein–Uhlenbeck process in different forms is commonly used for modeling the temperature. We propose a regime-switching model for the temperature dynamics, where the parameters depend on a Markov chain. Also, since the jumps in the temperature are directly related to the regime switch, we model them by the chain itself. The jumps can be considered as the shifts in the level of the temperature due to the transitions of the state of the atmospheric conditions.

Moreover, the estimation of the proposed model is considered. The daily temperature data (in degrees Celsius) for Chicago O'Hare International Airport, USA, over the period from 1 January 2001 to 31 December 2013 is analysed. We consider various existing models and the proposed model. The models are estimated for the period from 1 January 2001 to 31 December 2010. The EM algorithm, which is an alternative method for maximizing the likelihood function for models with missing observations or unobserved variables, is considered. To estimate the proposed model, we modify the EM algorithm. We apply repeatedly the modified EM algorithm until the jump times and transition times coincide. And we forecast the models for the period from 1 January 2011 to 31 December 2013. We compare forecast performance of the proposed model with the existing models and conclude that the forecast performance of the proposed model is in line with existing models considered..

After modeling the temperature dynamics, to price the derivatives, the risk-neutral probability is to be specified. Since temperature (and thus the index) is not a tradeable asset, any probability measure being equivalent to the objective probability is a risk-neutral probability. A generalized version of the Esscher transform is considered to select an equivalent measure. Then the prices of weather derivatives written on several temperature indices are derived using the temperature model proposed.

As a further study, the proposed model can be extended by allowing the speed of mean-reversion parameter to depend on the Markov chain. Moreover, in the proposed model dynamics the rate matrix of the chain can be taken as time varying.

REFERENCES

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

[2] A. K. Alexandridis and A. D. Zapranis, The weather derivatives market, in *Weather Derivatives*, pp. 1–20, Springer, 2013.

[3] O. E. Barndorff-Nielsen and N. Shephard, Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics, Journal of the Royal Statistical Society, Series B (Statistical Methodology), 63(2), pp. 167–241, 2001.

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

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

[6] F. E. Benth and J. Šaltytė Benth, Weather derivatives and stochastic modelling of temperature, International Journal of Stochastic Analysis, 2011, 2011.

[7] F. E. Benth and J. Šaltytė Benth, *Modeling and Pricing in Financial Markets for Weather Derivatives*, volume 17, World Scientific, 2013.

[8] F. E. Benth, J. Šaltytė Benth, and S. Koekebakker, Putting a price on temperature, Scandinavian Journal of Statistics, 34(4), pp. 746–767, 2007.

[9] D. C. Brody, J. Syroka, and M. Zervos, Dynamical pricing of weather derivatives, Quantitative Finance, 2(3), pp. 189–198, 2002.

[10] O. Cappé, E. Moulines, and T. Rydén, *Inference in Hidden Markov Models*, New York, Springer, 2005.

[11] P. Carr and D. Madan, Option valuation using the fast Fourier transform, Journal of Computational Finance, 2(4), pp. 61–73, 1999.

[12] CME, *Weather Products, http://www.cmegroup.com/trading/weather/.*

[13] A. P. Dempster, N. M. Laird, and D. B. Rubin, Maximum likelihood from incomplete data via the EM algorithm, Journal of the Royal Statistical Society. Series B (Methodological), 39(1), pp. 1–38, 1977, ISSN 00359246.

[14] F. X. Diebold, Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of Diebold–Mariano tests, Journal of Business & Economic Statistics, 33(1), pp. 1–1, 2015.

[15] F. X. Diebold and R. S. Mariano, Comparing predictive accuracy, Journal of Business & Economic Statistics, 2012.

[16] F. Dornier and M. Queruel, Caution to the wind, Energy & Power Risk Management, 13(8), pp. 30–32, 2000.

[17] F. Dufour and R. Elliott, Filtering with discrete state observations, Applied Mathematics and Optimization, 40(2), pp. 259–272, 1999.

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

[19] R. J. Elliott, L. Aggoun, and J. B. Moore, *Hidden Markov Models: Estimation and Control*, volume 29, Springer Science & Business Media, 1995.

[20] W. Enders and T. Doan, *RATS Programming Manual*, 2014.

[21] G. Folland, *Real Analysis: Modern Techniques and Their Applications Pure and Applied Mathematics*, John Wiley & Sons, Inc., New York, 1984.

[22] P. H. Franses and D. Van Dijk, *Non-linear Time Series Models in Empirical Finance*, Cambridge University Press, 2000.

[23] P. Glasserman, *Monte Carlo Methods in Financial Engineering*, volume 53, Springer Science & Business Media, 2003.

[24] J. D. Hamilton, A new approach to the economic analysis of nonstationary time series and the business cycle, Econometrica: Journal of the Econometric Society, pp. 357–384, 1989.

[25] J. D. Hamilton, Analysis of time series subject to changes in regime, Journal of Econometrics, 45(1), pp. 39–70, 1990.

[26] R. J. Hyndman and A. B. Koehler, Another look at measures of forecast accuracy, International Journal of Forecasting, 22(4), pp. 679–688, 2006.

[27] J. Janczura and R. Weron, *http://ideas.repec.org/s/wuu/hscode.html*, 2011.

[28] J. Janczura and R. Weron, Efficient estimation of Markov regime-switching models: An application to electricity spot prices, AStA Advances in Statistical Analysis, 96(3), pp. 385–407, 2012.

[29] J. Kallsen and A. N. Shiryaev, The cumulant process and Esscher's change of measure, Finance and Stochastics, 6(4), pp. 397–428, 2002.

[30] C. J. Kim, Dynamic linear models with Markov-switching, Journal of Econometrics, 60(1), pp. 1 – 22, 1994, ISSN 0304-4076.

[31] R. S. Mamon and R. J. Elliott, *Hidden Markov Models in Finance*, volume 104, Springer Science & Business Media, 2007.

[32] R. S. Mamon and R. J. Elliott, *Hidden Markov Models in Finance*, Springer, 2014.

[33] M. Mraoua and D. Bari, Temperature stochastic modeling and weather derivatives pricing: Empirical study with Moroccan data, Afrika Statistika, 2(1), 2007.

[34] NCDC, *http://www.ncdc.noaa.gov/oa/ncdc.html*, National Climatic Data Center.

[35] C. R. Nelson and C. J. Kim, *State-Space Models with Regime-Switching*, MIT Press Edition, 1999.

[36] J. R. Norris, *Markov Chains*, Cambridge University Press, 1998.

[37] Y. Shen, K. Fan, and T. K. Siu, Option valuation under a double regime-switching model, Journal of Futures Markets, 34(5), pp. 451–478, 2014.

[38] A. Swishchuk and K. Cui, Weather derivatives with applications to Canadian data, Journal of Mathematical Finance, 3, pp. 81–95, 2013.

[39] H. M. Taylor and S. Karlin, *An Introduction to Stochastic Modeling*, Academic Press, New York, 1998.

[40] H. Tong, Threshold models, in *Threshold Models in Non-linear Time Series Analysis*, pp. 59–121, Springer, 1983.

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

APPENDIX A

AN OVERVIEW OF MARKOV CHAINS

We work throughout with a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. This part mainly follows from [39] and [36].

A.1 Discrete-Time Markov Chains

A *Markov process* $(\zeta_t)_{t\geq 0}$ is a stochastic process with the property that, given the value of ζ_t, the values of ζ_s for $s > t$ are not influenced by the values of ζ_u for $u < t$. That is, the probability of any particular future behavior of the process, when its current state is known exactly, is not altered by additional knowledge concerning its past behavior.
A *discrete-time Markov chain* is a Markov process whose state space is a finite or countable set, and whose (time) index set is $\{0, 1, 2, ...\}$. In formal terms, the *Markov property* is that

$$\begin{aligned}&\mathbb{P}[\zeta_{n+1} = j | \zeta_0 = i_0, \ldots, \zeta_{n-1} = i_{n-1}, \zeta_n = i] \\ &= \mathbb{P}[\zeta_{n+1} = j | \zeta_n = i],\end{aligned} \tag{A.1}$$

for all time points n and all states $i_0, \ldots, i_{n-1}, i, j$.
It is frequently convenient to label the state space of the Markov chain by $\{0, 1, 2, ...\}$, and it is customary to speak of ζ_n as being in state i if $\zeta_n = i$.
The probability of ζ_{n+1} being in state j given that ζ_n is in state i is called the *one-step transition probability* and is denoted by $p_{ij}^{n,n+1}$. That is,

$$p_{ij}^{n,n+1} = \mathbb{P}[\zeta_{n+1} = j | \zeta_n = i]. \tag{A.2}$$

In general, the transition probabilities are functions of not only the initial and final states, but also the time of transition as well. When the one-step transition probabilities are independent of the time variable n, it is said that the Markov chain has *stationary transition probabilities*. We consider Markov chains having stationary transition probabilities. Then $p_{ij}^{n,n+1} = p_{ij}$ is independent of n, and p_{ij} is the conditional probability that the state value undergoes a transition from i to j in one trial. It is customary to arrange these numbers p_{ij} in a matrix, in the infinite square array

$$\mathbf{P} = \begin{bmatrix} p_{00} & p_{01} & p_{02} & p_{03} & \cdots \\ p_{10} & p_{11} & p_{12} & p_{13} & \cdots \\ p_{20} & p_{21} & p_{22} & p_{23} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \\ p_{i0} & p_{i1} & p_{i2} & p_{i3} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{bmatrix}$$

and refer to $\mathbf{P} = [p_{ij}]$ as the Markov matrix or *transition probability matrix* of the process.

The $(i+1)$th row of $\mathbf{P}$ is the probability distribution of the values of ζ_{n+1} under the condition that $\zeta_n = i$, for $i = 0, 1, \ldots$. If the number of states is finite, then $\mathbf{P}$ is a finite square matrix whose order (the number of rows) is equal to the number of states. Clearly, the quantities p_{ij} satisfy the conditions

$$p_{ij} \geq 0, \qquad \text{for } i, j = 0, 1, 2, \ldots, \tag{A.3}$$

$$\sum_{j=0}^{\infty} p_{ij} = 1, \qquad \text{for } i = 0, 1, 2, \ldots. \tag{A.4}$$

The condition given by Equation (A.4) merely expresses the fact that some transition occurs at each trial. (For convenience, one says that a transition has occurred even if the state remains unchanged.)

We write $p_{i_n i_{n+1}} = \mathbb{P}[\zeta_{n+1} = i_{n+1} | \zeta_n = i_n]$.

It can be easily shown that

$$\mathbb{P}[\zeta_0 = i_0, \zeta_1 = i_1, \ldots, \zeta_n = i_n] = p_{i_0} p_{i_0 i_1} \cdots p_{i_{n-2} i_{n-1}} p_{i_{n-1} i_n}, \tag{A.5}$$

where $p_i = \mathbb{P}[\zeta_0 = i]$. Thus, a Markov process is completely defined once its transition probability matrix and initial state ζ_0 (or, more generally, the probability distribution of ζ_0) are specified.
Moreover, Equation (A.1) is equivalent to the Markov property in the form

$$\begin{aligned} &\mathbb{P}[\zeta_{n+1} = j_1, \ldots, \zeta_{n+m} = j_m | \zeta_0 = i_0, \ldots, \zeta_n = i_n] \\ &= \mathbb{P}[\zeta_{n+1} = j_1, \ldots, \zeta_{n+m} = j_m | \zeta_n = i_n], \end{aligned} \tag{A.6}$$

for all time points n, m and all states $i_0, \ldots, i_n, j_1, \ldots, j_m$. In other words, once Equation (A.6) is established for the value $m = 1$, it holds for all $m \geq 1$.

The formal definition of a Markov chain is given below.

Definition A.1. Let I be a countable set. Each $i \in I$ is called a *state* and I is called the *state-space*. $(\zeta_n)_{n \geq 0}$ is a Markov chain with initial distribution λ and transition matrix $\mathbf{P} = (p_{ij})_{i,j \in I}$ if

(i) ζ_0 has distribution λ,

(ii) for $n \geq 0$, conditional on $\zeta_n = i$, ζ_{n+1} has distribution $p_{ij}, j \in I$, and is independent of $\zeta_0, \ldots, \zeta_{n-1}$.

More explicitly, these conditions state that for $n \geq 0$ and $i_0, \ldots, i_{n+1} \in I$,

(i) $\mathbb{P}[\zeta_0 = i_0] = \lambda_{i_0}$,

(ii) $\mathbb{P}[\zeta_{n+1} = i_{n+1} | \zeta_0 = i_0, \ldots, \zeta_n = i_n] = p_{i_n i_{n+1}}$.

We say that $(\zeta_n)_{n\geq 0}$ is $Markov(\lambda, \mathbf{P})$ for short. If $(\zeta_n)_{0\leq n\leq N}$ is a finite sequence of random variables satisfying (i) and (ii) for $n = 0, 1, \ldots, N-1$, then we again say $(\zeta_n)_{0\leq n\leq N}$ is $Markov(\lambda, \mathbf{P})$.

Theorem A.1. *A discrete-time random process* $(\zeta_n)_{0\leq n\leq N}$ *is* $Markov(\lambda, \mathbf{P})$ *if and only if for all* $i_0, \ldots, i_N \in I$

$$\mathbb{P}[\zeta_0 = i_0, \zeta_1 = i_1, \ldots, \zeta_N = i_N] = \lambda_{i_0} p_{i_0 i_1} p_{i_1 i_2} \ldots p_{i_{N-1} i_N}, \tag{A.7}$$

where $\lambda_{i_0} = \mathbb{P}[\zeta_0 = i_0]$.

The analysis of a Markov chain concerns mainly the calculation of the probabilities of the possible realizations of the process. Central in these calculations are the n-step transition probability matrices $\mathbf{P}^{(n)} = [p_{ij}^{(n)}]$. Here, $p_{ij}^{(n)}$ denotes the probability that the process goes from state i to state j in n transitions. Formally,

$$p_{ij}^{(n)} = \mathbb{P}[\zeta_{m+n} = j | \zeta_m = i]. \tag{A.8}$$

Observe that we are dealing only with temporally homogeneous processes having stationary transition probabilities, since otherwise the left side of Equation (A.8) would also depend on m.
The Markov property allows us to express Equation (A.8) in terms of $[p_{ij}]$ as stated in the following theorem.

Theorem A.2. *The n-step transition probabilities of a Markov chain satisfy*

$$p_{ij}^{(n)} = \sum_{k=0}^{\infty} p_{ik} p_{kj}^{(n-1)}, \tag{A.9}$$

where we define

$$p_{ij}^{(0)} = \begin{cases} 1, & \textit{if } i = j, \\ 0, & \textit{if } i \neq j. \end{cases}$$

Remark A.1. We recognize the relation given by Equation (A.9) as the formula for matrix multiplication, so that $\mathbf{P}^{(n)} = \mathbf{P} \cdot \mathbf{P}^{(n-1)}$. By iterating this formula, we obtain

$$\mathbf{P}^{(n)} = \underbrace{\mathbf{P} \cdot \mathbf{P} \cdot \cdots \cdot \mathbf{P}}_{n \text{ factors}} = \mathbf{P}^n. \tag{A.10}$$

That is, the n-step transition probabilities $p_{ij}^{(n)}$ are the entries in the matrix $\mathbf{P}^n$, the nth power of $\mathbf{P}$.

We write $p_{ij}^{(n)}$ for the (i,j) entry in the matrix $\mathbf{P}^n$, the nth power of $\mathbf{P} = [p_{ij}]$.

Example A.1. Let

$$\mathbf{P} = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix},$$

where $0 < a, b < 1$, be the transition matrix of a two-state Markov chain $(\zeta_n)_{n\geq 0}$. When $a = 1-b$, so that the rows of $\mathbf{P}$ are the same, then $\zeta_1, \zeta_2, \ldots$ are independent identically distributed random variables with $\mathbb{P}[\zeta_n = 0] = b$ and $\mathbb{P}[\zeta_n = 1] = a$. When $a \neq 1-b$, the probability distribution for ζ_n varies depending on the outcome ζ_{n-1} at the previous stage.
For the two-state Markov chain, the n-step transition matrix is given by

$$\mathbf{P}^n = \frac{1}{a+b}\begin{bmatrix} b & a \\ b & a \end{bmatrix} + \frac{(1-a-b)^n}{a+b}\begin{bmatrix} a & -a \\ -b & b \end{bmatrix}.$$

Note that $|1-a-b| < 1$ when $0 < a, b < 1$, and thus $|1-a-b|^n \to 0$ as $n \to \infty$ and

$$\lim_{n\to\infty} \mathbf{P}^n = \begin{bmatrix} \dfrac{b}{a+b} & \dfrac{a}{a+b} \\ \dfrac{b}{a+b} & \dfrac{a}{a+b} \end{bmatrix}.$$

That is, such a system, in the long run, will be in state 0 with probability $b/(a+b)$ and in state 1 with probability $a/(a+b)$, irrespective of the initial state in which the system started.

Definition A.2. State j is said to be accessible from state i if for some integer $n \geq 0$, $p_{ij}^{(n)} > 0$, that is, there is positive probability that state j can be reached starting from state i in some finite number of transitions. Two states i and j, each accessible to the other, are said to communicate, and we write $i \leftrightarrow j$. If two states i and j do not communicate, then either

$$p_{ij}^{(n)} = 0, \qquad \text{for all } n \geq 0,$$

or

$$p_{ji}^{(n)} = 0, \qquad \text{for all } n \geq 0,$$

or both are true. A Markov chain is irreducible if all states communicate with each other.

A.2 Continuous-Time Markov Chains

A continuous-time Markov chain $(\zeta_t)_{t>0}$ is a Markov process on the states $0, 1, 2, \ldots$. The exponential distribution plays a fundamental role in continuous-time Markov chains because of the memorylessness property.

Definition A.3. A random variable $X : \Omega \to [0, \infty]$ has exponential distribution of parameter $\lambda,\ 0 \leq \lambda < \infty$, if

$$\mathbb{P}(X > t) = e^{-\lambda t}, \qquad \text{for all } t \geq 0.$$

If $\lambda > 0$, then X has density function

$$f_X(t) = \lambda e^{-\lambda t} \mathbf{1}_{t \geq 0}.$$

The mean of X is given by

$$\mathbb{E}(X) = \lambda^{-1}.$$

Theorem A.3. ***(Memorylessness property)*** *A random variable $X : \Omega \to (0, \infty]$ has exponential distribution if and only if it has the following memoryless property:*

$$\mathbb{P}[X > s + t | X > s] = \mathbb{P}(X > t), \qquad \textit{for all } s, t \geq 0.$$

Definition A.4. Let I be a countable set. A Q-matrix on I is a matrix $\mathbf{Q} = (q_{ij})_{i,j \in I}$ satisfying the following conditions:

(i) $0 \leq -q_{ii} < \infty$, for all i,

(ii) $q_{ij} \geq 0$, for all $i \neq j$,

(iii) $\sum_{j \in I} q_{ij} = 0$, for all i.

Note that in each row of $\mathbf{Q}$ we can choose the off-diagonal entries to be any nonnegative real number, subject only to the constraint that the off-diagonal row sum is finite, that is,

$$q_i = \sum_{j \neq i} q_{ij} < \infty.$$

The diagonal entry q_{ii} is then $-q_i$, making the total row sum zero.

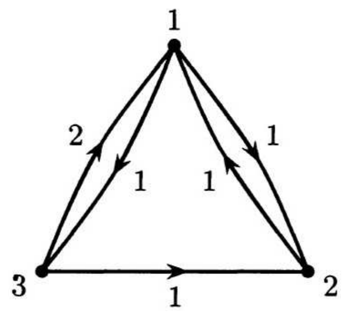
We set $q_i = -q_{ii}$.

For the state space $\{0, 1, 2, \ldots, N\}$, we have

$$\mathbf{Q} = \begin{bmatrix} -q_0 & q_{01} & \cdots & q_{0N} \\ q_{10} & -q_1 & \cdots & q_{1N} \\ \vdots & \vdots & \vdots & \vdots \\ q_{N0} & q_{N1} & \cdots & -q_N \end{bmatrix}.$$

We call each off-diagonal entry q_{ij} *the rate of going from* i to j, and q_i *the rate of leaving* i. A convenient way to present the data for a continuous-time Markov chain is by means of a diagram. Each diagram then corresponds to a unique Q-matrix.

Example A.2. Consider the following diagram.



The associated Q-matrix is given by

$$\mathbf{Q} = \begin{bmatrix} -2 & 1 & 1 \\ 1 & -1 & 0 \\ 2 & 1 & -3 \end{bmatrix}.$$

Theorem A.4. *Let $\mathbf{Q}$ be a matrix on a finite set I. Set $\mathbf{P}(t) = e^{t\mathbf{Q}}$. Then $\mathbf{P}(t), t \geq 0$, has the following properties:*

(i) $\mathbf{P}(s+t) = \mathbf{P}(s)\mathbf{P}(t)$, *for all* s, t,

(ii) $\mathbf{P}(t), t \geq 0$, *is the unique solution to the forward equation*

$$\frac{d\mathbf{P}(t)}{dt} = \mathbf{P}(t)\mathbf{Q}, \qquad \mathbf{P}(0) = \mathbf{I},$$

(iii) $\mathbf{P}(t), t \geq 0$, *is the unique solution to the backward equation*

$$\frac{d\mathbf{P}(t)}{dt} = \mathbf{Q}\mathbf{P}(t), \qquad \mathbf{P}(0) = \mathbf{I},$$

(iv) for $k = 0, 1, 2, \ldots,$ *we have*

$$\frac{d^k\mathbf{P}(t)}{dt^k} \,|_{t=0} = \mathbf{Q}^k.$$

Definition A.5. A matrix $\mathbf{P} = (p_{ij})_{i,j \in I}$ is stochastic if it satisfies

(i) $0 \leq p_{ij} < \infty$, for all i, j,

(ii) $\sum_{j \in I} p_{ij} = 1$, for all i.

Theorem A.5. *A matrix $\mathbf{Q}$ on a finite set I is a Q-matrix if and only if $\mathbf{P}(t) = e^{t\mathbf{Q}}$ is a stochastic matrix for all $t \geq 0$.*

Definition A.6. The jump matrix $\mathbf{\Pi} = (\pi_{ij})_{i,j\in I}$ of a Q-matrix $\mathbf{Q} = (q_{ij})_{i,j\in I}$ is defined by

$$\pi_{ij} = \begin{cases} q_{ij}/q_i, & \text{if } i \neq j \text{ and } q_i \neq 0, \\ 0, & \text{if } i \neq j \text{ and } q_i = 0, \end{cases}$$

$$\pi_{ii} = \begin{cases} 0, & \text{if } q_i \neq 0, \\ 1, & \text{if } q_i = 0. \end{cases}$$

Note that $\mathbf{\Pi}$ is a stochastic matrix.

Definition A.7. $(\zeta_t)_{t\geq 0}$ is called a right-continuous process if for all $\omega \in \Omega$ and $t \geq 0$ there exists $\varepsilon > 0$ such that

$$\zeta_s(\omega) = \zeta_t(\omega), \qquad \text{for } t \leq s \leq t + \varepsilon.$$

The jump times $J_0, J_1, \ldots$ of $(\zeta_t)_{t\geq 0}$ are given by

$$J_0 = 0, \; J_{n+1} = \inf\{t \geq J_n : \; \zeta_t \neq \zeta_{J_n}\},$$

for $n = 0, 1, \ldots$ and $\inf\{\emptyset\} = \infty$.
The holding times $S_1, S_2, \ldots$ of $(\zeta_t)_{t\geq 0}$ are given by

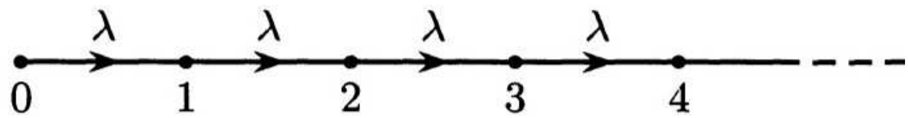
$$S_n = \begin{cases} J_n - J_{n-1}, & \text{if } J_{n-1} < \infty, \\ \infty, & \text{otherwise}, \end{cases}$$

for $n = 1, 2, \ldots$. The right-continuity forces $S_n > 0$ for all n. If $J_{n+1} = \infty$ for some n, we define $\zeta_\infty = \zeta_{J_n}$, the final value, otherwise ζ_∞ is undefined. The (first) explosion time ϑ is defined by

$$\vartheta = \sup_n J_n = \sum_{n=1}^{\infty} S_n.$$

The discrete-time process $(\xi_n)_{n\geq 0}$ given by $\xi_n = \zeta_{J_n}$ is called the jump process of $(\zeta_t)_{t\geq 0}$, or the jump chain if it is a discrete-time Markov chain. This is simply the sequence of values taken by $(\zeta_t)_{t\geq 0}$ up to explosion.

Example A.3. Poisson processes are some of the simplest examples of continuous-time Markov chains. A right-continuous process $(\zeta_t)_{t\geq 0}$ with values in $\{0, 1, 2, \ldots\}$ is a *Poisson process of rate* λ, $0 < \lambda < \infty$, if its holding times $S_1, S_2, \ldots$ are independent exponential random variables of parameter λ and its jump chain is given by $\xi_n = n$. For the diagram given below



the associated Q-matrix is given by

$$\mathbf{Q} = \begin{bmatrix} -\lambda & \lambda & & & \\ & -\lambda & \lambda & & \\ & & \ddots & \ddots & \\ & & & & \end{bmatrix},$$

where the entries off the diagonal and super-diagonal are all zero. A simple way to construct a Poisson process of rate λ is to take a sequence $S_1, S_2, \ldots$ of independent exponential random variables of parameter λ to set $J_0 = 0,\ J_n = S_1 + \cdots + S_n$ and then set

$$\zeta_t = n, \quad \text{if } J_n \leq t < J_{n+1}.$$

Remark A.2. $f(t) = o(t)$ means $\frac{f(t)}{t} \to 0$ as $t \to 0$.

Theorem A.6. *Let $(\zeta_t)_{t\geq 0}$ be a right-continuous process with values in a finite set I and λ be the distribution of ζ_0. Let $\mathbf{Q} = (q_{ij})_{i,j\in I}$ be a Q-matrix on I with jump matrix $\mathbf{\Pi} = (\pi_{ij})_{i,j\in I}$. Then the following three conditions are equivalent:*

1. *(jump chain/holding time definition) the jump chain $(\xi_n)_{n\geq 0}$ of $(\zeta_t)_{t\geq 0}$ is discrete-time Markov$(\lambda, \mathbf{\Pi})$ and for each $n \geq 1$, conditional on $\xi_0, \ldots, \xi_{n-1}$, the holding times $S_1, \ldots, S_n$ are independent exponential random variables of parameters $q_{\xi_0}, \ldots, q_{\xi_{n-1}}$, respectively;*

2. *(infinitesimal definition) for all $t, h \geq 0$, conditional on $\zeta_t = i$, ζ_{t+h} is independent of ζ_s, $s \leq t$, and as $h \downarrow 0$, uniformly in t, for all j*

$$\mathbb{P}[\zeta_{t+h} = j | \zeta_t = i] = \delta_{ij} + q_{ij}h + o(h);$$

3. *(transition probability definition) for all $n = 0, 1, 2, \ldots$, all times $0 \leq t_0 \leq t_1 \leq \cdots \leq t_{n+1}$ and all states $i_0, \ldots, i_{n+1}$*

$$\mathbb{P}[\zeta_{t_{n+1}} = i_{n+1} | \zeta_{t_0} = i_0, \ldots, \zeta_{t_n} = i_n] = p_{i_n i_{n+1}}(t_{n+1} - t_n), \tag{A.11}$$

where $p_{ij}(t), i, j \in I, t \geq 0$, is the solution of the forward equation

$$\mathbf{P}'(t) = \mathbf{P}(t)\mathbf{Q}, \qquad \mathbf{P}(0) = \mathbf{I}.$$

If $(\zeta_t)_{t\geq 0}$ satisfies any of these conditions, then it is called a Markov chain with initial distribution λ and generator matrix $\mathbf{Q}$. We say that $(\zeta_t)_{t\geq 0}$ is $Markov(\lambda, \mathbf{Q})$ for short.

Remember that, by Theorem A.4, for I finite, the forward and backward equations have the same solution. So in the above theorem, the forward equation can be replaced with the backward equation. Thus, we have

$$\mathbf{P}'(t) = \mathbf{P}(t)\mathbf{Q} = \mathbf{Q}\mathbf{P}(t), \tag{A.12}$$

where $\mathbf{P}'(t)$ denotes the matrix whose elements are $p'_{ij}(t) = \dfrac{dp_{ij}(t)}{dt}$. Then with the initial condition $\mathbf{P}(0) = \mathbf{I}$, we have

$$\mathbf{P}(t) = e^{t\mathbf{Q}} = \mathbf{I} + \sum_{n=1}^{\infty} \frac{t^n \mathbf{Q}^n}{n!} \tag{A.13}$$

and $\mathbf{Q}$ is the matrix derivative of $\mathbf{P}(t)$ at $t = 0$, that is, $\mathbf{Q} = \mathbf{P}'(0)$. Notice that for a continuous-time finite-state Markov chain $(\zeta_t)_{t\geq 0}$ with Q-matrix $\mathbf{Q}$, the *transition probability* from i to j in time t is given by

$$p_{ij}(t) = \mathbb{P}[\zeta_t = j | \zeta_0 = i], \tag{A.14}$$

where $p_{ij}(t)$ is (i, j) entry in $\mathbf{P}(t) = e^{t\mathbf{Q}}$.
Moreover, the holding time definition can also be expressed as follows: Starting in state i, the process waits there for a duration that is exponentially distributed with parameter q_i. The process then jumps to state $j \neq i$ with probability q_{ij}/q_i; the waiting time in state j is exponentially distributed with parameter q_j, and so on. The sequence of states visited by the process, denoted by $\xi_0, \xi_1, \ldots$, is a Markov chain with discrete parameter, called the *embedded Markov chain*. Conditioned on the state sequence $\xi_0, \xi_1, \ldots$, the successive holding times $S_1, S_2, \ldots$ are independent exponentially distributed random variables with parameters $q_{\xi_0}, q_{\xi_1}, \ldots$, respectively.

Example A.4. Consider a Markov chain $(\zeta_t)_{t\geq 0}$ with states $\{0, 1\}$ with generator matrix

$$\mathbf{Q} = \begin{bmatrix} -\alpha & \alpha \\ \beta & -\beta \end{bmatrix}.$$

The process alternates between states 0 and 1. The holding times in state 0 are independent and exponentially distributed with parameter α. Those in state 1 are independent and exponentially distributed with parameter β. It can be shown that

$$\mathbf{Q}^n = (-(\alpha + \beta))^{n-1}\mathbf{Q}.$$

Then, by Equation (A.13), we have

$$\mathbf{P}(t) = \begin{bmatrix} 1-\pi & \pi \\ 1-\pi & \pi \end{bmatrix} + \begin{bmatrix} \pi & -\pi \\ -(1-\pi) & 1-\pi \end{bmatrix} e^{-\tau t},$$

where $\pi = \alpha/(\alpha + \beta)$ and $\tau = \alpha + \beta$.

When a Markov chain on states $\{0, 1, \ldots, N\}$ is irreducible (all states communicate), then $p_{ij}(t) > 0$ for $i, j = 0, 1, \ldots, N$ and $\lim_{t\to\infty} p_{ij}(t) = \pi_j > 0$ exists independently of the initial state i. The limiting distribution may be found by passing to the limit in Equation (A.12), noting that $\lim_{t\to\infty} \mathbf{P}'(t) = 0$. The resulting equations for $\pi = \begin{bmatrix} \pi_0 & \pi_1 & \ldots & \pi_N \end{bmatrix}$ are

$$0 = \pi\mathbf{Q} = \begin{bmatrix} \pi_0 & \pi_1 & \ldots & \pi_N \end{bmatrix} \begin{bmatrix} -q_0 & q_{01} & \cdots & q_{0N} \\ q_{10} & -q_1 & \cdots & q_{1N} \\ \vdots & \vdots & \vdots & \vdots \\ q_{N0} & q_{N1} & \cdots & -q_N \end{bmatrix},$$

which is the same as

$$\pi_j q_j = \sum_{i \neq j} \pi_i q_{ij}, \qquad j = 0, 1, \ldots, N. \tag{A.15}$$

Then, the limiting distribution is determined by Equation (A.15) together with

$$\pi_0 + \pi_1 + \cdots + \pi_N = 1. \tag{A.16}$$

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Türkvatan, Aysun
Nationality: Turkish
Place of Birth: Ankara
e-mail: aysun.turkvatan@metu.edu.tr

EDUCATION

Degrree	Institution	Year of Graduation
Ph.D.	Middle East Technical University Institute of Applied Mathematics Department of Financial Mathematics	2016
M.S.	Middle East Technical University Institute of Applied Mathematics Department of Financial Mathematics	2008
B.S.	Middle East Technical University Department of Mathematics	2006

PROFESSIONAL EXPERIENCE

Year	Place	Enrollment
2009-Present	Central Bank of the Republic of Turkey	Assistant Specialist