UNIVERSITÉ D'ORLÉANS

ÉCOLE DOCTORALE MATHÉMATIQUES, INFORMATIQUE, PHYSIQUE THÉORIQUE ET INGÉNIERIE DES SYSTÈMES

LABORATOIRE : Mathématiques - Analyse, Probabilités, Modélisation - Orléans

THÈSE PRÉSENTÉE PAR :

VAN LY TRAN

soutenue le **12 décembre 2013**

pour obtenir le grade de : **Docteur de l'université d'Orléans**

Discipline/ Spécialité : **MATHÉMATIQUES APPLIQUÉES**

Modèles Stochastiques des Processus de Rayonnement Solaire

Stochastic Models of Solar Radiation Processes

THÈSE DIRIGÉE PAR :

Richard ÉMILION	Professeur, Université d'Orléans
Romain ABRAHAM	Professeur, Université d'Orléans

RAPPORTEURS :

Sophie DABO-NIANG	Professeur, Université de Lille
Jean-François DELMAS	Professeur, École des Ponts ParisTech

JURY :

Romain ABRAHAM	Professeur, Université d'Orléans
Didier CHAUVEAU	Professeur, Université d'Orléans, Président du jury
Sophie DABO-NIANG	Professeur, Université de Lille
Jean-Franccois DELMAS	Professeur, École des Ponts ParisTech
Richard ÉMILION	Professeur, Université d'Orléans
Philippe POGGI	Professeur, Université de Corse

P H D T H E S I S

PhD of Science

of the University of Orléans

Specialty : Applied Mathematics

Defended by

TRAN Van Ly

STOCHASTIC MODELS OF SOLAR RADIATION PROCESSES

Thesis Advisors : Richard Emilion and Romain Abraham

12th December, 2013

Jury :

Reviewers :	Sophie DABO-NIANG	-	University of Lille
	Jean-François DELMAS	-	École des Ponts ParisTech
Advisors :	Richard EMILION	-	University of Orleans
	Romain ABRAHAM	-	University of Orleans
President :	Didier CHAUVEAU	-	University of Orleans
Examinator :	Philippe POGGI	-	University of Corse

Remerciements

Plusieurs personnes m'ont aidé durant ce travail de thèse.

La première personne que je tiens à vivement remercier est mon Directeur de thèse, le professeur Richard Émilion, pour le choix du sujet, pour sa confiance en moi, sa patience, et son apport considérable sans lequel ces travaux n'auraient pas pu être menés à terme. Je lui suis reconnaissant pour tout le temps qu'il a consacré à répondre à mes questions et à corriger ma rédaction. Ce fut pour moi une expérience extrêmement enrichissante.

J'adresse mes vifs remerciements à mon codirecteur de thèse, le professeur Romain Abraham, Directeur du laboratoire MAPMO, pour le choix du sujet, pour ses explications et ses précieux conseils qui m'ont éclairé, pour son accueil et son aide dans le laboratoire durant toute ces années.

Je tiens à vivement remercier les professeurs Jean-François Delmas et Sophie Dabo-Niang d'avoir accepté et d'accomplir la délicate tâche de rapporteurs de cette thèse.

Mes vifs remerciements aux membres du jury d'avoir accepté d'évaluer ce travail de recherche.

Je remercie très spécialement Dr. Ted Soubdhan, Maître de conférences en physique à l'université d'Antilles-Guyane qui nous a introduit à la problématique de l'énergie solaire, a orienté nos recherches et a mis à notre disposition ses mesures de rayonnement solaire de la Guadeloupe.

Je remercie très spécialement M. Mathieu Delsaut, ingénieur logiciel, et toute l'équipe du projet RCI-GS de l'université de La Réunion, qui ont mis à notre disposition les mesures de rayonnement solaire de La Réunion.

Je tiens à remercier tous ceux qui m'ont aidé à obtenir le financement de cette thèse.

Je tiens à remercier Mesdames Anne Liger, Marie-France Grespier, Marie-Laurence Poncet, Marine Cizeau, M. Romain Theron et toutes les personnes du laboratoire MAPMO, pour leur accueil chaleureux et tout l'aide qu'ils m'ont apportée.

Ce travail de thèse aurait été impossible sans le soutien affectif de ma petite famille : ma femme Thao Nguyen et ma petite fille Anh Thu qui m'ont permis de persévérer toutes ces années. Je voudrais également remercier profondément mes parents, mes frères et mes soeurs, qui m'ont toujours aidé à chaque étape de mes études.

J'ai été grandement soutenu et encouragé par Nicole Nourry et mes amis : Vo Van Chuong, Ngoc Linh, Hong Dan, Minh Phuong, Loic Piffet, Sébastient Dutercq, Thuy Nga, Xuan Lan, Hiep Thuan, Trang Dai, Thuy Lynh, Thanh Binh, Xuan Hieu et d'autres que j'oublie de citer.

A eux tous, j'adresse mes plus sincères remerciements pour la réalisation de cette thèse.

Orléans, décembre 2013,
Van Ly TRAN

To my wife and my daughter

Résumé

Les caractéristiques des rayonnements solaire dépendent fortement de certains événements météorologiques **non observés** (fréquence, taille et type des nuages et leurs propriétés optiques; aérosols atmosphériques, albédo du sol, vapeur d'eau, poussière et turbidité atmosphérique) tandis qu'une séquence du rayonnement solaire peut être **observée** et mesurée à une station donné. Ceci nous a suggéré de modéliser les processus de rayonnement solaire (ou d'indice de clarté) en utilisant un modèle Markovien caché (HMM), paire corrélée de processus stochastiques.

Notre modèle principal est un HMM à temps continu $(X_t, y_t)_{t\geq 0}$ tel que (y_t), le processus observé de rayonnement, soit une solution de l'équation différentielle stochastique (EDS) :

$$dy_t = [g(X_t)I_t - y_t]dt + \sigma(X_t)y_t dW_t,$$

où I_t est le rayonnement extraterrestre à l'instant t, (W_t) est un mouvement Brownien standard et $g(X_t)$, $\sigma(X_t)$ sont des fonctions de la chaîne de Markov non observée (X_t) modélisant la dynamique des régimes environnementaux.

Pour ajuster nos modèles aux données réelles observées, les procédures d'estimation utilisent l'algorithme EM et la méthode du changement de mesures par le théorème de Girsanov. Des équations de filtrage sont établies et les équations à temps continu sont approchées par des versions *robustes*.

Les modèles ajustés sont appliqués à des fins de comparaison et classification de distributions et de prédiction.

Abstract

Characteristics of solar radiation highly depend on some **unobserved** meteorological events (frequency, height and type of the clouds and their optical properties; atmospheric aerosols, ground albedo, water vapor, dust and atmospheric turbidity) while a sequence of solar radiation can be **observed** and measured at a given station. This has suggested us to model solar radiation (or clearness index) processes using a hidden Markov model (HMM), a pair of correlated stochastic processes.

Our main model is a continuous-time HMM $(X_t, y_t)_{t \geq 0}$ such that the solar radiation process $(y_t)_{t \geq 0}$ is a solution of the stochastic differential equation (SDE):

$$dy_t = [g(X_t)I_t - y_t]dt + \sigma(X_t)y_t dW_t,$$

where I_t is the extraterrestrial radiation received at time t, (W_t) is a standard Brownian motion and $g(X_t)$, $\sigma(X_t)$ are functions of the unobserved Markov chain (X_t) modelling environmental regimes.

To fit our models to observed real data, the estimation procedures combine the Expectation Maximization (EM) algorithm and the measure change method due to Girsanov theorem. Filtering equations are derived and continuous-time equations are approximated by *robust* versions.

The models are applied to pdf comparison and classification and prediction purposes.

Introduction

Context

The aim of the present thesis is to propose some probabilistic models for sequences of solar radiation which is defined as the energy given off by the sun (W/m^2) at the earth suface. Our main model concerns a Stochastic Differential Equations (SDE) in random environment, the latter being modelized by a hidden Markov chain. Statistical fitting of such models hinges on filtering equations that we establish in order to update the estimations in the steps of EM algorithm. Experiments are done using real large datasets recorded by some terrestrial captors that have measured solar radiation.
Such a modelling problem is of greatest importance in the domain of renewable energy where short-term and very short-term time horizon prediction is a challenge, particularly in the domain of solar energy.

Random aspects

Probabilistic models turn out to be relevant as the measured solar radiation is actually a *global radiation*, or *total radiation*, which results from two components, a deterministic one and a random one, namely
- the *direct radiation* which is the energy coming through a straight line from the sun to a specific geographical position of the earth surface. At a given time this deterministic radiation can be computed quite precisely and as it roughly corresponds to a measurement during a perfectly *clear-sky* weather, it is also known as the *extra-terrestrial radiation*
- the *diffuse radiation* which is reflected by the environment and depends on meteorolgical conditions, and is therefore highly random.
Both components can be measured by captors.

The total solar radiation can also be studied indirectly by considering its dimensionless form, the so-called *clearness index* (CI), which is defined as the ratio of the total radiation to the direct radiation and thus is a nice descriptor of the atmospheric transmittance.
Our approach will therefore consist in considering a discrete (resp. continous) sequence of solar observations as a *path* of a discrete-time (resp. continuous-time) *stochastic process*.
The following figure (Figure 1) illustrates our arguments: the iobserved irregular falls are due to frequent cloud passages which depend on some random conditions such as wind speed, type of clouds and some other meteorological variables:

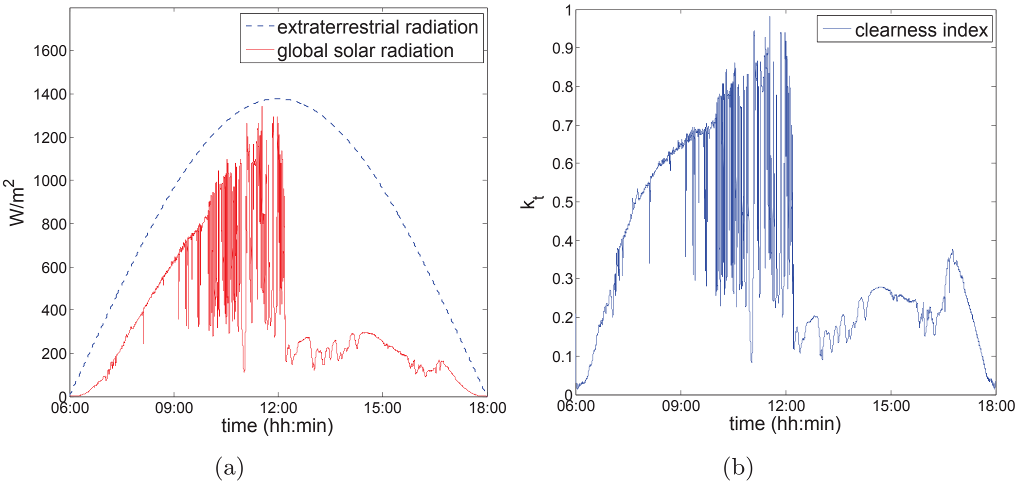


Figure 1: Measurements of total solar radiation and extraterrestrial radiation (a). Corresponding clearness index (b). [Soubdhan 2009]

Two possible approaches

In the present understanding, the establishment of meteorological radiation models are usually based on physical processes as well as on statistical techniques [Gueymard 1993, Kambezidis 1989, Muneer 1997, Psiloglou 2000, Psiloglou 2007].

The physical modelling studies the physical processes occurring in the atmosphere and influencing solar radiation. Accordingly, the solar radiation is absorbed, reflected, or diffused by solid particles in any location of space and especially by the earth, which depends on its arrival for many activities such as weather, climate, agriculture, ... The physical calculation method is exclusively based on physical considerations including the geometry of the earth, its distance from the sun, geographical location of any point on the earth, astronomical coordinates, the composition of the atmosphere, ... The incoming irradiation at any given point takes different shapes.
The second approach, "statistical solar climatology" branches into multiple aspects: modelling of the observed empirical frequency distributions, forecasting of solar radiation values at a given place based on historical data, looking for statistical interrelationships between the main solar irradiation components and other available meteorological parameters such as sunshine duration, cloudiness, temperature, and so on.
Our work has clearly taken the second approach.

HMM and SDE

Stochastic characteristics of solar radiation highly depend on some **unobserved** meteorological events such as frequency, height and type of the clouds and their optical

properties, atmospheric aerosols, ground albedo, water vapor, dust and atmospheric turbidity (Woyte et al. (2007)) while a sequence of solar radiation can be **observed** and measured at a given station. This has suggested us to model a random sequence of clearness index (resp. a stochastic process of solar radiation) by using a HMM which is a pair of correlated stochastic processes: the first (unobserved) one, called the *state process*, is a finite-state Markov chain in discrete-time (resp. in continous-time) representing meteorological regimes while the second (observed) one depends on the first one and describes the sequence of clearness index (resp. the process of solar radiation) as a discrete process (resp. a continuous one, solution of a SDE).
The idea of using HMM and SDE in the study of solar radiation sequences was mentioned by T. Soubdhan and R. Emilion in [Soubdhan 2009, Soubdhan 2011]. After a classification of daily solar radiation distributions, the authors thought that the sequence of class labels can be governed by a HMM in discrete time with some underlying unobservable regimes. The same authors have also proposed a SDE to model a continuous-time clearness index sequence but their data-driven approach fails for prediction during high variability regimes. However our work has been developed starting from these ideas. Our results can be summarized as follows:

1. We propose a discrete time HMM to model a daily (resp. hourly, resp. monthly) clearness index sequence.

2. We propose a continuous time HMM to model the clearness index process over a time interval $[0, T]$ in a solar day.

3. We propose a continuous time HMM and a SDE to directly model the total solar radiation process over a time interval $[0, T]$.

Estimation procedures

To fit our models to observed data, the estimation procedures will combine Maximum Likelihood Estimators (MLE) and Expectation Maximization (EM) algorithm for partially observed systems [Dempster 1977, Celeux 1989].

Filtering

A crucial notion in our estimation procedure is that of filter which is a time-indexed increasing family of σ-algebras, each one being generated by the events occured up to time t. The filtering process is defined as the family of conditional expectations w.r.t. these σ-algebras. A large part of our contribution deal with recursive equations of the filtering process needed in the estimation algorithms. They hinge on the work of [Dembo 1986, Campillo 1989, Elliott 1995, Elliott 2010].

Continuous-time filtering equations will be approximated by *robust* versions, following an approach due to [Clark 1977] and using some results of [James 1996, Krishnamurthy 2002, Clark 2005].

Reference probability method. Girsanov theorem

A great part of our computations concerns the so-called *reference probability method* which refers to a procedure where a probability measure change is introduced to reformulate the original estimation and control task into a new probability space (fictitious world) in which well-known results for identically and independently distributed (i.i.d.) random variables can be applied. Then the results are reinterpreted back to the original probability space (real world) by applying [Elliott 1995, chp. 1]. The Radon-Nykodim derivative of the new probabily measure w.r.t. the original one is given by the famous Girsanov theorem in both its discrete and continuous time version.

Thesis organization

Our thesis is divided into five Chapters following this introductory part.

Chapter 1.

In the first chapter we present some backgroung notions concerning solar radiation: direct, diffuse and global radiations, clearness index. The computation of the direct solar radiation is detailed. The end of the chapter briefly presents some points concerning measurement devices and datasets the we have dealt with.

Chapter 2.

In the second chapter, we recall some mathematical results that will be needed in chapters 3 and 4: conditional Bayes formula, Ito product, Ito formula, Girsanov theorem, HMM, EM algorithm.

Chapter 3.

In this third chapter we introduce three models for clearness index sequences (CISs):
- DTM-K, a Discrete-Time Model for discrete daily CISs, $(K_h)_{h=1,2,\ldots}$
- CTM-k, a Continuous-Time Model for continuous processes of CI, $(k_t), t \in [0, T]$, and its Discrete-Time Approximate Model, DTAM-k, obtained from time dicretization by uniformly partitionning $[0, T]$ into intervals of width Δ.

For each model, we define the state process, the observation process and the parameter vector. The state process of these models are finite-state homogeneous Markov chains. For CTM-k, the transition matrix of the chain is a rate matrix. For DTAM-k, the Δ width in the time partition is chosen to be small enough so that the transition matrix of the chain be a stochastic matrix. The observation process is a function of the chain which values are corrupted by a Gaussian noise (for DTM-K and DTAM-k) and by a standard Brownian motion (for CTM-k).

The filtering equations are established with complete proofs. Computations to obtain MLE updating formulas in the iterations of EM algorithm are detailed. Using DTAM-k, we first establish the computable approximation of the continuous time equations in CTM-k, and then we provide the estimates for the noisy variance.

Chapter 3 ends with some experiments with real data. Parameters of DTM-K are estimated from La Réunion island (France) data with daily CISs having similar characteristics while parameters of the CTM-k approximated by parameters of DTAM-k are estimated from Guadeloupe island (France) data which were sampled at 1Hz (i.e. at each second).

Chapter 4.

In this fourth chapter, we propose our main model, a continuous-time HMM for the total solar radiation sequence $(y_t)_{t\geq 0}$ under the random effects of meteorological events, denoted CTM-y.

The state process is similar to the CTM-k case but the observation process (y_t) modelling total solar radiation process, is assumed to be of the SDE:

$$dy_t = [g(X_t)I_t - y_t]dt + \sigma(X_t)y_t dW_t,$$

where I_t is the extraterrestrial radiation received at time t, (W_t) is a standard Brownian motion and $g(X_t)$, $\sigma(X_t)$ are functions of the Markov chain (X_t).

Again, the change-of-measure technique and the steps of EM algorithm establishing the filtering equations for updating the parameter vector g, are fully detailed.

Here too, we propose an approximation of state filter equation and we build a Discrete-Time Approximate Model (DTAM-y) to provide discrete-time approximate equations. Our computations hinge on a robust discretization of continuous-time filters recently obtained by [Elliott 2010, chap. 1]. Estimation of the noisy variance is studied.

Experimentations with real data and parameter estimations are performed from various samples of data sampled at 1Hz. Using the model with estimated parameters, we generate some simulations of solar radiation process paths.

Chapter 5.

In this fifth chapter, we first use DTM-K, with estimated parameters from La Réunion island data, to generate a large number of paths. A distribution of daily clearness index is then estimated from these simulated data.

Next, using the estimations for our two models CTM-k and CTM-y from 1Hz solar radiation (or clearness index) Guadeloupe island data, measured over time interval $[0, T]$, we simulate a large number of paths in the next hour $[T, T+1]$ and we propose a confidence interval for total solar radiation in $[T, T+1]$. Such predictions are compared to observations.

Given the data up to hour T and predicting total solar radiation during the next hour $[T, T+1]$ is of great interest for solar energy suppliers.

Chapter 6.

In this concluding part we discuss about some problems concerning parameter estimations, predictions, and comparison between the physical model approach and the statistical model approach. Some perspectives for future works are also proposed.

Contents

Chapter 1

Solar radiation

Contents

Résumé

Dans ce chapitre, nous rappelons d'abord quelques notions de physique en énergie solaire : rayonnement solaire extraterrestre, calcul du rayonnement direct, rayonnement diffus, rayonnement total ou global, indice de clarté. Nous parlerons brièvement des instruments de mesure du rayonnement et nous décrirons enfin les données réelles que nous avons utilisées.

Abstract

In this chapter, we first recall some physics notions in solar energy: extraterrestrial solar radiation, direct radiation computation, diffuse radiation, total or global radiation, clearness index. Then, we will briefly talk about radiation measurement instruments and last, we will describe the real data that we have dealt with.

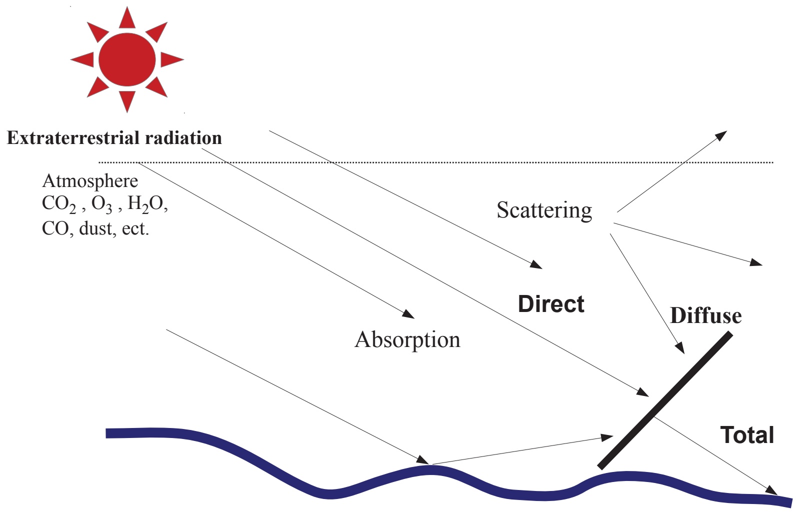


Figure 1.1: The types of solar radiation.

1.1 Introduction

Solar radiation emission from the sun into every corner of space appears in the form of electromagnetic waves that carry energy at the speed of light. The solar radiation is absorbed, reflected, or diffused by solid particles in any location of space and especially by the earth (Figure 1.1). This process depends on many environment conditions such as weather, climate, pollution, ... The incoming radiation at any given point takes different shapes depending on its geographical location, its astronomical coordinates, its distance from the sun, the composition of the local atmosphere and the local topgraphy.
This section provides some basic concepts, definitions, and astronomical equations which are used in our thesis. These concepts, definitions and equations are referenced from [Liu 1960, Psiloglou 2000, Sen 2008, Tovar-Pescador 2008].

1.2 Extraterrestrial solar radiation

1.2.1 Extraterrestrial normal radiation

The extraterrestrial normal radiation, denoted I_0, also called top of the atmosphere radiation, is the solar radiation arriving at the top of the atmosphere. It can simply be considered as the product of a solar constant denoted by I_{CS} and a correction factor of the earth's orbit, namely its excentricity, denoted by ε:

$$I_0 = I_{CS} \cdot \varepsilon. \tag{1.1}$$

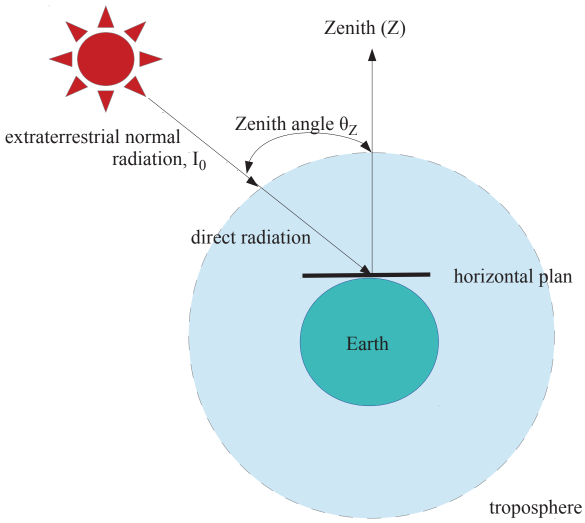


Figure 1.2: Horizontal plane for the extraterrestrial horizontal radiation.

The World Radiation Center has adopted that $I_{CS} = 1367\ W/m^2$ with an uncertainty of 1% [Duffie 2006] and introducing I_{CS} is justified as follows. As already said, the sun radiation is subject to many absorbing, diffusing, and reflecting effects within the earth's atmosphere which is about 10 km average thick and, therefore, it is necessary to know the power density, *i.e.*, watts per meter per minute on the earth's outer atmosphere and at right angles to the incident radiation. The density defined in this manner is referred to as the *solar constant* I_{CS}. It is equivalent to the energy from the sun, per unit time, received on a unit area of surface perpendicular to the direction of propagation of the radiation, at mean earth-sun distance, outside of the atmosphere.
The excentricity ε, as suggested by [Spencer 1972], is given by

$$\varepsilon = 1.00011 + 0.034221 \cos\Gamma + 0.00128 \sin\Gamma + 0.000719 \cos 2\Gamma + 0.000077 \sin 2\Gamma, \tag{1.2}$$

where the day angle Γ (in radians) is equal to:

$$\Gamma = \frac{2\pi n_d - 1}{365},$$

n_d denoting the number of the day in the year (1 for first of January, 365 for December 31).

A simple approximation for ε was suggested by [Duffie 1980, Duffie 1991]:

$$\varepsilon = 1 + 0.033 \cos\left(\frac{2\pi n_d}{365}\right). \tag{1.3}$$

1.2.2 Extraterrestrial horizontal radiation

At time t of a day, the amount incident radiation per horizontal surface area unit along the zenith direction, called the extraterrestrial horizontal radiation and denoted I_t, is related to the extraterrestrial normal radiation I_0 as follows:

$$I_t = I_0 \cos \theta_z, \tag{1.4}$$

where θ_z is the zenith angle at time t between the normal to the surface and the direction of the direct beam (Figure 1.2). The zenith angle calculation is described in Section 1.3 assuming, for sake of simplicity of modelling, that land is horizontal.

1.3 Zenith angle calculation

1.3.1 Equation of time

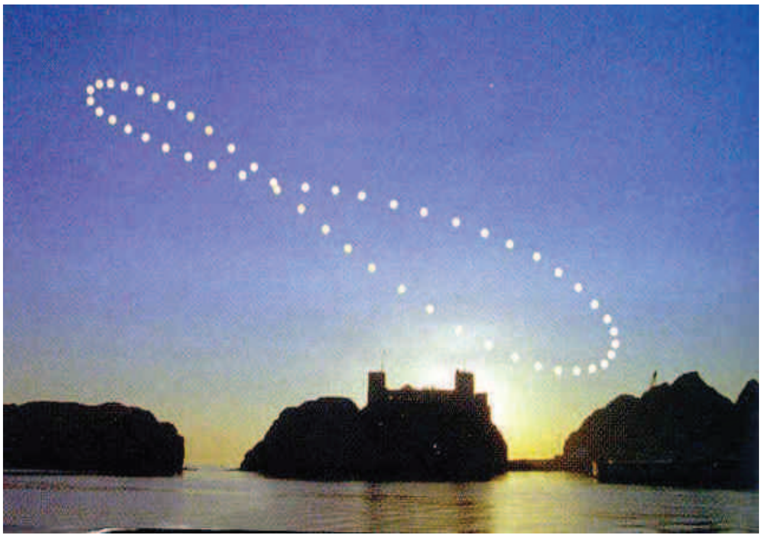


Figure 1.3: Relative positions of the Sun (photo: [Tingilinde 2006]).

The solar day is defined as the time that is needed by the Sun to achieve a complete tour of the Earth [Lanini 2010]. This does not necessarily correspond to 24 hours and varies from year to year. Figure 1.3 illustrates how the relative Sun position is moving across the sky: pictures of the Sun taken by an immobile photographer at the same time of the day have been superimposed. It can be seen that after one whole year of observations, the Sun is computing a eight-shape circuit. The principal causes of this phenomena are the elliptical shape of the terrestrial orbit around the Sun and the tilt of the Earth in relation to the plane of its orbit.

As a consequence, at 12 noon the Sun does not have the same position in the different months. The curve described by the so called the *equation of time*, (E_t), was proposed by [Spencer 1972] and then truncated by [Iqbal 1986]:

$$E_t = 229.18(0.000075+0.001868\cos\Gamma - 0.032077\sin\Gamma \\ -0.014615\cos 2\Gamma - 0.04089\sin 2\Gamma), \qquad (1.5)$$

where $\Gamma = \frac{2\pi(n_d-1)}{365}$, $n_d = 1, 2, \ldots, 365$.

1.3.2 Apparent solar time

Most meteorological measurements are recorded in terms of local standard time. In many solar energy calculations, it is necessary to obtain irradiation, wind, and temperature data for the same instant. It is, therefore, necessary to compute local apparent time, which is also called the true solar time. Solar time is the time to be used in all solar geometry calculations. It is necessary to apply the corrections due to the difference between the local longitude, L_{loc}, and the longitude of the standard time meridian, L_{stm}. The apparent time, L_{at}, can be calculated by considering the standard time, L_{st} according to [Iqbal 1986] as:

$$L_{at} = L_{st} \pm (L_{stm} - L_{loc}) + E_t. \qquad (1.6)$$

where E_t is caculated as in (1.5).

In this expression, "+" applies to west direction and "−" applies to east direction. All terms in the above equation are to be expressed in hours.

1.3.3 Hour angle

The hour angle, denoted ω, is the angular displacement of the sun east or west of the local meridian due to rotation of the earth on its axis at 15^0 per hour as morning negative and afternoon positive [Iqbal 1986]:

$$\omega = 15(12 - L_{at}), \qquad (1.7)$$

where L_{at} is the apparent time which is calculated as in (1.6).

1.3.4 Declination

The solar declination, denoted δ, is the angle between a line joining the centers of the Sun and the Earth to the equatorial plane, depends on the date and on the location (Figure 1.4): north direction has positive value, its maximum is equivalent to $+23.45^0$ at the summer solstice and its minimum to -23.45^0 at the winter solstice.

We can consider the following expressions for the approximate calculation of δ ([Iqbal 1986]) as:

$$\delta = 23.45\sin\left[\frac{360(284+n_d)}{365}\right], \qquad (1.8)$$

where n_d is the day number of year: $n_d = 1, 2, \ldots, 365$.

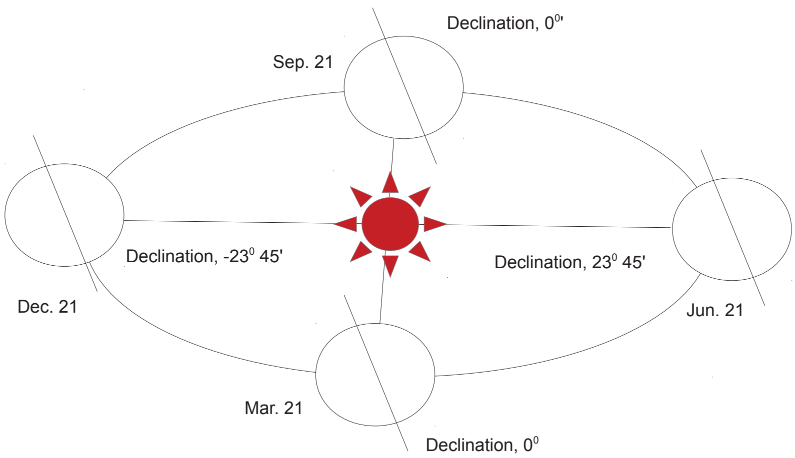


Figure 1.4: The declination angles.

1.3.5 Zenith angle

The zenith angle, denoted θ_z, is the angle between the vertical and the line to the sun *i.e.*, the angle of incidence of beam radiation on a horizontal surface (see again Figure 1.2). At solar noon zenith angle is zero, in the sunrise and sunset this angle is 90^0. The zenith angle θ_z can be actually calculated by [Sen 2008, page 86]:

$$\cos\theta_z = \cos\phi\cos\delta\cos\omega + \sin\phi\sin\delta, \tag{1.9}$$

where ϕ is the latitude of the location and δ calculated by (1.8), ω calculated by (1.7).

1.4 Total solar radiation

Total (global) solar radiation, G_t, is the sum of the direct beam, I_b, and the diffuse solar radiation, I_d, on a horizontal surface (Figure 1.5).

Solar radiation from the sun after traveling in space enters the atmosphere at the space-atmosphere interface, where the ionization layer of the atmosphere ends. Afterwards, a certain amount of solar radiation is absorbed by the atmosphere, by the clouds, and by particles in the atmosphere. A certain amount is reflected back into the space, and a certain amount is absorbed by the earth's surface. The combination of reflection, absorption (filtering), refraction, and scattering result in highly dynamic radiation levels at any given location on the earth. As a result of the cloud cover and scattering sunlight, the radiation received at any point is both direct (or beam) and diffuse (or scattered), see again Figure 1.1.

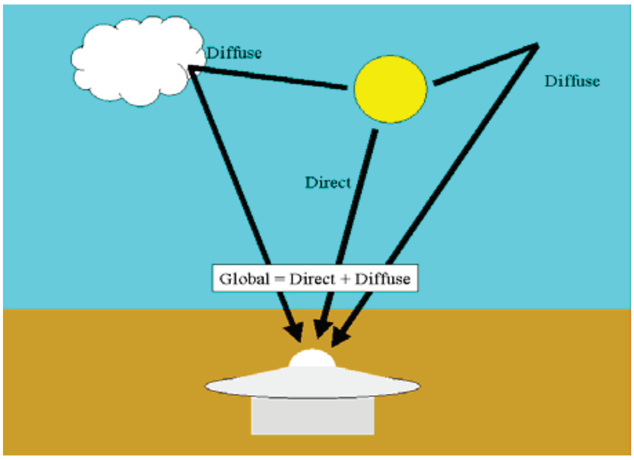


Figure 1.5: $G_t = I_b + I_d$.

1.4.1 Direct solar radiation

Direct solar radiation is defined as the radiation which travels in a straight line from the sun to the earth's surface. It is the solar radiation received from the sun without scatter by the atmosphere and without any disturbances. The quantity of direct solar radiation reaching any particular parts of the earth's surface is determined by the position of the point, time of year, shape of the surface, ... To model this would require knowledge of intensities and direction at different times of the day .... As examples, we can refer to the models in [Psiloglou 2000, Psiloglou 2007].

1.4.2 Diffuse solar radiation

After the solar radiation enters the earth's atmosphere, it is partially scattered and partially absorbed. The scattered radiation is called diffuse radiation. Again, a portion of this diffuse radiation goes back to space and a portion reaches the ground.

Diffuse radiation is first intercepted by the constituents of the air such as water vapor, CO_2, dust, aerosols, clouds, etc., and then it is released as scattered radiation in many directions. This is the main reason why diffuse radiation scattering in all directions and closed to the earth's surface as a source does not give rise to sharp shadows. When the solar radiation in the form of an electromagnetic wave hits a particle, a part of the incident energy is scattered in all directions and is called diffuse radiation. Diffuse radiation is scattered out of the solar beam by gases (Rayleigh scattering) and by aerosols (which include dust particles, as well as sulfate particles, soot, sea salty particles, pollen, etc.). The Reflected radiation is mainly reflected from the terrain and is therefore more important in mountainous areas.

Diffuse radiation occurs when small particles and gas molecules diffuse part of the incoming solar radiation in random directions without any alteration in the wavelength of the electromagnetic energy. Diffuse cloud radiation would require modeling of clouds and this is considered as quite impossible because of a great daily variability.

1.5 Clearness index

The ratio of the total solar radiation G_t to the extraterrestrial horizontal radiation I_t is defined as the clearness index and is denoted k_t:

$$k_t = \frac{G_t}{I_t}, \tag{1.10}$$

Clearness index is the quantity needed to focus on the analysis of fluctuations in solar radiance. It gives the ratio of the actual energy on the ground to that initially available at the top of the atmosphere accounting, therefore evaluating at time t the transparency of the atmosphere. Alternatively, this index can be considered as an instantaneous class membership degree, the class being an ideal perfect clear-sky day, the more this index is closed to one, the more the day is clear at time t.

For long-term predictions, the clearness index is often considered over a given time interval Δt. It is denoted by $K^{\Delta t}$ and is defined as the relation between the horizontal total radiation on the ground and the extraterrestrial horizontal radiation over the same time interval Δt:

$$K^{\Delta t} = \frac{\int_{\Delta t} G_s ds}{\int_{\Delta t} I_s ds}. \tag{1.11}$$

The usually used integration periods are the day and the hour, termed daily clearness index and hourly clearness index, respectively.

1.6 Solar radiation measurement

This section is designed to be a concise introduction for the instrumentation used to measure the components of solar radiation as well as for the climatic characteristics and geographical location of the areas where the observed data were recorded.

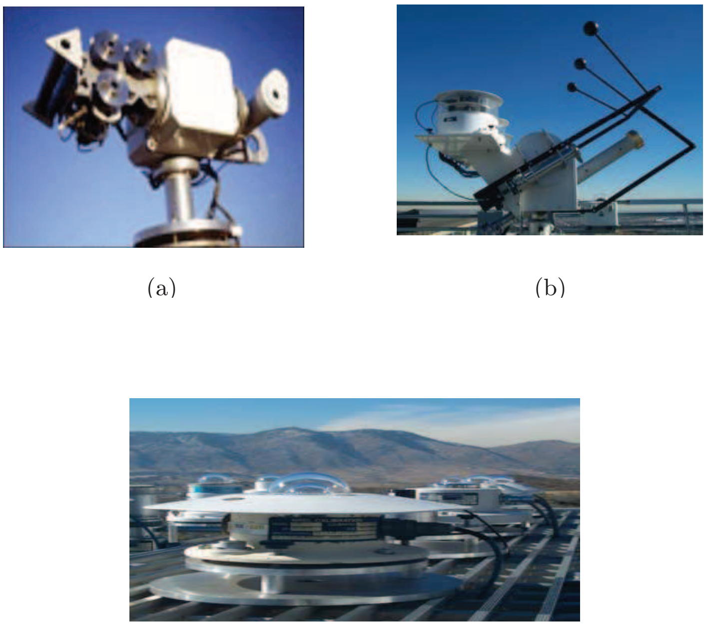
1.6.1 Solar radiometers

Pyrheliometer
The *Pyrheliometer* is a solar radiometer which is used to measure the "direct normal radiation" I_{bn} (note that $I_b = I_{bn} \cos\theta_z$). *Pyrheliometers* have a narrow aperture (generally between 5^0 and 6^0 total solid angle), admitting only beam radiation with some inadvertent circumsolar contribution from the Sun's aureole within the field

of view of the instrument, but still excluding all diffuse radiation from the rest of the sky. *Pyrheliometers* must be pointed at, and track the Sun throughout the day. Their sensor is always normal to the direct beam, so that I_{bn} is often called "direct normal radiation" (Figure 1.6a).

Pyranometer

A *Pyranometer* is used to perfomed the horizontal total radiation G_t or the diffuse radiarion I_d. Pyranometers have a 180^0 field of view. The horizontal total radiation G_t is measured by a *Pyranometer* with a horizontal sensor (Figure 1.6c) while the diffuse radiarion I_d is measured by a shaded *Pyranometer* under a tracking ball (Figure 1.6b).



(a) (b) (c)

Figure 1.6: Typical instruments for measuring solar radiation components:
(a) *Pyrheliometer*, (b) shaded *Pyranometer*, (c) *Pyranometer* with a horizontal sensor.

1.6.2 Data observed in Guadeloupe and La Réunion islands

The total solar radiation measurements used in our estimation procedures were performed in two French islands, namely Guadeloupe and La Réunion, located in the West Indies and the Indian Ocean, respectively. These areas are exposed to an important solar radiation and are characterized by a humid tropical climate.

Guadeloupe island is located at $16^0$15N latitude and $60^0$30W longitude. The average solar load for a horizontal surface is between 4 kWh/m^2 and 7 kWh/m^2 per day. The air temperature varies between 17^0C and 33^0C. Relative humidity ranges from 70% to 80% and the trade winds are relatively constant throughout the year. The total solar radiation measurements were performed in this island in 2006 by a *Pyranometer* from KIPP&ZONEN, model SP-Lite, a sensor having a response time inferior to 1 s. The SP-Lite measures the solar energy received from the entire hemisphere (180^0 field of view). These data were recorded, pretreated, analyzed ans interpreted by Dr. T. Soubdhan, Assistant Professor in Physics at University of Antilles-Guyane, Guadeloupe.

La Réunion is a Southern Hemisphere volcanic island with an average temperature oscillating from 24^0C to 32^0C in the coastal regions and oscillating from 15^0C to 22^0C in the regions located above an altitude of 1500 m in the interior of the island. The combination of a very steep terrain, with large variations in altitude, and prevailing trade winds from south-southeast induce local contrasts in weather patterns at ground level. The radiations were captured by a SPN1 *Pyranometer* ([Delta-T-Devices 2012]), a sensor rated as a "good quality" one by World Meteorological Organization. This sensor is actually based on a set of seven thermopiles, symmetrically arranged below a shadow dome according to a specific geometry, ensuring by that way that, at any time of the day, wherever in the world the measurement is made, there is always one sensor fully exposed to the sun and one sensor fully shadowed. The recorded data used for our daily clearness index sequences modelling were sampled at 0.1 s at Moufia campus location ($20^0$54S, $55^0$29E) from 2009 to 2011 in the setting of a La Réunion region project titled RCI-GS. They were managed by software engineer M. Delsaut and were averaged to give one collected point per minute for final storage purpose.

Both data collections and storages used a datalogger from CAMMPBELL SCIENTIFIC (the "burning" sunshine recorders were first developed by John Francis Campbell in 1853 and later modified in 1879 by Sir George Gabriel Stokes, ...).

Chapter 2

Mathematical recalls

Contents

Résumé

Nous supposons connu les ésprances conditionnelles, les chaînes de Markov à temps discret et à temps continu, les martingales et le calcul stochastique.
Nous précisons dans ce chapitre formalisme, notations et méthodes utilisés dans les prochains chapitres en suivant la présentation de [Elliott 2010] : modèles Markoviens cachés (HMM) à temps discret et méthode de changement de probabilité dite *méthode de la probabilité référente* basée sur le théoreme de Girsanov en version discrète, rappels de certains résultats de calcul stochastique, HMM à temps continu et méthode de la probabilité référente, algorithme EM en temps continu.

Abstract

Conditional expectations, discrete-time and continuous time Markov chains, martingales and stochastic calculus are assumed to be known.
We precise in this chapter the formalism, the notations and the methods used in the next chapters, following the presentation in [Elliott 2010] : Hidden Markovian Models (HMM) in discrete time and reference probability method based on a discrete version of Girsanov theorem, recalls of some results in stochastic calculus, continuous time HMM and reference probability method, EM algorithm in continuous time.

2.1 Conditional expectations

2.1.1 Radon-Nikodym derivative

Theorem 2.1. *(Radon-Nikodym) If P and $\overline{P}$ are two probability measures on $(\Omega, \mathcal{B})$ such that for each $B \in \mathcal{B}$, $P(B) = 0$ implies $\overline{P}(B) = 0$, then there are exists a non-negative random variable Λ, such that $\overline{P}(C) = \int_C \Lambda dP$ for all $C \in \mathcal{B}$. We write $d\overline{P}/dP|_{\mathcal{B}} = \Lambda$.*

For a proof, see [Wong 1985].

Definition 2.1. *Let $X \in L_1$ and $\mathcal{A}$ be a sub-σ-field of $\mathcal{B}$. If X is non-negative and integrable we can use the Radon-Nikodym to deduce the existence of an $\mathcal{A}$-measurable random variable, denoted by $E(X|\mathcal{A})$, which is uniquely determined except on an even of probability zero, such that*

$$\int_A X dP = \int_A E(X|\mathcal{A}) dP, \tag{2.1}$$

for all $A \in \mathcal{A}$.

$E(X|\mathcal{A})$ is called the conditional expectation of X give $\mathcal{A}$. For a general integrable random variable we define $E(X|\mathcal{A})$ as $E(X^+|\mathcal{A}) - E(X^-|\mathcal{A})$.

The following is a list of classical results. If $\mathcal{A}_1$ and $\mathcal{A}_2$ are two sub-σ-fields of $\mathcal{B}$ such that $\mathcal{A}_1 \subset \mathcal{A}_2$, then

$$E(E(X|\mathcal{A}_1)|\mathcal{A}_2) = E(E(X|\mathcal{A}_2)|\mathcal{A}_1) = E(X|\mathcal{A}_1). \tag{2.2}$$

If $X, Y, XY \in L_1$, and Y is $\mathcal{A}$-measurable, then

$$E(XY|\mathcal{A}) = YE(X|\mathcal{A}). \tag{2.3}$$

If X and Y are independent, then

$$E(Y|\sigma(X)) = E(Y). \tag{2.4}$$

2.1.2 Jensen inequality

Theorem 2.2. *Let $\{\Omega, \mathcal{F}, P\}$ be a probability space and $\mathcal{G}$ a subfield of $\mathcal{F}$. Let $\phi : \mathbf{R} \to \mathbf{R}$ be convex and let X be an integrable random variable such that $\phi(X)$ is integrable. Then, we have*

$$\phi(E(X|\mathcal{G})) \leq E(\phi(X)|\mathcal{G}).$$

A proof can be found in [Elliott 1982].

2.1.3 Conditional Bayes formula

Theorem 2.3. *Suppose $(\Omega, \mathcal{F}, P)$ is a probability space and $\mathcal{G} \subset \mathcal{F}$ is a sub-σ-field. Suppose $\overline{P}$ is another probability measure absolutely continuous with respect to P and with Radon-Nikodym derivative $d\overline{P}/dP = \Lambda$. Then if ϕ is any $\overline{P}$ integrable random variable*

$$\overline{E}[\phi|\mathcal{G}] = \psi \qquad \text{where} \quad \psi = \frac{E[\Lambda\phi|\mathcal{G}]}{E[\Lambda|\mathcal{G}]} \quad if\ E[\Lambda|\mathcal{G}] > 0$$
$$\text{and} \quad \psi = 0 \quad \text{otherwise.}$$

Theorem 2.3 was proved by [Elliott 2010, Theorem 3.2].

2.2 Martingale difference sequence

Definition 2.2. *Let $(X_h) = \{X_h, h = 1, 2, \dots\}$ be an adapted discrete stochastic process on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_h), P)$. (X_h) is called a martingale difference sequence (MDS) if it satisfies the following two conditions:*

i. $E(|X_h|) < \infty$,

ii. $E(X_h|\mathcal{F}_{h-1}) = 0$, *a.s., for all* h.

So, a stochastic serie is an MDS if its expectation with respect to the past is zero. MDS is an extremely useful concept in modern probability theory because it implies much milder restrictions on the memory of the sequence than independence. Most of limit theorems that hold for an independent sequence also hold for an MDS.

The definition implies that:

- if (Y_h) is a martingale, then $(U_h) = (Y_h - Y_{h-1})$ will be an MDS,
- $(V_h) = (X_h - E[X_h|\mathcal{F}_{h-1}])$ is an MDS.

The sequences (U_h), (V_h) above are also called *sequences of martingale increments*.

2.3 Binary vector representation of a Markov chain

Consider a (discrete or continuous time) Markov chain $(\mathbb{X}_{time})$ with finite state space $\mathbf{S}_{\mathbb{X}} = \{s_1, s_2, ..., s_N\}$.

Let $\mathbf{1}_{(\mathbb{X}_{time}=s_i)}$ denote the so-called indicator function defined as $\mathbf{1}_{(\mathbb{X}_{time}=s_i)} = 1$ if $\mathbb{X}_{time} = s_i$ and $\mathbf{1}_{(\mathbb{X}_{time}=s_i)} = 0$ if $\mathbb{X}_{time} \neq s_i$. The prime symbol denoting transpose, we consider the following binary vector representation:

$$X_{time} = (\mathbf{1}_{(\mathbb{X}_{time}=s_1)}, \mathbf{1}_{(\mathbb{X}_{time}=s_2)}, ..., \mathbf{1}_{(\mathbb{X}_{time}=s_N)})',$$

so that at any *time*, just one component of X_{time} is one while the others are zero.

Then, for sake of simplicity in computations, we will now consider the chain (X_{time}) which is derived from $(\mathbb{X}_{time})$, the state space of (X_{time}) being the set of unit vectors e_i with all components 0 but 1 at the i-th component:

$$\mathbf{S} = \{e_1, e_2, ..., e_N\}.$$

Let $\langle \cdot\,, \cdot \rangle$ denote the usual inner product in $\mathbf{R}^N$. Noticing that $\langle X_{time}, e_i \rangle = \mathbf{1}_{(\mathbb{X}_{time}=s_i)}$, $i = 1, 2, \ldots, N$, we will write the vector representation of the Markov chain as

$$X_{time} = (\langle X_{time}, e_1 \rangle, \langle X_{time}, e_2 \rangle, ..., \langle X_{time}, e_N \rangle)'.$$

Throughout this thesis, we will assume without loss of generality, that the state space of the finite-state Markov chain (X_{time}) is a set of unit vectors defined as above and that X_0 is given or its distribution π_0 is known.

2.4 Hidden Markov models

A Hidden Markov Model (HMM) is a pair of stochastic processes called the *state process* and the *observation process*, respectively. The state process is a hidden, that is an unobserved, homogeneous Markov chain modelling the environment, each state of the chain representing a specific regime of the environment. The observation process is a real valued function of the chain corrupted by a Gaussian noise (in discrete time) or is assumed to satisfy a stochastic differential equation (in continuous time). Such processes will be defined on a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_{time}), P)$.

2.5 Discrete-time HMM

In this section we present a discrete time HMM that will be used to model daily clearness index sequences.

Consider a system whose states are described by a discrete-time homogeneous Markov chain $(X_h)_{h=0,1,2,...}$, called *the state process*, with state space $\mathbf{S} = \{e_1, e_2, ..., e_N\}$.

For $h = 1, 2, \ldots$, we will write

$$X_h = (\langle X_h, e_1\rangle, \langle X_h, e_2\rangle, \ldots, \langle X_h, e_N\rangle)'.$$

Recall that $\langle X_h, e_i\rangle = \mathbf{1}_{(X_h=e_i)}$, $i = 1, 2, \ldots, N$.

Remark 2.1. *For $i = 1, 2, \ldots, N$, we have*

$$E(\langle X_h, e_i\rangle) = \sum_{j=1}^{N} \langle e_j, e_i\rangle P(X_h = e_j) = P(X_h = e_i). \tag{2.5}$$

Let $\mathcal{X}_h$ denote the σ-algebra generated by $\{X_0, X_1, \ldots, X_h\}$ and let $A = (a_{ji}) \in \mathbf{R}^{N\times N}$ denote the probability transition matrix of $(X_h)_{h=1,2,...}$ defined as

$$a_{ji} = P(X_h = e_j | X_{h-1} = e_i), \quad i, j = 1, 2, \ldots, N.$$

Note that $a_{ii} = 1 - \sum_{j\neq i} a_{ji}$ $(i = 1, 2, \ldots, N)$.

The equation of the process $(X_h)_{h=1,2,...}$, the so-called *state equation*, will be obtained from the following lemma.

Lemma 2.1. *Let $(V_h)_{h=1,2,...}$ be a sequence defined by*

$$V_h \triangleq X_h - AX_{h-1}. \tag{2.6}$$

Then $(V_h)_{h=1,2,...}$ is a sequence of martingale increments.

Recall that a sequence of martingale increments is a random discrete series whose expectation with respect to the past is 0 (see Section 2.2).

Proof. The Markov property implies that

$$P(X_h = e_j | \mathcal{X}_{h-1}) = P(X_h = e_j | X_{h-1}). \tag{2.7}$$

From (2.7) and Remark 2.1, we have

$$E(X_h | \mathcal{X}_{h-1}) = E(X_h | X_{h-1}) = AX_{h-1}. \tag{2.8}$$

Thus,

$$E(V_h | \mathcal{X}_{h-1}) = E(X_h - AX_{h-1} | X_{h-1}) = AX_{h-1} - AX_{h-1} = 0. \tag{2.9}$$

□

From Lemma 2.1, the state process $(X_h)_{h=1,2,\dots}$ can be represented by the *state equation*:

$$X_h = AX_{h-1} + V_h, \tag{2.10}$$

where $(V_h)_{h=1,2,\dots}$ is a sequence of martingale increments.

We suppose that the state process (X_h) is not observed directly, but rather observed through a function of the Markov chain (X_h), say $(K_h)_{h=1,2,\dots}$.

Definition 2.3. *A discrete-time HMM is a pair of processes $(X_h, K_h)_{h=1,2,\dots}$ determined by the following equations:*

$$X_h = AX_{h-1} + V_h, \tag{2.11}$$
$$K_h = b(X_h) + \alpha(X_h)w_h, \tag{2.12}$$

where (V_h) is sequence of martingales, (w_h) is a sequence of i.i.d. $\mathcal{N}(0,1)$ random variables, w_h is independent of $\sigma\{X_1, X_2, \dots, X_h, K_1, K_2, \dots, K_h\}$ and $b(X_h) = \langle X_h, b\rangle$, $\alpha(X_h) = \langle X_h, \alpha\rangle$, with $b = (b_1, b_2, \dots, b_N)'$, $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_N)'$. The parameter vector of the model is defined as the vector:

$$\theta = (a_{ji}, 1 \le j \ne i \le N; b_1, b_2, \dots, b_N; \alpha_1, \alpha_2, \dots, \alpha_N),$$

We shall assume $\alpha_i \neq 0$ and thus without loss of genrality that $\alpha_i > 0, 1 \le i \le N$.

We will assume that θ belongs to a compact set $\Theta \subset \mathbf{R}^{N\times N+N}$ and will denote by P_θ a probability measure on $(\Omega, \mathcal{F})$ for which the process $(X_h, K_h)_{h=1,2,\dots}$ satisfies equations (2.12) with parameter θ.

In practice the number of states N will be suggested by the user.

2.5.1 Filtrations, number of jumps, occupation time and level sums

The following notions will be useful for estimating the model parameters.

Definition 2.4. *For $h = 1, 2, \dots,$ let*

$$\mathcal{G}_h^K \triangleq \sigma\{X_1, X_2, \dots, X_h, K_1, K_2, \dots, K_h\},$$
$$\mathcal{Y}_h^K \triangleq \sigma\{K_1, K_2, \dots, K_h\}$$

be the $\sigma-$algebras generated by (X_h, K_h) and (K_h), respectively.

These $\sigma-$algebras containing all the available information up to time h form increasing sequences, and thus are *filtrations*. $\mathcal{G}_h^K$ is called the filtration of *complete data* and $\mathcal{Y}_h^K$ the filtration of *incomplete observation data*.

Note that $\mathcal{Y}_h^K \subset \mathcal{G}_h^K \subset \mathcal{F}_h$ for all h.

Estimating the model parameters requires the computation of the conditional expectations of the following quantities given the observation history $\mathcal{Y}_h^K$.

Definition 2.5. *Let $(X_h, K_h)_{h=1,2,\dots}$ be a discrete-time HMM and let $f(\cdot)$ be any bounded function. Let us define:*

1. *the number of jumps of the Markov chain from e_i to e_j until time h by*

$$\mathcal{J}_h^{ij} = \sum_{l=1}^{h} \langle X_{l-1}, e_i \rangle \langle X_l, e_j \rangle, \tag{2.13}$$

2. *the occupation time of the Markov chain in state e_i until time h by*

$$\mathcal{O}_h^{i} = \sum_{l=1}^{h} \langle X_l, e_i \rangle, \tag{2.14}$$

3. *and the level sums of the observation process in state e_i up to time h by*

$$\mathcal{T}_h^{i}(f) = \sum_{l=1}^{h} f(K_l) \langle X_l, e_i \rangle. \tag{2.15}$$

2.5.2 Reference Probability Method of measure change

In the so-called Reference Probability Method [Elliott 1995, Elliott 2010], the main technique for obtaining ML (Maximum Likelihood) estimates of parameters is the change of measure which is a discrete time version of Girsanov theorem (see Theorem 2.5). To achieve such a mathematical objective, we are going to work in a "fictitious world" $(\Omega, (\mathcal{F}_h), \overline{P})$, called the reference probability space, with a reference probability measure $\overline{P}$ which is determined by a fix parameter vector $\theta_0 \in \Theta$ and then we will relate these results in the "real world" $(\Omega, (\mathcal{F}_h), P_\theta)$ by using a back change of measure and by applying Bayes's rule.

The reference probability measure $\overline{P}$, a convenient measure to work with, is defined such that under $\overline{P}$ the observations are i.i.d. random variables. Working under $\overline{P}$, we will reformulate the initial estimate of the parameter vector θ in a "fictitious world" $(\Omega, (\mathcal{F}_{time}), \overline{P})$ where the well-known results for i.i.d. random variables can be applied. Then the results will be reinterpreted back to the real world $(\Omega, (\mathcal{F}_h), P_\theta)$ with the initial probability measure P_θ (see [Elliott 2010, chap. 1, pg. 3-11] for more details).

Introduce a reference probability measure $\overline{P}$ from P_θ such that under $\overline{P}$:

($C3.1$) $(K_h)_{h=1,2,\ldots}$ is a sequence of $\mathcal{N}(0,1)$ i.i.d. random variables which are independent of the Markov chain (X_h),

($C3.2$) $(X_h)_{h=1,2,\ldots}$ is a Markov chain with transition matrix A so that $\overline{E}(X_h|\mathcal{G}_{h-1}^K) = AX_{h-1}$, where $\overline{E}$ is the expectation under $\overline{P}$.

The existence of $\overline{P}$ with the characteristics above is described by Lemma 2.2 below.

Lemma 2.2. *Let P_θ be the measure determined by the parameter vector θ (Definition 2.3) and let $\overline{P}$ be a new probability measure such that*

$$\left.\frac{\overline{P}}{P_\theta}\right|_{\mathcal{G}_h^K} = \overline{\Lambda}_h^{K,\theta} = \prod_{l=1}^{h} \overline{\lambda}_l, \tag{2.16}$$

where $\overline{\lambda}_l = \frac{\langle X_l,\alpha\rangle\phi(K_l)}{\phi\left(\frac{K_l-\langle X_l,b\rangle}{\langle X_l,\alpha\rangle}\right)}, l = 1,2,\dots,h$, $\phi(\cdot)$ *is the* $\mathcal{N}(0,1)$ *density function.*

Then $\overline{P}$ *satisfy conditions* $(C3.1)$, $(C3.2)$ *stated above.*

Proof. Our proof starts by observing that the existence of $\overline{P}$ such that $(\overline{P}/P_\theta)|_{\mathcal{G}^K_{h-1}} = \overline{\Lambda}^{K,\theta}_h$ follows from Kolmogorov's Extension Theorem.

We first prove that $\Lambda^{K,\theta}_h$ is integrable and a martingale under P_θ by (a) and (b) below.

(a) For $l = 2,3,\dots,h$, we have $K_l = \langle X_l, b\rangle + \langle X_t,\alpha\rangle w_l$, so that $w_l = \frac{K_l-\langle X_l,b\rangle}{\langle X_t,\alpha\rangle}$ with $w_l \sim \mathcal{N}(0,1))$. Moreover as $\mathcal{G}^K_{l-1} \subset \mathcal{G}^K_{l-1} \vee X_l$, it follows that

$$\begin{aligned}
E_\theta\left(\overline{\lambda}_l|\mathcal{G}^K_{l-1}\right) =& E_\theta\left\{E_\theta\left(\frac{\langle X_l,\alpha\rangle\phi\left(K_l\right)}{\phi\left(\frac{K_l-\langle X_l,b\rangle}{\langle X_l,\alpha\rangle}\right)}\middle|\mathcal{G}^K_{l-1}\vee X_l\right)\middle|\mathcal{G}^K_{l-1}\right\}\\
=& E_\theta\left\{\int_{-\infty}^{+\infty}\frac{\langle X_l,\alpha\rangle\phi\left(K_l\right)}{\phi(w_l)}\phi(w_l)dw_l\middle|\mathcal{G}^K_{l-1}\right\}\\
=& \int_{-\infty}^{+\infty}\phi(K_l)dK_l\\
=& 1,
\end{aligned} \tag{2.17}$$

because w_l is independent of $\mathcal{G}^K_{h-1}$.

From this,

$$E_\theta(\overline{\Lambda}^{K,\theta}_h) = E_\theta\left(\prod_{l=1}^{h}\overline{\lambda}_l\right) = 1 < \infty.$$

(b) By (2.17),

$$E_\theta(\overline{\Lambda}^{K,\theta}_h|\mathcal{G}^K_{h-1}) = \overline{\Lambda}^{K,\theta}_{h-1}E_\theta\left(\overline{\lambda}_h|\mathcal{G}^K_{h-1}\right) = \overline{\Lambda}^{K,\theta}_{h-1}.$$

We now note that $\overline{P}(K_h \le x|\mathcal{G}^K_{h-1}) = \overline{E}(\mathbf{1}_{(K_h\le x)}|\mathcal{G}^K_{h-1})$. By a version of Bayes theorem,

$$\begin{aligned}
\overline{E}(\mathbf{1}_{(K_h\le x)}|\mathcal{G}^K_{h-1}) =& \frac{E_\theta(\overline{\Lambda}^{K,\theta}_h\mathbf{1}_{(K_h\le x)}|\mathcal{G}^K_{h-1})}{E_\theta(\overline{\Lambda}^{K,\theta}_h|\mathcal{G}^K_{h-1})}\\
=& \frac{\overline{\Lambda}^{K,\theta}_{h-1}}{\overline{\Lambda}^{K,\theta}_{h-1}}\frac{E_\theta\left(\overline{\lambda}_h\mathbf{1}_{(K_h\le x)}|\mathcal{G}^K_{h-1}\right)}{E_\theta\left(\overline{\lambda}_h|\mathcal{G}^K_{h-1}\right)}.
\end{aligned} \tag{2.18}$$

From (2.18) and (2.17), we have

$$\begin{aligned}
\overline{P}(K_h \le x|\mathcal{G}^K_{h-1}) =& E_\theta\left(\frac{\langle X_h,\alpha\rangle\phi\left(K_h\right)}{\phi\left(\frac{K_h-\langle X_h,b\rangle}{\langle X_h,\alpha\rangle}\right)}\mathbf{1}_{(K_h\le x)}\middle|\mathcal{G}^K_{h-1}\right)\\
=& \int_{-\infty}^{+\infty}\frac{\langle X_h,\alpha\rangle\phi\left(K_h\right)}{\phi(w_h)}\phi(w_h)\mathbf{1}_{(K_h\le x)}dw_h\\
=& \int_{-\infty}^{x}\phi(K_h)dK_h.
\end{aligned}$$

Then $\overline{P}(K_h \le x|\mathcal{G}^K_{h-1}) = \overline{P}(K_h \le x))$, $K_h \sim \mathcal{N}(0,1)$ under $\overline{P}$.

Let $\mathcal{X}_h = \sigma\{X_0, X_1, \dots, X_h\}$, by considering similarly,

$$\begin{aligned}\overline{P}(K_h \le x|\mathcal{X}_h) =& E_\theta\left\{ E_\theta\left(\frac{\langle X_h,\alpha\rangle \phi\left(K_h\right)}{\phi\left(\frac{K_h-\langle X_h,b\rangle}{\langle X_h,\alpha\rangle}\right)} \mathbf{1}_{(K_h\le x)} \middle| \mathcal{G}^K_h \right) \middle| \mathcal{X}_h \right\} \\ =& E_\theta\left\{ \int_{-\infty}^{+\infty} \frac{\langle X_h,\alpha\rangle\phi\left(K_h\right)}{\phi(w_h)}\phi(w_h)\mathbf{1}_{(K_h\le x)}dw_l \middle| \mathcal{X}_h\right\} \\ =& \int_{-\infty}^{+\infty} \frac{\langle X_h,\alpha\rangle\phi\left(K_h\right)}{\phi(w_h)}\phi(w_h)\mathbf{1}_{(K_h\le x)}dw_h \\ =& \int_{-\infty}^{x} \phi(K_h)dK_h,\end{aligned}$$

a quantity which is independent of $\mathcal{X}_h$. Condition $(C3.1)$ follows.

Consider now condition $(C3.2)$. We have

$$\begin{aligned}\overline{E}(X_h|\mathcal{G}^Y_{h-1}) =& \frac{E_\theta(\Lambda^{K,\theta}_h X_h|\mathcal{G}^Y_{h-1})}{E_\theta(\Lambda^{K,\theta}_h|\mathcal{G}^Y_{h-1})} \\ =& E_\theta\left(\frac{\langle X_h,\alpha\rangle\phi\left(K_h\right)}{\phi\left(\frac{K_h-\langle X_h,b\rangle}{\langle X_h,\alpha\rangle}\right)} X_h \middle| \mathcal{G}^K_{h-1}\right) \\ =& E_\theta(X_h|\mathcal{G}^K_{h-1}) \\ =& AX_{h-1}.\end{aligned}$$

So, under $\overline{P}$, (X_h) remains a Markov chain with transition matrix A. This completes our proof. □

Remark 2.2. *Starting with the probability measure $\overline{P}$ we can recover the measure P_θ by observing that*

$$\left.\frac{dP_\theta}{d\overline{P}}\right|_{\mathcal{G}^K_h} = \Lambda^{K,\theta}_h = \prod_{l=1}^{h} \frac{\phi\left(\frac{K_l-\langle X_l,b\rangle}{\langle X_l,\alpha\rangle}\right)}{\langle X_l,\alpha\rangle\phi\left(K_l\right)}, \tag{2.19}$$

Note that $\Lambda^{K,\theta}_h \overline{\Lambda}^{K,\theta}_h = 1$.

2.5.3 Normalized and unnormalized filters

Definition 2.6. *For any $\mathcal{G}^K_h$-adapted sequence (H_h) (for instance $H_h \equiv \mathcal{J}^{ij}_h, \mathcal{O}^i_h$ or $\mathcal{T}^i_h(f)$), define*

$$\gamma(H_h) \triangleq \overline{E}(\Lambda^{K,\theta}_h H_h|\mathcal{Y}^K_h), \tag{2.20}$$

$$\pi(H_h) \triangleq E_\theta(H_h|\mathcal{Y}^K_h), \tag{2.21}$$

where $\Lambda^{K,\theta}_h$ is determined as in (2.19).

The processes $\gamma(H_h)$ and $\pi(H_h)$ are called the unnormalized filter and the normalized filter of the process (H_h), respectively.

In the problem of parameter estimation, we have to compute the normalized filters $\pi(H_h)$, for $H_h \equiv \mathcal{J}_h^{ij}, \mathcal{O}_h^i$ or $\mathcal{T}_h^i(f)$. We first work on the reference probability space $(\Omega, (\mathcal{F}_h), \overline{P})$ to establish the equation of unnormalized filters $\gamma(H_h)$. These results will be used to obtain the normalized filters $\pi(H_h)$ in the real world $(\Omega, (\mathcal{F}_h), P_\theta)$. By using Bayes rule:

$$E_\theta(H_h|\mathcal{Y}_h^K) = \frac{\overline{E}(\Lambda_h^{K,\theta} H_h|\mathcal{Y}_h^K)}{\overline{E}(\Lambda_h^{K,\theta}|\mathcal{Y}_h^K)}, \tag{2.22}$$

which is equivalent to

$$\pi(H_h) = \frac{\gamma(H_h)}{\gamma(1)}. \tag{2.23}$$

On the other hand, it is not possible to compute directly the filter $\gamma(H_h)$, for $H_h \equiv \mathcal{J}_h^{ij}, \mathcal{O}_h^i$ or $\mathcal{T}_h^i(f)$, but we can compute its associated process $\gamma(H_h X_h) = \overline{E}(\Lambda_h^{K,\theta} H_h X_h|\mathcal{Y}_h^K)$ and note that

$$\begin{aligned}\gamma(H_h) =& \gamma(H_h\langle X_h, \underline{1}\rangle), \quad \text{because } 1 = \sum_{i=1}^N \langle X_h, e_i\rangle = \langle X_h, \underline{1}\rangle, \\ =& \langle \gamma(H_h X_h), \underline{1}\rangle.\end{aligned}$$

Consequently, we can rewrite (2.23) as

$$\pi(H_h) = \frac{\langle \gamma(H_h X_h), \underline{1}\rangle}{\langle \gamma(X_h), \underline{1}\rangle}. \tag{2.24}$$

Thus, for obtaining the normalized filters $\pi(H_h)$ ($H_h \equiv \mathcal{J}_h^{ij}, \mathcal{O}_h^i$ or $\mathcal{T}_h^i(f)$) in the real world $(\Omega, (\mathcal{F}_h), P_\theta)$, we must determine the equation of unnormalized filter processes $\gamma(H_h X_h)$. This will be computed in Chapter 3.

2.6 Some recalls on stochastic calculus

2.6.1 Ito product rule

For two semimartingales X_t and Y_t, Ito product rule gives

$$X_t Y_t = \int_0^t X_s dY_s + \int_0^t Y_s dX_s + [X, Y]_t, \tag{2.25}$$

where

$$\begin{aligned}[X, Y]_t = \lim_{n\to\infty} (\text{in prob.}) \Big\{ X_0 Y_0 + \sum_{0\le k<2^n} \big[&(X_{t(k+1)2^{-n}} - X_{tk2^{-n}}) \\ &\times (Y_{t(k+1)2^{-n}} - Y_{tk2^{-n}})\big] \Big\}\end{aligned}$$

is the *quadratic variation* of X_t and Y_t.

Remark 2.3. *If the process $[X,Y]_t$ has a compensator denoted by $\langle X,Y\rangle_t$, it will be called the predictabe quadratic variation of X_t and Y_t. If the martingale part of the semimartingale is discontinuous, then $[X,Y]_t = X_0Y_0 + \sum_{0<s\leq t}\Delta X_s \Delta Y_s$.*

Remark 2.4. *If X_t and Y_t are Ito processes, then:*

$$X_t = X_0 + \int_0^t \mu_s^X ds + \int_0^t \sigma_s^X dW_s,$$
$$Y_t = X_0 + \int_0^t \mu_s^Y ds + \int_0^t \sigma_s^Y dW_s,$$

where W_t is a standard Brownian motion, and

$$[X,Y]_t = X_0H_0 + \int_0^t \sigma_s^X \sigma_s^Y ds. \tag{2.26}$$

2.6.2 Ito formula

Let $(\xi_t)_{t\in[0,T]}$ be a random process satisfying the stochastic differential equation (SDE)

$$d\xi = a_t dt + b_t dW_t, \tag{2.27}$$

where (W_t) is a Wiener process, and the nonanticipative functions a_t, b_t are such that

$$P\left\{\int_0^T |a_t| dt < \infty\right\} = 1, \tag{2.28}$$
$$P\left\{\int_0^T b_t^2 dt < \infty\right\} = 1. \tag{2.29}$$

Theorem 2.4. *Let $f(t,x) : [0,+\infty) \times \mathbf{R} \to \mathbf{R}$ be a continuous function satisfying $f_t \in C^1([0,+\infty))$ and $f_x \in C^2(\mathbf{R})$. Then the process $f(t,\xi_t)$ satisfies the SDE*

$$df(t,\xi_t) = \left[f_t'(t,\xi_t) + f_x'(t,\xi_t)a_t + \frac{1}{2}f_{xx}''(t,\xi_t)b_t^2\right]dt + f_x'(t,\xi_t)b_t dW_t. \tag{2.30}$$

Theorem 2.4 which is Theorem 4.4 of [Lipster 2010]. The formula given by (2.30) was obtained by K. Itô and is called the *Itô formula*.

2.6.3 Girsanov theorem

Theorem 2.5. *(Girsanov, [Elliott 2010, page 355]) Suppose that (Y_t), $t \in [0,T]$, is a standard Brownian motion on a filtered space $\{\Omega, \mathcal{F}, (\mathcal{F}_t), Q\}$. Let $f : \Omega \times [0,T] \to \mathbf{R}$ be a predictable process such that* [1]

$$\int_0^T |f_t|^2 dt < \infty \; a.s.$$
$$E\left[\exp\left(\frac{1}{2}\int_0^T f_t^2 dt\right)\right] < \infty$$

[1] Note that if f_t is bounded ($|f_t| \leq L < \infty$) then $E\left[\exp\left(\frac{1}{2}\int_0^T f_t^2 dt\right)\right] \leq \exp\left(\frac{T}{2}L^2\right) < \infty$

Let

$$\Lambda_t = \exp\left\{\int_0^t f_s dY_s - \frac{1}{2}\int_0^t |f_s|^2 ds\right\}, \tag{2.31}$$

and suppose that

$$E_Q[\Lambda_t] = 1 \text{ for all } t \in [0, T].$$

If P is the probability measure on $\{\Omega, \mathcal{F}\}$ defined by

$$\frac{dP}{dQ} = \Lambda_T,$$

then the process defined by

$$W_t = Y_t - \int_0^t f_s ds.$$

is a standard Brownian motion on $\{\Omega, \mathcal{F}, \mathcal{F}_t, P\}$.

For a proof, refer to [Elliott 1982].

Remark 2.5. *By Ito formula, the process (Λ_t) defined as in (2.31) satisfies the following equation:*

$$\Lambda_t = 1 + \int_0^t \Lambda_s f_s dY_s, \tag{2.32}$$

and $E_Q[\Lambda_t] = 1$ [Krishnamurthy 2002].

Remark 2.6. *Let (Y_t) be a process determined on $\{\Omega, \mathcal{F}, \mathcal{F}_t, P\}$ by the SDE:*

$$dY_t = f_t dt + dW_t, \tag{2.33}$$

where (W_t) is a standard Brownian motion and the function f is defined as in Theorem 2.5. If Q is the measure defined by

$$\frac{dQ}{dP} = (\Lambda_t)^{-1} = \exp\left\{-\int_0^t f_s dW_s - \frac{1}{2}\int_0^t |f_s|^2 ds\right\},$$

then (Y_t) is a standard Brownian motion under Q and we have, by Itô's formula,

$$(\Lambda_t)^{-1} = 1 - \int_0^t (\Lambda_s)^{-1} f_s dW_s. \tag{2.34}$$

2.7 Continuous-time homogeneous Markov chain

Consider a continuous-time Markov chain $(X)_{t\in(0,T)}$ with state space $\mathbf{S} = \{e_1, e_2, ..., e_N\}$. Let $p_{ji}(t, h)$ be the transition probability from state e_i at time t to state e_j at time $t + h$:

$$p_{ji}(t, h) = P(X_{t+h} = e_j | X_t = e_i).$$

The Markov chain is said homogeneous if $p_{ji}(t, h)$ does not depend on t and only depends on h :

$$P(X_{t+h} = e_j | X_t = e_i) \triangleq p_{ji}(h).$$

The time interval between state transition from state e_i to state e_j is a random variable H_{ji} and note that its distribution is exponential with a given parameter λ_{ji} [Gross 2012].

Therefore

$$P(H_{ji} \leq h) = 1 - \exp(-\lambda_{ji} h),$$

and

$$p_{ji}(h) = 1 - \exp(-\lambda_{ji} h).$$

Let $h = \Delta t$ be a small increment of time so that:

$$p_{ji}(\Delta t) = 1 - \exp(-\lambda_{ji}\Delta t) \approx 1 - (1 - \lambda_{ji}\Delta t) = \lambda_{ji}\Delta t.$$

Assume that $p_{ii}(\Delta t)$ is approximately an affine function of Δt:

$$p_{ii}(\Delta t) = 1 - \lambda_i \Delta t.$$

We can define the transition rates using

$$\lambda_{ji} \triangleq \lim_{\Delta t \to 0} \frac{p_{ji}(\Delta t)}{\Delta t},$$
$$\lambda_i \triangleq \lim_{\Delta t \to 0} \frac{1 - p_{ii}(\Delta t)}{\Delta t}.$$

Note that

$$p_{ii}(\Delta t) + \sum_{j \neq i} p_{ji}(\Delta t) = 1.$$

From this, we have the following relationship

$$\lambda_i = \lim_{\Delta t \to 0} \frac{1 - p_{ii}(\Delta t)}{\Delta t} = \lim_{\Delta t \to 0} \frac{\sum_{j \neq i} p_{ji}(\Delta t)}{\Delta t} = \sum_{j \neq i} \lambda_{ji}, \quad \forall i.$$

The column sums of the probability transition matrix $P(\Delta t) = (p_{ji}(\Delta t))$ is 1. Define the transition rate matrix for (X_t) as

$$A^{\Delta} \triangleq \begin{pmatrix} -\lambda_1 & \lambda_{12} & \dots & \lambda_{1N} \\ \lambda_{21} & -\lambda_2 & \dots & \lambda_{2N} \\ \vdots & \vdots & \dots & \vdots \\ \lambda_{N1} & \lambda_{N2} & \dots & -\lambda_N \end{pmatrix}.$$

It follows from the relations above that:

$$A^{\Delta} = \lim_{\Delta t \to 0} \frac{P(\Delta t) - \mathbf{I}}{\Delta t},$$

where $\mathbf{I}$ is the $N \times N$ unit matrix.

From this, it is seen that

$$P(\Delta t) \approx \Delta t A^{\Delta} + \mathbf{I}.$$

This approximation will be used to establish the transition probability matrix in ours approximate models later (Section 3.3.1 and Section 4.4.1).

2.8 Continuous-time HMM

Following [Dembo 1986, James 1996, Elliott 1995, Elliott 2010], we review now some of the standard definitions and notions of a continuous-time HMM based on a SDE with a noise variance taken to be one.

The *state process* of the model is a continuous-time homogeneous Markov chain $(X_t)_{t\in[0,T]}$ with state space $\mathbf{S} = \{e_1, e_2, ..., e_N\}$ and transition probabilities determined by:

$$P_{ji}(\Delta) = P(X_{t+\Delta} = e_j | X_t = e_i) = \begin{cases} \lambda_{ji}\Delta + o(\Delta), & \text{if } j \neq i,\ \Delta \to 0, \\ 1 + \lambda_{ii}\Delta + o(\Delta), & \text{if } i = j,\ \Delta \to 0. \end{cases} \tag{2.35}$$

where $\lambda_{ii} = -\sum\limits_{j\neq i} \lambda_{ji}$, $i = 1, 2, \dots, N$.

In this case, the transition matrix $A^{\Delta} = (\lambda_{ji}) \in \mathbf{R}^{N\times N}$ is a transition rate matrix (or infinitesimal generator, see Section 2.7). It is well-known that (X_t) now has a semimartingale representation [Elliott 2010, page 198]:

$$X_t = X_0 + \int_0^t A^{\Delta} X_s ds + V_t, \tag{2.36}$$

where (V_t) is an $\mathcal{F}_t$-martingale.

The process (X_t) is not observed directly, it is observed through a scalar process (Y_t) which is assumed to satisfy a stochastic differential equation.

Definition 2.7. *A continuous-time HMM is a pair of processes $(X_t, Y_t)_{t\in[0,T]}$ such that*

$$X_t = X_0 + \int_0^t A^{\Delta} X_s ds + V_t, \tag{2.37}$$

$$dY_t = h(X_t, Y_t, t)dt + dW_t, \tag{2.38}$$

where (V_t) is an $\mathcal{F}_t$-martingale, (W_t) is a standard Brownian motion which is independent of (X_t) and $h_t = h(X_t, Y_t, t)$ is a predictable process such that

$$\int_0^T |h_t|^2 dt < \infty \ a.s.$$

$$h_t \text{ is bounded.}$$

Equation 2.38 is called *the observation equation* of the model.

2.8.1 Filtrations, number of jumps, occupation time and level sums

We next precise some notations in the continuous-time standard HMM which will be used in the problem of parameter estimation.

Definition 2.8. *Let $(X_t, Y_t)_{t\in[0,T]}$ be a continuous-time standard HMM, define:*

1. *the filtration of complete data*

$$\mathcal{G}_t^Y \triangleq \sigma\{X_s, Y_s; 0 \le s \le t\}, \tag{2.39}$$

2. *the filtration of incomplete data*

$$\mathcal{Y}_t^Y \triangleq \sigma\{Y_s; 0 \le s \le t\}, \tag{2.40}$$

3. *the number of jumps of the Markov chain from e_i to e_j until time t*

$$\mathcal{J}_t^{ij} = \int_0^t \langle X_s, e_i\rangle \langle dX_s, e_j\rangle, \tag{2.41}$$

4. *the occupation time of the Markov chain in state e_i until time t*

$$\mathcal{O}_t^i = \int_0^t \langle X_s, e_i\rangle ds, \tag{2.42}$$

5. *the level sums of the observation process in state e_i up to time t*

$$\mathcal{T}_t^i(f) = \int_0^t f(Y_s)\langle X_s, e_i\rangle dY_s, \tag{2.43}$$

   where $f(\cdot)$ is any bounded function.

Remark 2.7. *We have $\mathcal{Y}_t^Y \subset \mathcal{G}_t^Y$, for all $t \in [0, T]$.*

2.8.2 Change of measure

Let θ be the parameter set of the model (X_t, Y_t) (Definition 2.8) and write P_θ for the measure determined by the set θ.

Again, in the computations of parameter estimation, we will also use the reference probability method with the change-of-measure technique similar to the discrete-time case.

For this, we first introduce a reference probability measure $\overline{P}$ from the initial measure P_θ by putting:

$$\left.\frac{d\overline{P}}{dP_\theta}\right|_{\mathcal{G}_t^Y} = \overline{\Lambda}_t^{Y,\theta} = \exp\left\{-\int_0^t h_s dW_s - \frac{1}{2}\int_0^t |h_s|^2 ds\right\}. \tag{2.44}$$

By Girsanov's theorem (see Theorem 2.5, Remark 2.6), (Y_t) is a standard Brownian motion under $\overline{P}$.

We will work with a standard Brownian motion (Y_t) defined on the reference probability space $(\Omega, \mathcal{F}, \overline{P})$.

Then, the obtained results will be reinterpreted in the initial probability space $(\Omega, \mathcal{F}, P_\theta)$ by a back change of measure form $\overline{P}$ to P_θ:

$$\left.\frac{dP_\theta}{d\overline{P}}\right|_{\mathcal{G}_t^Y} = \Lambda_t^{Y,\theta} = \exp\left\{\int_0^t h_s dY_s - \frac{1}{2}\int_0^t |h_s|^2 ds\right\}. \tag{2.45}$$

Specifically, the results obtained on $(\Omega, \mathcal{F}, \overline{P})$ for the unnormalized filters $\gamma(H_t) = \overline{E}(\Lambda_t^{Y,\theta} H_t | \mathcal{Y}_t^Y)$, $H_t \equiv \mathcal{J}_t^{ij}, \mathcal{O}_t^i, \mathcal{T}_t^i(f)$, will be related to $(\Omega, \mathcal{F}, P_\theta)$ for obtaining the normalized filters $\pi(H_t) = E_\theta(H_t|\mathcal{Y}_t^Y)$ by Bayes's rule:

$$E_\theta(H_t|\mathcal{Y}_t^Y) = \frac{\overline{E}(\Lambda_t^{Y,\theta} H_t|\mathcal{Y}_t^Y)}{\overline{E}(\Lambda_t^{Y,\theta}|\mathcal{Y}_t^Y)}, \tag{2.46}$$

or

$$\pi(H_t) = \frac{\gamma(H_t)}{\gamma(1)}. \tag{2.47}$$

Remark 2.8. *As in the discrete-time HMM, we have*

$$\pi(H_t) = \frac{\langle \gamma(H_t X_t), \underline{1} \rangle}{\langle \gamma(X_t), \underline{1} \rangle}. \tag{2.48}$$

Remark 2.9. *From* (2.45) *and by applying Ito formula (see Section 2.6.2, Theorem 2.4) with* $\xi_t = \int_0^t h_s dY_s - \frac{1}{2}\int_0^t |h_s|^2 ds$, $f(t, \xi_t) = \Lambda_t^{Y,\theta} = e^{\xi_t}$, *we obtain*

$$\Lambda_t^{Y,\theta} = 1 + \int_0^t \Lambda_s^{Y,\theta} h_s dY_s, \tag{2.49}$$

and

$$\overline{\Lambda}_t^{Y,\theta} = 1 - \int_0^t \overline{\Lambda}_s^{Y,\theta} h_s dW_s. \tag{2.50}$$

Note that $\overline{\Lambda}_t^{Y,\theta} \Lambda_t^{Y,\theta} = 1$.

2.9 Parameter estimation

Because the state process is unobserved, the parameter estimation can be only based on incomplete data $\mathcal{Y}_{time}^{\cdot}$ ($\mathcal{Y}_{time}^{\cdot} \equiv \mathcal{Y}_h^K$ or $\mathcal{Y}_{time}^{\cdot} \equiv \mathcal{Y}_t^Y$). In this case, the Maximum Likelihood Estimator cannot be easily handled [James 1996, Charalambous 2000, Elliott 2010] so that several algorithms were proposed by statisticians and engineers, among the most successful ones is the EM algorithm.

We first describe the *likelihood function*, the *pseudo log-likelihood* used in the EM algorithm and the two main steps of this algorithm. We then summarize without proofs some relevant theorems on the stationary and the converging properties of the sequence of estimates.

2.9.1 Likelihood function

Let $\{P_\theta, \theta \in \Theta\}$ be a family of probability measures and let $\overline{P}$ be a reference probability mesure defined as in Section 2.4, the *likelihood function* for computing an estimate θ' of parameter vector θ based on the available information $\mathcal{Y}_{time}^{\cdot}$ is:

$$L(\theta) = \overline{E}\left(\frac{dP_\theta}{d\overline{P}} | \mathcal{Y}_{time}^{\cdot}\right),$$

and the ML estimate θ_{MLE} is defined by

$$\theta_{MLE} = \arg\max_{\theta \in \Theta} L(\theta).$$

2.9.2 Pseudo log-likelihood function

The EM algorithm provides an iterative numerical method which can be used to generate a sequence $\{\theta^{(p)}, p \geq 0\}$ for updating the ML estimates of parameter vector θ. To this end, we first consider the following *log-likelihood function*:

$$\mathcal{L}(\theta) = \log \left[\overline{E} \left(\frac{dP_\theta}{d\overline{P}} \middle| \mathcal{Y}^{\cdot}_{time} \right) \right].$$

Due to the monotonicity of the log function, the maximization of $\mathcal{L}(\theta)$ is equivalent to the maximization of $L(\theta)$.

Let P^y_θ denote the restriction of P_θ to $\mathcal{Y}^{\cdot}_{time}$. It can be proved [Dembo 1986, Lipster 2010], that

$$\frac{dP^y_{\theta'}}{dP^y_\theta} = E_\theta \left(\frac{dP_{\theta'}}{dP_\theta} \middle| \mathcal{Y}^{\cdot}_{time} \right).$$

Then

$$\begin{aligned} \mathcal{L}(\theta') - \mathcal{L}(\theta) =& \log \frac{dP^y_{\theta'}}{d\overline{P}^y} - \log \frac{dP^y_\theta}{d\overline{P}^y} \\ =& \log \frac{dP^y_{\theta'}}{dP^y_\theta}, \end{aligned}$$

and

$$\mathcal{L}(\theta') - \mathcal{L}(\theta) = \log E_\theta \left(\frac{dP_{\theta'}}{dP_\theta} \middle| \mathcal{Y}^{\cdot}_{time} \right).$$

Jensen inequality then yields

$$\mathcal{L}(\theta') - \mathcal{L}(\theta) \geq E_\theta \left(\log \frac{dP_{\theta'}}{dP_\theta} \middle| \mathcal{Y}^{\cdot}_{time} \right),$$

Therefore, for all $\theta, \theta' \in \Theta$, we have obtained

$$\mathcal{L}(\theta') - \mathcal{L}(\theta) \geq Q(\theta', \theta),$$

with equality if and only if $\theta = \theta'$, where

$$Q(\theta', \theta) \triangleq E_\theta \left(\log \Lambda^{\cdot, \theta\theta'}_{time} | \mathcal{Y}^{\cdot}_{time} \right) \text{ with } \Lambda^{\cdot, \theta\theta'}_{time} \triangleq \left. \frac{dP_{\theta'}}{dP_\theta} \right|_{\mathcal{G}^{\cdot}_{time}}. \tag{2.51}$$

The function $Q(\theta', \theta)$ as defined in (2.51) is called the *pseudo log-likelihood function.*

2.9.3 EM Algorithm

Starting with an initial parameter vector $\theta^{(0)}$, each iteration of EM algorithm consists in the two following steps:

E-Step *(Expectation Step)* : From (2.51), set $\theta = \theta'^{(p)}$ and compute the pseudo log-likelihood function

$$Q(\theta', \theta'^{(p)}) = E_{\theta'^{(p)}} \left[\log \Lambda_{time}^{\cdot, \theta'^{(p)}\theta'} | \mathcal{Y}_{time}^{\cdot} \right],$$

M-Step *(Maximization Step)* : Find $\theta'^{(p+1)} \in \arg\max_{\theta' \in \Theta} Q(\theta', \theta'^{(p)})$.

Repeat from **E-Step** with $p = p + 1$, unless a stopping test is satisfied. The stationary and converging properties of the EM algorithm were evaluated by [Dempster 1977, Dembo 1986] via Theorem 2.6, Theorem 2.7 and Theorem 2.8 below.

Theorem 2.6. *For every sequence* $\{\theta^{(p)}, p = 0, 1, 2, \ldots\}$ *generated by EM algorithm,*

$$\mathcal{L}(\theta^{(p+1)}) \geq \mathcal{L}(\theta^{(p)}) \tag{2.52}$$

with equality if and only if $\theta^{(p+1)} = \theta^{(p)}$.

Theorem 2.6 is Theorem 1 in [Dembo 1986] and the infinite dimensional analogue of Theorem 1 in [Dempster 1977].

Theorem 2.7. *Suppose that* $\{\theta^{(p)}, p = 0, 1, 2, \ldots\}$ *is an instance generated by EM algorithm such that*

(1) *the sequence* $\mathcal{L}(\theta^{(p)})$ *is bounded, and*

(2) $Q(\theta^{(p+1)}|(\theta^{(p)})) - Q(\theta^{(p)}|(\theta^{(p)})) \geq \lambda(\theta^{(p+1)} - \theta^{(p)})^2$ *for some* $\lambda > 0$ *and all* p.

Then the sequence $\{\theta^{(p)}, p = 0, 1, 2, \ldots\}$ *converges to some* θ^* *in the closure of* Θ.

The proof can be found in [Dempster 1977, Theorem 2].

Theorem 2.8. *Assume that*

(1) $\Theta_{\theta_0} = \{\theta \in \Theta : \mathcal{L}(\theta) \geq \mathcal{L}(\theta_0)\}$ *is compact for any* $\theta_0 \in \Theta$.

(2) $\mathcal{L}(\theta)$ *is continuous in* Θ *and differentiable in the interior of* Θ.

(3) $Q(\theta', \theta)$ *is continuous with respect to* θ *and* θ'.

(4) *All the EM instances* $\{\theta^{(p)}, p = 0, 1, 2, \ldots\}$ *are in the interior of* Θ.

Then, all the limit points of any instance $\{\theta^{(p)}, p = 0, 1, 2, \ldots\}$ *of the EM algorithm (and there exists at least one such limit point) are stationary points of* $\mathcal{L}(\cdot)$*, having the same value* L^* *(of* $\mathcal{L}(\cdot)$*). Furthermore,* $\mathcal{L}(\theta^{(p)})$ *converges monotonically to* L^*.

Theorem 2.8 was proved by [Dembo 1986, Theorem 2].

Chapter 3

Stochastic models for clearness index processes

Contents

Résumé

Dans ce chapitre, nous proposons un modèle de type HMM à temps discret, noté DTM-K, pour une suite d'indices journaliers de clarté $(K_h)_{h=1,2,\dots}$, ainsi qu'un modèle de type HMM à temps continu, noté CTM-k, pour un processus d'indice de clarté $(k_t)_{t\in[0,T]}$ sur un intervalle de temps $[0,T]$. Les estimations numériques des paramètres de CTM-k nécessite de construire un modèle approché à temps discret, noté DTAM-k, obtenu en discrétisant le temps par une partition de $[0,T]$.

Pour chacun de ces modèles, nous définissons un processus des états, un processus des observations, et un vecteur de paramètres. Nous détaillons les équations de filtrages pour obtenir les formules de mise à jour des estimations dans les itérations de l'algorithme EM, en particulier pour estimer la variance du bruit.

Les applications numériques ont été faites en utilisant des données provenant de mesures effectuées à La Réunion pour DTM-K et en Guadeloupe pour CTM-k et DTAM-k.

Les résultats de ce chapitre ont été présentés à la conférence [Tran 2013] (voir aussi [Tran a]).

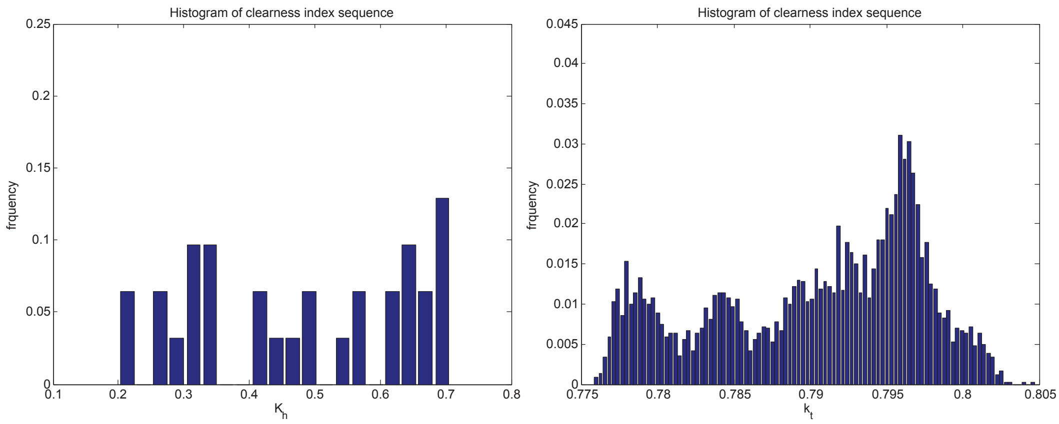
Abstract

In this chapter, we propose a discrete-time HMM model, denoted DTM-K, for a sequence of daily clearness index $(K_h)_{h=1,2,\dots}$, and also a continuous-time HMM model, denoted CTM-k, for a process $(k_t)_{t\in[0,T]}$ of clearness index on a time interval $[0,T]$. Numerical estimations of CTM-k parameters requires to build a discrete-time approximating model, denoted by DTAM-k, obtained by discretizing time, using a partition of $[0,T]$.

For each of these models, we define a state process, an observation process, and a parameter vector. We detail the filtering equations in order to get estimation update formulas used in the iterations of EM algorithm, in particular for estimating the noise variance.

Experiments for DTM-K (resp. for CTM-k and DTAM-k) are done using data coming from measurements performed in La Réunion island (resp. in Guadeloupe island).

The results of this chapter have been presented in [Tran 2013] conference (also see [Tran a]).



(a) (b)

Figure 3.1: Histogram of observed data: (a) daily CIS measured in October 2010, La Réunion; (b) k-DATA-II.1 (see Table 3.3).

3.1 Modelling a daily clearness index sequence

The empirical distribution of a daily Clearness Index Sequence (CISs) during a period (for instance during October 2010 Figure 3.1a) suggests that the daily CIS distribution could be a Gaussian mixture, each Gaussian component corresponding, may be, to some specific meteorological regime. This has lead us to modelize the dynamic of the sequence by a discrete-time HMM where

1. the unobserved state process is a Markov chain representing the dynamic of regimes, each daily index belonging to a regime, several daily indices belonging eventually to a same regime,

2. the observed process is such that, given (or within) regime i, the various observed daily clearness indices are outcomes of a Gaussian distribution whose mean b_i and standard deviation α_i depend on regime i, $i = 1, 2, \dots, N$.

Actually, each regime corresponds to a Gaussian component of the suggested Gaussian mixture, and in terms of probabilistic classification, each regime corresponds to a (Gaussian) class. The advantage of considering a HMM is that it provides a parametric description of the random dynamic of the regimes, which is not the case in a classification setting.

In this section we propose a discrete-time HMM $(X_h, K_h)_{h=1,2,\dots}$ to model a daily clearness index sequence in random environment. We first describe the *state process*, the *observation process* and the parameters of the model denoted DTM-K. Then we will detail the pseudo log-likelihood function, the filtering equations and the computations used in the EM algorithm. Experiments with real data will be presented in Section 3.4.

3.1.1 State process

The random dynamic of meteorological regimes will be modellized by a *state process*, an unobserved discrete-time, finite-state homogeneous Markov chain $(X_h)_{h=0,1,2,\ldots}$ (Figure 3.2).

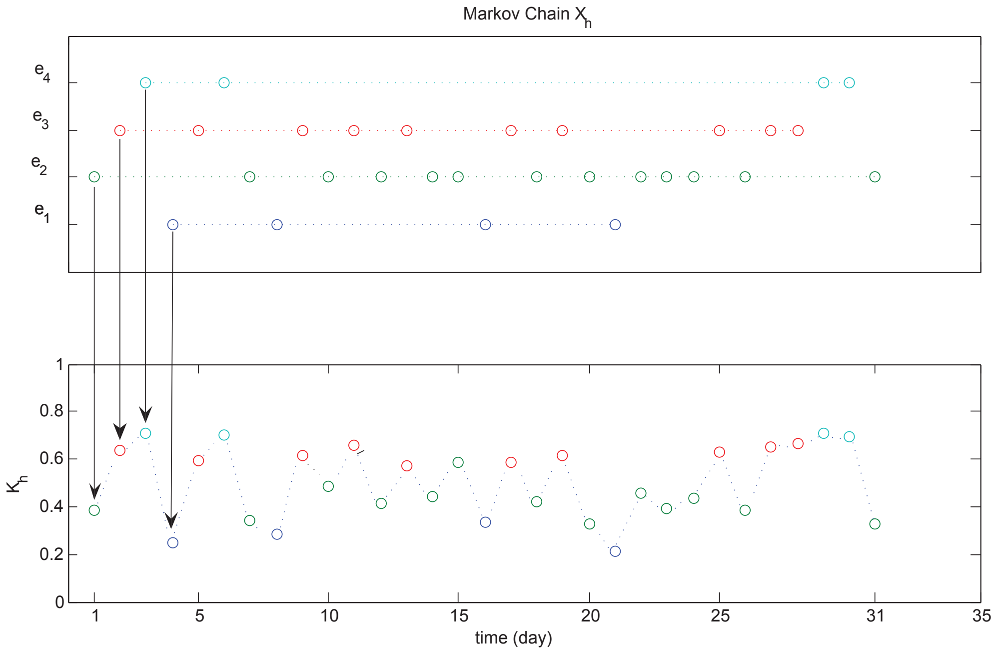


Figure 3.2: An illustration for the DTM-K: $(X_h, K_h)_{h=1,2,\ldots}$.

The state space of (X_h) is the set of unit vectors:

$$\mathbf{S} = \{e_1, e_2, \ldots, e_N\}, \quad e_i = (0, \ldots, 0, \underbrace{1}_{i-th}, 0, \ldots, 0)' \in \mathbf{R}^N.$$

Recall that we can write $X_h = (\langle X_h, e_1\rangle, \langle X_h, e_2\rangle, \ldots, \langle X_h, e_N\rangle)'$ and the process $(X_h)_{h=1,2,\ldots}$ can be represented by the following equation, called the state equation,

$$X_h = AX_{h-1} + V_h, \tag{3.1}$$

where $(V_h)_{h=1,2,\ldots}$ is a sequence of martingale increments.

3.1.2 Observation Process and model parameters

The random values of a daily CIS (K_h) are modelled by the so-called *observation process* as follows. In regime i, that is when the Markov chain is in state e_i, $(i = 1, 2, \ldots, N)$ the daily clearness index K_h will be considered as an outcome of a Gaussian distribution $\mathcal{N}(b_i, \alpha_i^2)$ depending on regime i. In other words:

$$\mathbf{1}_{(X_h=e_i)}K_h = \mathbf{1}_{(X_h=e_i)}(b_i + \alpha_i w_h), \quad h = 1, 2, \ldots, \tag{3.2}$$

where w_h are $\mathcal{N}(0,1)$ i.i.d. random variables and b_i, α_i are parameters to be estimated.

The DM-K model of daily CIS under the random effects of meteorological events will be the HMM $(X_h, K_h)_{h=1,2,\ldots}$, where the state process (X_h) is determined by (3.1) and the *observation process* (K_h) is defined as

$$K_h = \sum_{i=1}^{N} \mathbf{1}_{(X_h=e_i)} K_h = \sum_{i=1}^{N} \mathbf{1}_{(X_i=e_i)} \left(b_i + \alpha_i w_h\right), \quad h = 1, 2, \ldots. \tag{3.3}$$

If $b = (b_1, b_2, \ldots, b_N)'$ and $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_N)'$, then we have equivalently

$$K_h = \langle X_h, b\rangle + \langle X_h, \alpha\rangle w_h, \quad h = 1, 2, \ldots. \tag{3.4}$$

The parameter vector for the proposed model DM-K is

$$\theta = (a_{ji}, 1 \leq j \neq i \leq N;\ b_1, b_2, \ldots, b_N;\ \alpha_1, \alpha_2, \ldots, \alpha_N),$$

where the number of states N will be suggested by the user.

3.1.3 Parameter estimation

As mentioned in Chapter 2, our aim is to determine a new parameter vector $\theta' = \{(a'_{ji}), 1 \leq j \neq i \leq N; b', \alpha'\}$ which maximizes the pseudo log-likelihood function at each iteration of EM algorithm, where $b' = (b'_1, b'_2, \ldots, b'_N)'$, $\alpha' = (\alpha'_1, \alpha'_2, \ldots, \alpha'_N)'$.

Updating the transition probabilities a_{ji} is as follows ([Elliott 1995, chap. 3, pg. 68-70] and [James 1996]):

$$a'_{ji} = \frac{\pi(\mathcal{J}_h^{ji})}{\pi(\mathcal{O}_h^{i})} \quad 1 \leq j \neq i \leq N, \tag{3.5}$$

where $\pi(\mathcal{J}_h^{ji})$ and $\pi(\mathcal{O}_h^{i})$ are the normalized filters (see Section 2.5.3) of the number of jumps and the occupation time, respectively.

Note that in (3.5), π is built using θ (corresponding to $\theta'^{(p)}$ in *p-th* iteration of the EM algorithm), and not θ' (corresponding to $\theta'^{(p+1)}$).

We now consider the update from b and α to b' and α', respectively.

3.1.3.1 Pseudo log-likelihood function

Given the filtration $\mathcal{Y}_h^K$ of observed data at day h, consider the conditional expectation of the log-likelihood function:

$$Q(\theta', \theta) = E_\theta \left[\log \Lambda_h^{K,\theta\theta'} | \mathcal{Y}_h^K\right], \tag{3.6}$$

where $\Lambda_h^{K,\theta\theta'} = \frac{dP_{\theta'}}{dP_\theta}\big|_{\mathcal{G}_h^K}$.

We first have

$$\Lambda_h^{K,\theta\theta'} = \prod_{l=1}^{h} \frac{\langle X_l, \alpha\rangle \phi\left(\frac{K_l - \langle X_l, b'\rangle}{\langle X_l, \alpha'\rangle}\right)}{\langle X_l, \alpha'\rangle \phi\left(\frac{K_l - \langle X_l, b\rangle}{\langle X_l, \alpha\rangle}\right)}, \tag{3.7}$$

where $\phi(\cdot)$ denotes the $\mathcal{N}(0,1)$ density.

From (3.6) and (3.7), we get:

$$Q(\theta',\theta) = E_\theta\left\{\sum_{l=1}^{h}\left[\log\frac{1}{\langle X_l,\alpha'\rangle} - \frac{1}{2}\left(\frac{K_l-\langle X_l,b'\rangle}{\langle X_l,\alpha'\rangle}\right)^2\right]|\mathcal{Y}_h^K\right\} + R(\theta,\mathcal{Y}_h^K), \quad (3.8)$$

where the function $R(\theta,\mathcal{Y}_h^K)$ does not depend on θ'.

3.1.3.2 Computations in EM algorithm

E-Step *(Expectation Step)* : Set $\theta = \theta'^{(p)}$ and rewrite (3.8) as

$$\begin{aligned} Q(\theta',\theta'^{(p)}) &= \sum_{i=1}^{N} E_{\theta'(p)}\left\{\sum_{l=1}^{h}\left[\langle X_l,e_i\rangle\log\frac{1}{\alpha_i'} - \frac{1}{2}\frac{\langle X_l,e_i\rangle}{\alpha_i'^2}(K_l-b_i')^2\right]|\mathcal{Y}_h^K\right\} \\ &\quad + R(\theta'^{(p)},\mathcal{Y}_h^K) \\ &= \sum_{i=1}^{N}\left\{\pi(\mathcal{O}_h^i)\log\frac{1}{\alpha_i'} - \frac{1}{2\alpha_i'^2}\left[\pi(\mathcal{T}_h^i(K_h^2)) - 2b_i'\pi(\mathcal{T}_h^i(K_h)) + b_i'^2\pi(\mathcal{O}_h^i)\right]\right\} \\ &\quad + R(\theta'^{(p)},\mathcal{Y}_h^K), \end{aligned}$$

where

$$\mathcal{O}_h^i = \sum_{l=1}^{h}\langle X_l,e_i\rangle, \quad (3.9)$$

$$\mathcal{T}_h^i(K_h) = \sum_{l=1}^{h} K_h\langle X_l,e_i\rangle, \quad (3.10)$$

$$\mathcal{T}_h^i(K_h^2) = \sum_{l=1}^{h} K_h^2\langle X_l,e_i\rangle, \quad (3.11)$$

and $\pi(H_h) = E_{\theta'(p)}(H_h|\mathcal{Y}_h^K)$ for $H_h \equiv \mathcal{O}_h^i$, $\mathcal{T}_h^i(K_h)$ or $\mathcal{T}_h^i(K_h^2)$.

M-Step *(Maximization Step)* : Let us find now $\theta'^{(p+1)} \in \arg\max_{\theta'\in\Theta} Q(\theta',\theta'^{(p)})$.

1. Taking derivative of $Q(\theta',\theta'^{(p)})$ with respect to $b_i, i = 1,2,\ldots,N$, we obtain

$$\frac{\partial}{\partial b_i'}Q(\theta',\theta'^{(p)}) = -\frac{1}{2\alpha_i'^2}\left[-2\pi(\mathcal{T}_h^i(K_h)) + 2b_i'\pi(\mathcal{O}_h^i)\right]. \quad (3.12)$$

Now $\frac{\partial}{\partial b_i'}Q(\theta',\theta'^{(p)}) = 0$ yields

$$b_i' = \frac{\pi(\mathcal{T}_h^i(K_h))}{\pi(\mathcal{O}_h^i)}. \quad (3.13)$$

2. Similarly, for $i = 1,2,\ldots,N$, $\frac{\partial}{\partial\alpha_i'}Q(\theta',\theta'^{(p)}) = 0$ yields

$$\alpha_i'^2 = \frac{1}{\pi(\mathcal{O}_h^i)}\left[\pi(\mathcal{T}_h^i(K_h^2)) - 2b_i'\pi(\mathcal{T}_h^i(K_h)) + b_i'^2\pi(\mathcal{O}_h^i)\right]. \quad (3.14)$$

3.1.3.3 Updating parameter

For $H_h \equiv \mathcal{O}_h^i$, $\mathcal{T}_h^i(K_h)$ or $\mathcal{T}_h^i(K_h^2)$, as $\pi(H_h) = \frac{\langle\gamma(H_h X_h),\underline{1}\rangle}{\langle\gamma(X_h),\underline{1}\rangle}$, (3.5), (3.13), (3.14) become:

$$a'_{ji} = \frac{\langle\gamma(X_h \mathcal{J}_h^{ij}),\underline{1}\rangle}{\langle\gamma(X_h\mathcal{O}_h^i),\underline{1}\rangle}, \quad b'_i = \frac{\langle\gamma(X_h\mathcal{T}_h^i(K_h)),\underline{1}\rangle}{\langle\gamma(X_h\mathcal{O}_h^i),\underline{1}\rangle}, \quad 1 \le j \ne i \le N, \tag{3.15}$$

$$\alpha_i'^2 = \left(\langle\gamma(X_h\mathcal{O}_h^i),\underline{1}\rangle\right)^{-1} \times \left[\langle\gamma(X_h\mathcal{T}_h^i(K_h^2)),\underline{1}\rangle - 2b'_i\langle\gamma(X_h\mathcal{T}_h^i(K_h)),\underline{1}\rangle + b_i'^2\langle\gamma(X_h\mathcal{O}_h^i),\underline{1}\rangle\right]. \tag{3.16}$$

It remains to establish some recursive equations to compute $\gamma(X_h\mathcal{J}_h^{ji})$, $\gamma(X_h\mathcal{O}_h^i)$, $\gamma(X_h\mathcal{T}_h^i(K_h))$ and $\gamma(X_h\mathcal{T}_h^i(K_h^2))$. This is done in the next section.

3.1.4 Filtering equations

Definition 3.1. *Let $\Lambda_h^{K,\theta}$ be the process determined as in (2.19), $h = 1, 2, \ldots$. Define*

$$q_h = \gamma(X_h) = \overline{E}(\Lambda_h^{K,\theta} X_h | \mathcal{Y}_h^K). \tag{3.17}$$

The process (q_h) is called the state filter process of model (X_h, K_h).

In oder to obtain a recursive equation for $(q_h), h = 1, 2, \ldots$, introduce the following quantities:

$$B_h^{(i)} \triangleq \frac{\phi\left(\frac{K_h - b_i}{\alpha_i}\right)}{\alpha_i \phi(K_h)}, \quad i = 1, 2, \ldots, N,$$
$$B_h \triangleq \operatorname{diag}(B_h^{(1)}, B_h^{(2)}, \ldots, B_h^{(N)}).$$

Lemma 3.1. *The recursive equation of state process q_h is given by*

$$q_h = B_h A q_{h-1}, \quad h = 1, 2, \ldots \tag{3.18}$$

with initial condition $q_0 = \pi_0$.

Proof. By definition of $(q_h)_{h=1,2,\ldots}$ we have

$$\begin{aligned}
q_h =& \overline{E}(\Lambda_h^{K,\theta} X_h | \mathcal{Y}_h^K) \\
=& \overline{E}\left[\Lambda_{h-1}^{K,\theta} \frac{\phi\left(\frac{K_h - \langle X_h, b\rangle}{\langle X_h \alpha\rangle}\right)}{\langle X_h, \alpha\rangle \phi(K_h)} X_h \middle| \mathcal{Y}_h^K\right] \\
=& \sum_{i=1}^N \overline{E}\left[\Lambda_{h-1}^{K,\theta}\langle X_h, e_i\rangle | \mathcal{Y}_{h-1}^K\right] B_h^{(i)} e_i \\
=& \sum_{i=1}^N \overline{E}\left\{\overline{E}\left[\Lambda_{h-1}^{K,\theta}\langle X_h, e_i\rangle | \mathcal{G}_{h-1}^K\right] | \mathcal{Y}_{h-1}^K\right\} B_h^{(i)} e_i,
\end{aligned}$$

Noticing that $\mathcal{Y}_{h-1}^K \subset \mathcal{G}_{h-1}^K$, we get

$$
\begin{aligned}
q_h =& \sum_{i=1}^{N} \overline{E}\left\{\Lambda_{h-1}^{K,\theta}\overline{E}\big[\langle X_h, e_i\rangle|\mathcal{G}_{h-1}^K\big]\big|\mathcal{Y}_{h-1}^K\right\} B_h^{(i)} e_i \\
=& \sum_{i=1}^{N} \overline{E}\big[\Lambda_{h-1}^{K,\theta}\langle AX_{h-1}, e_i\rangle|\mathcal{K}_{h-1}\big] B_h^{(i)} e_i \\
=& \sum_{i=1}^{N} \langle Aq_{h-1}, e_i\rangle B_h^{(i)} e_i \\
=& B_h A q_{h-1}.
\end{aligned}
$$

□

Using the state filter process (q_h) computed in Lemma 3.1 above, the recursive equation for the number of jumps, for the occupation time and for the level sum are now stated in Theorem 3.1 and Theorem 3.2 below.

Theorem 3.1. *Let $\mathcal{J}_h^{ij} = \sum_{l=1}^{h} \langle X_{l-1}, e_i\rangle\langle X_h, e_j\rangle$ be the number of jumps from state e_i to state e_j up to time h. The process $\gamma(\mathcal{J}_h^{ij} X_h)$ is computed by the following recursive equation:*

$$
\gamma(\mathcal{J}_h^{ij} X_h) = B_h A\gamma(\mathcal{J}_{h-1}^{ij} X_{h-1}) + \langle q_{h-1}, e_i\rangle\langle B_h A e_i, e_j\rangle e_j, \quad h = 1, 2, \ldots, \tag{3.19}
$$

with initial condition $\gamma(\mathcal{J}_0^{ij} X_0) = \underline{0}$.

Proof. We have

$$
\begin{aligned}
\gamma(\mathcal{J}_h^{ij} X_h) =& \overline{E}(\Lambda_h^{K,\theta}\mathcal{J}_h^{ij} X_h|\mathcal{Y}_h^K) \\
=& \overline{E}\left[\Lambda_{h-1}^{K,\theta}\frac{\phi\left(\frac{K_h - \langle X_h, b\rangle}{\langle X_h, \alpha\rangle}\right)}{\langle X_h, \alpha\rangle\phi(K_h)}\left(\mathcal{J}_{h-1}^{ij} + \langle X_{h-1}, e_i\rangle\langle X_h, e_j\rangle\right) X_h\middle|\mathcal{Y}_h^K\right] \\
=& \sum_{l=1}^{N} \overline{E}\big[\Lambda_{h-1}^{K,\theta}\langle X_h, e_l\rangle\left(\mathcal{J}_{h-1}^{ij} + \langle X_{h-1}, e_i\rangle\langle X_h, e_j\rangle\right)|\mathcal{Y}_{h-1}^K\big] B_h^{(l)} e_l \\
=& \sum_{l=1}^{N} \overline{E}\big[\Lambda_{h-1}^{K,\theta}\langle X_h, e_l\rangle\mathcal{J}_{h-1}^{ij}|\mathcal{Y}_{h-1}^K\big] B_h^{(l)} e_l \\
&+ \overline{E}\big[\Lambda_{h-1}^{K,\theta}\langle X_h, e_j\rangle\langle X_{h-1}, e_i\rangle|\mathcal{Y}_{h-1}^K\big] B_h^{(j)} e_j.
\end{aligned}
\tag{3.20}
$$

As $\mathcal{Y}_h^K \subset \mathcal{G}_h^K$, conditional expectation properties imply that

$$
\begin{aligned}
\overline{E}\big[\Lambda_{h-1}^{K,\theta}\langle X_h, e_l\rangle\mathcal{J}_{h-1}^{ij}|\mathcal{Y}_{h-1}^K\big] =& \overline{E}\left\{\overline{E}\big[\Lambda_{h-1}^{K,\theta}\langle X_h, e_l\rangle\mathcal{J}_{h-1}^{ij}|\mathcal{G}_{h-1}^K\big]\big|\mathcal{Y}_{h-1}^K\right\} \\
=& \overline{E}\big[\Lambda_{h-1}^{K,\theta}\langle AX_{h-1}, e_l\rangle\mathcal{J}_{h-1}^{ij}|\mathcal{Y}_{h-1}^K\big] \\
=& \langle A\gamma(\mathcal{J}_{h-1}^{ij} X_h), e_l\rangle.
\end{aligned}
\tag{3.21}
$$

Similarly,

$$
\begin{aligned}
\overline{E}[\Lambda_{h-1}^{K,\theta}\langle X_h, e_l\rangle\langle X_{h-1}, e_i\rangle|\mathcal{Y}_{h-1}^K] =& \overline{E}\left\{\overline{E}[\Lambda_{h-1}^{K,\theta}\langle X_h, e_j\rangle\langle X_{h-1}, e_i\rangle|\mathcal{G}_{h-1}^K]\,|\mathcal{Y}_{h-1}^K\right\}\\
=& \overline{E}[\Lambda_{h-1}^{K,\theta}\langle AX_{h-1}, e_j\rangle\langle X_{h-1}, e_i\rangle|\mathcal{Y}_{h-1}^K]\\
=& a_{ji}\langle q_{h-1}, e_i\rangle. \qquad (3.22)
\end{aligned}
$$

Plugging (3.19), (3.21) into (3.22) and noticing that $a_{ji}B_h^{(j)} = \langle B_h A e_i, e_j\rangle$, we get

$$
\begin{aligned}
\gamma(\mathcal{J}_h^{ij}X_h) =& \sum_{l=1}^{N}\langle A\gamma(\mathcal{J}_{h-1}^{ij}X_h), e_l\rangle B_h^{(l)}e_l + a_{ji}\langle q_{h-1}, e_i\rangle B_h^{(j)}e_j\\
=& B_h A\gamma(\mathcal{J}_{h-1}^{ij}X_{h-1}) + \langle q_{h-1}, e_i\rangle\langle B_h A e_i, e_j\rangle e_j, \quad h = 1, 2, \ldots. \qquad (3.23)
\end{aligned}
$$

□

Theorem 3.2. *Let $\mathcal{T}_h^i(f) = \sum_{l=1}^{h} f(K_l)\langle X_{l-1}, e_i\rangle$ be the level sum for state e_i, where $f(\cdot)$ is any bounded function. The process $\gamma(\mathcal{T}_h^i(f)X_h)$ is determined by:*

$$
\gamma(\mathcal{T}_h^i(f)X_h) = B_h A\gamma(\mathcal{T}_{h-1}^i(f)X_{h-1}) + f(K_h)\langle q_{h-1}, e_i\rangle B_h A e_i, \qquad (3.24)
$$

with initial condition $\gamma(\mathcal{T}_0^i(f)X_0) = \underline{0}$.

Proof. Starting with the definition of $\gamma(\mathcal{T}_h^i(f)X_h)$ and noticing that $\mathcal{Y}_h^K \subset \mathcal{G}_h^K$, similar arguments as those used in Lemma 3.1 and Theorem 3.1 give :

$$
\begin{aligned}
\gamma(\mathcal{T}_h^i(f)X_h) =& \overline{E}(\Lambda_h^{K,\theta}\mathcal{T}_h^i(f)X_h|\mathcal{Y}_h^K)\\
=& \overline{E}\left[\Lambda_{h-1}^{K,\theta}\frac{\phi\left(\frac{K_h - \langle X_h, b\rangle}{\langle X_h, \alpha\rangle}\right)}{\langle X_h, \alpha\rangle\phi(K_h)}\left(\mathcal{T}_{h-1}^i(f) + f(K_h)\langle X_{h-1}, e_i\rangle\right)X_h\middle|\mathcal{Y}_h^K\right]\\
=& \sum_{l=1}^{N}\overline{E}[\Lambda_{h-1}^{K,\theta}\langle X_h, e_l\rangle\mathcal{T}_{h-1}^i(f)|\mathcal{Y}_{h-1}^K]B_h^{(l)}e_l\\
&+ \sum_{l=1}^{N}\overline{E}[\Lambda_{h-1}^{K,\theta}\langle X_h, e_l\rangle\langle X_{h-1}, e_i\rangle|\mathcal{Y}_{h-1}^K]f(K_h)B_h^{(l)}e_l\\
=& \sum_{l=1}^{N}\overline{E}\left\{\overline{E}[\Lambda_{h-1}^{K,\theta}\langle X_h, e_l\rangle\mathcal{T}_{h-1}^i(f)|\mathcal{G}_{h-1}^K]\middle|\mathcal{Y}_{h-1}^K\right\}B_h^{(l)}e_l\\
&+ \sum_{l=1}^{N}\overline{E}\left\{\overline{E}[\Lambda_{h-1}^{K,\theta}\langle X_h, e_l\rangle\langle X_{h-1}, e_i\rangle|\mathcal{G}_{h-1}^K]\middle|\mathcal{Y}_{h-1}^K\right\}f(K_h)B_h^{(l)}e_l.
\end{aligned}
$$

It follows that

$$\begin{aligned}
\gamma(\mathcal{T}_h^i(f)X_h) =& \sum_{l=1}^{N} \overline{E}\big[\Lambda_{h-1}^{K,\theta}\langle AX_{h-1}, e_l\rangle \mathcal{T}_{h-1}^i(f)\big|\mathcal{Y}_{h-1}^K\big] B_h^{(l)} e_l \\
&+ \sum_{l=1}^{N} \overline{E}\big[\Lambda_{h-1}^{K,\theta}\langle X_{h-1}, e_i\rangle \big|\mathcal{Y}_{h-1}^K\big] a_{li} f(K_h) B_h^{(l)} e_l \\
=& \sum_{l=1}^{N} \langle A\gamma(\mathcal{T}_{h-1}^i(f)X_{h-1}), e_l\rangle B_h^{(l)} e_l \\
&+ \sum_{l=1}^{N} \langle q_{h-1}, e_i\rangle \big] a_{li} f(K_h) B_h^{(l)} e_l \\
=& B_h A\gamma(\mathcal{T}_{h-1}^i(f)X_{h-1}) + f(K_h)\langle q_{h-1}, e_i\rangle B_h A e_i.
\end{aligned}$$

□

Now, we apply Theorem 3.2 to get the recursive equations for $\gamma(\mathcal{O}_h^i X_h)$, $\gamma(\mathcal{T}_h^i(K_h)X_h)$ and $\gamma(\mathcal{T}_h^i(K_h^2)X_h)$.

Taking $f(K_h) \equiv 1$ in (3.24), we get:

Corollary 3.1.

$$\gamma(\mathcal{O}_h^i X_h) = B_h A\gamma(\mathcal{O}_{h-1}^i(f)X_{h-1}) + \langle q_{h-1}, e_i\rangle B_h A e_i, \tag{3.25}$$

with initial condition $\gamma(\mathcal{O}_0^i X_0) = \underline{0}$.

Applying Theorem 3.2 with $f(K_h) \equiv K_h$ and then with $f(K_h) \equiv K_h^2$, we obtain the recursive equations for $\gamma(\mathcal{T}_h^i(K_h)X_h)$ and $\gamma(\mathcal{T}_h^i(K_h^2)X_h)$:

Corollary 3.2.

$$\gamma(\mathcal{T}_h^i(K_h)X_h) = B_h A\gamma(\mathcal{T}_{h-1}^i(K_h)X_{h-1}) + K_h\langle q_{h-1}, e_i\rangle B_h A e_i, \tag{3.26}$$

$$\gamma(\mathcal{T}_h^i(K_h^2)X_h) = B_h A\gamma(\mathcal{T}_{h-1}^i(K_h^2)X_{h-1}) + K_h^2\langle q_{h-1}, e_i\rangle B_h A e_i, \tag{3.27}$$

with the initial conditions $\gamma(\mathcal{T}_0^i(K_h)X_0) = \gamma(\mathcal{T}_0^i(K_h^2)X_0) = \underline{0}$.

3.2 Modelling a clearness index process on a time interval

3.2.1 CTM-k model

We now describe CTM-k, a continuous-time HMM $(X_t, k_t)_{t\in[0,T]}$ for modelling a stochastic process of clearness index $(k_t)_{t\in[0,T]}$ during a continuous-time interval of a solar day.

We define the *state process* $(X_t)_{t\in[0,T]}$ as a continuous-time homogeneous Markov chain (Section 2.8): the state space is again $\mathbf{S} = \{e_1, e_2, \ldots, e_N\}$, where e_i denotes the unit N-vector with 1 at the i-th position, and the transition rate matrix $A^{\Delta} = (\lambda_{ji}) \in \mathbf{R}^{N\times N}$ with $\lambda_{ii} = -\sum_{j\neq i} \lambda_{ji}$ for $i = 1, 2, \ldots, N$.

We define the *observation process* (k_t), k_t representing the random clearness index at time $t \in [0, T]$, as a solution of the following SDE:

$$dk_t = \langle X_t, c\rangle dt + \langle X_t, \beta\rangle dW_t, \quad 0 \le t \le T, \tag{3.28}$$

where (W_t) is a standard Brownian motion, and $c = (c_1, c_2, \ldots, c_N)'$, $\beta = (\beta_1, \beta_2, \ldots, \beta_N)'$.

Our model CTM-k (X_t, k_t) depends on the parameter vector:

$$\theta = (\lambda_{ji}, 1 \le j \ne i \le N;\ c_1, c_2, \ldots, c_N;\ \beta_1, \beta_2, \ldots, \beta_N).$$

Let $\mathcal{G}_t^k = \sigma\{X_s, k_s; 0 \le s \le t\}$ and $\mathcal{Y}_t^k = \sigma\{k_s; 0 \le s \le t\}$ denote the filtrations of the complete data and the incomplete data, respectively.

3.2.2 Change of measure

For each parameter vector $\theta \in \Theta$, first consider the change of probability measure from the initial measure P_θ in the real world $(\Omega, (\mathcal{F}_t), P_\theta)$ to the reference probability measure $\overline{P}$ by putting

$$\left.\frac{d\overline{P}}{dP_\theta}\right|_{\mathcal{G}_t^k} = \overline{\Lambda}_t^{k,\theta} = \exp\left\{-\int_0^t \frac{\langle X_s, c\rangle}{\langle X_s, \beta\rangle} dW_s - \frac{1}{2}\int_0^t \left[\frac{\langle X_s, c\rangle}{\langle X_s, \beta\rangle}\right]^2 ds\right\}. \tag{3.29}$$

By Ito formula, equation (3.28) has the equivalent form:

$$d\left(\frac{k_t}{\langle X_t, \beta\rangle}\right) = \frac{\langle X_t, c\rangle}{\langle X_t, \beta\rangle} dt + dW_t. \tag{3.30}$$

By (3.29), (3.30) and Girsanov theorem, the process $\left(\frac{k_t}{\langle X_t, \beta\rangle}\right)$ is a standard Brownian motion under $\overline{P}$.

We now consider the reverse change, from $\overline{P}$ to P_θ, by putting:

$$\left.\frac{dP_\theta}{d\overline{P}}\right|_{\mathcal{G}_t^k} = \Lambda_t^{k,\theta} = \exp\left\{\int_0^t \frac{\langle X_s, c\rangle}{\langle X_s, \beta\rangle} d\left(\frac{k_t}{\langle X_t, \beta\rangle}\right) - \frac{1}{2}\int_0^t \left[\frac{\langle X_s, c\rangle}{\langle X_s, \beta\rangle}\right]^2 ds\right\}. \tag{3.31}$$

Then, Ito formula yields

$$\Lambda_t^{k,\theta} = 1 + \int_0^t \Lambda_s^\theta \frac{\langle X_s, c\rangle}{\langle X_s, \beta\rangle} d\left(\frac{k_s}{\langle X_s, \beta\rangle}\right), \tag{3.32}$$

and we have $\overline{E}(\Lambda_t^{k,\theta}) = 1$.

With this back change, Girsanov theorem implies that under P_θ, $\tilde{W}_t$ is a standard Brownian motion if we define

$$d\tilde{W}_t = d\left(\frac{k_t}{\langle X_t, \beta\rangle}\right) - \frac{\langle X_t, c\rangle}{\langle X_t, \beta\rangle} dt. \tag{3.33}$$

Thus, under P_θ, (k_t) satisfies the following equation equivalent to (3.28), due to Ito formula and (3.33):

$$dk_t = \langle X_t, c\rangle dt + \langle X_t, \beta\rangle d\tilde{W}_t, \tag{3.34}$$

where $(\tilde{W}_t)$ is a standard Brownian motion.

That is, under P_θ, (k_t) satisfies the real world dynamics. However, as $\overline{P}$ is a more convenient measure, we will work on the reference probability space $(\Omega, \mathcal{F}_t, \overline{P})$ to compute the unnormalized filters $\gamma(H_t) = \overline{E}(\Lambda_t^{k,\theta} H_t|\mathcal{Y}_t^k)$, where (H_t) is a $\mathcal{G}_t^k$-adopted scalar process (for example $H_t \equiv \mathcal{J}_t^{ij}, \mathcal{O}_t^i, \dots$). These results will be related to (and used in) the real world $(\Omega, (\mathcal{F}_t), P_\theta)$ by using a version of conditional Bayes Theorem:

$$E_\theta(H_t|\mathcal{Y}_t^k) = \frac{\overline{E}(\Lambda_t^{k,\theta} H_t|\mathcal{Y}_t^k)}{\overline{E}(\Lambda_t^{k,\theta}|\mathcal{Y}_t^k)}. \tag{3.35}$$

A complete proof for (3.35) can be found in [Elliott 1995, Theorem 3.2]. We here consider in cases where $\overline{E}(\Lambda_t^{k,\theta}|\mathcal{Y}_t^k) > 0$. Suppose B is any set in $G = \{\omega : \overline{E}(\Lambda_t^{k,\theta}|\mathcal{Y}_t^k) > 0)\} \subset \mathcal{Y}_t^k$, we must show

$$\int_B \frac{\overline{E}(\Lambda_t^{k,\theta} H_t|\mathcal{Y}_t^k)}{\overline{E}(\Lambda_t^{k,\theta}|\mathcal{Y}_t^k)} dP_\theta = \int_B E_\theta(H_t|\mathcal{Y}_t^k) dP_\theta.$$

Proof. By noting $\overline{E}\left[\overline{E}\left(\xi|\mathcal{Y}_t^k\right)\right] = \overline{E}(\xi)$ for any ξ, we have:

$$\begin{aligned}
\int_B \frac{\overline{E}(\Lambda_t^{k,\theta} H_t|\mathcal{Y}_t^k)}{\overline{E}(\Lambda_t^{k,\theta}|\mathcal{Y}_t^k)} dP_\theta =& E_\theta\left[\mathbf{1}_B \frac{\overline{E}(\Lambda_t^{k,\theta} H_t|\mathcal{Y}_t^k)}{\overline{E}(\Lambda_t^{k,\theta}|\mathcal{Y}_t^k)}\right] \\
=& \overline{E}\left[\mathbf{1}_B \Lambda_t^{k,\theta} \frac{\overline{E}(\Lambda_t^{k,\theta} H_t|\mathcal{Y}_t^k)}{\overline{E}(\Lambda_t^{k,\theta}|\mathcal{Y}_t^k)}\right] \\
=& \overline{E}\left[\overline{E}\left[\mathbf{1}_B \Lambda_t^{k,\theta} \frac{\overline{E}(\Lambda_t^{k,\theta} H_t|\mathcal{Y}_t^k)}{\overline{E}(\Lambda_t^{k,\theta}|\mathcal{Y}_t^k)}|\mathcal{Y}_t^k\right]\right] \\
=& \overline{E}\left[\mathbf{1}_B \overline{E}[\Lambda_t^{k,\theta}|\mathcal{Y}_t^k] \frac{\overline{E}(\Lambda_t^{k,\theta} H_t|\mathcal{Y}_t^k)}{\overline{E}(\Lambda_t^{k,\theta}|\mathcal{Y}_t^k)}\right] \\
=& \overline{E}\left[\mathbf{1}_B \overline{E}(\Lambda_t^{k,\theta} H_t|\mathcal{Y}_t^k)\right] \\
=& \overline{E}\left[\mathbf{1}_B \Lambda_t^{k,\theta} H_t\right] \\
=& \int_B \Lambda_t^{k,\theta} H_t d\overline{P} \\
=& \int_B H_t dP_\theta \\
=& \int_B E_\theta(H_t|\mathcal{Y}_t^k) dP_\theta.
\end{aligned}$$

□

Remark 3.1. *By Ito formula, the process in* (3.29) *can be determined by*

$$\overline{\Lambda}_t^{k,\theta} = 1 - \int_0^t \overline{\Lambda}_s^{k,\theta} \frac{\langle X_s, c\rangle}{\langle X_s, \beta\rangle} dW_s, \tag{3.36}$$

and we have $\Lambda_t^{k,\theta}\overline{\Lambda}_t^{k,\theta} = 1$.

3.2.3 Parameter estimation

At each iteration of EM algorithm, the new transition probabilities will be updated by [Elliott 1995, chap. 3, pg. 68-70] and [James 1996]:

$$\lambda'_{ji} = \frac{\pi(\mathcal{J}_t^{ji})}{\pi(\mathcal{O}_t^i)}, \quad 1 \le j \ne i \le N, \tag{3.37}$$

where $\pi(\mathcal{J}_t^{ji})$ and $\pi(\mathcal{O}_t^i)$ are the normalized filter (see Section 2.8.2).

Because measures corresponding to Wiener processes with different variances are not absolutely continuous [Lipster 2010, James 1996], we cannot obtain the ML estimates of parameters $\beta_1, \beta_2, \dots, \beta_N$ in continuous time. However, we get around this problem using discretization by introducing a Discrete Time Approximate Model (DTAM) in which we can use densities with respect to the Lebesgue measure and compute Radon-Nikodym derivatives of observation processes with different noisy variances.

We now use the EM algorithm mentioned in Section 2.9 to update the ML estimates of the parameter $c = (c_1, c_2, \dots, c_N)'$.

3.2.3.1 Expectation step

We consider the pseudo log-likelihood function defined as:

$$Q(\theta', \theta) = E_\theta\left(\log \Lambda_t^{k,\theta\theta'} | \mathcal{Y}_t^k\right),$$

where

$$\begin{aligned}
\Lambda_t^{k,\theta\theta'} = \left.\frac{dP_{\theta'}}{dP_\theta}\right|_{\mathcal{G}_t^k} &= \exp\left\{\int_0^t \frac{\langle X_s, c'\rangle}{\langle X_s, \beta'\rangle^2} dk_s - \frac{1}{2}\int_0^t \left(\frac{\langle X_s, c'\rangle}{\langle X_s, \beta'\rangle}\right)^2 ds\right\} \\
&\quad \times \left[\exp\left\{\int_0^t \frac{\langle X_s, c\rangle}{\langle X_s, \beta\rangle^2} dk_s - \frac{1}{2}\int_0^t \left(\frac{\langle X_s, c\rangle}{\langle X_s, \beta\rangle}\right)^2 ds\right\}\right]^{-1} \\
&= \exp\left\{\int_0^t \left[\frac{\langle X_s, c'\rangle}{\langle X_s, \beta'\rangle^2} - \frac{\langle X_s, c\rangle}{\langle X_s, \beta\rangle^2}\right] dk_s\right. \\
&\quad \left. - \frac{1}{2}\int_0^t \left[\left(\frac{\langle X_s, c'\rangle}{\langle X_s, \beta'\rangle}\right)^2 - \left(\frac{\langle X_s, c\rangle}{\langle X_s, \beta\rangle}\right)^2\right] ds\right\}.
\end{aligned}$$

Hence

$$\begin{aligned}Q(\theta',\theta) =&E_\theta\Bigg\{\int_0^t\left[\frac{\langle X_s,c'\rangle}{\langle X_s,\beta'\rangle^2}-\frac{\langle X_s,c\rangle}{\langle X_s,\beta\rangle^2}\right]dk_s\\&-\frac{1}{2}\int_0^t\left[\left(\frac{\langle X_s,c'\rangle}{\langle X_s,\beta'\rangle}\right)^2-\left(\frac{\langle X_s,c\rangle}{\langle X_s,\beta\rangle}\right)^2\right]ds\Bigg|\mathcal{Y}_t^k\Bigg\} \qquad (3.38)\\ =&E_\theta\Bigg\{\sum_{i=1}^N\frac{1}{\beta_i'^2}\Bigg[c_i'\int_0^t\langle X_s,e_i\rangle dk_s\\&-\frac{1}{2}c_i'^2\int_0^t\langle X_s,e_i\rangle ds\Bigg]\Bigg|\mathcal{K}_t\Bigg\}+R(\theta,\mathcal{Y}_t^k), \qquad (3.39)\end{aligned}$$

where $R(\theta,\mathcal{Y}_t^k)$ does not depend on θ'.

Then, by setting $\theta=\theta'^{(p)}$ at the p-th iteration of EM algorithm, we get

$$\begin{aligned}Q(\theta',\theta'^{(p)}) =&E_{\theta'^{(p)}}\Bigg\{\sum_{i=1}^N\frac{1}{\beta_i'^2}\Bigg[c_i'\int_0^t\langle X_s,e_i\rangle dk_s\\&-\frac{1}{2}c_i'^2\int_0^t\langle X_s,e_i\rangle ds\Bigg]\Bigg|\mathcal{Y}_t^k\Bigg\}+R(\theta'^{(p)},\mathcal{Y}_t^k). \qquad (3.40)\end{aligned}$$

Using the notation $\pi(H_t)=E_{\theta'^{(p)}}(H_t|\mathcal{Y}_t^k)$, for $H_t\equiv\mathcal{O}_t^i$ or $\mathcal{T}_t^i$ with

$$\mathcal{O}_t^i=\int_0^t\langle X_s,e_i\rangle ds,\ \mathcal{T}_t^i=\int_0^t\langle X_s,e_i\rangle dk_s,\quad i=1,2,\dots,N, \qquad (3.41)$$

we rewrite (3.40) as:

$$Q(\theta',\theta'^{(p)})=\sum_{i=1}^N\frac{1}{\beta_i'^2}\left[c_i'\pi(\mathcal{T}_t^i)-\frac{1}{2}c_i'^2\pi(\mathcal{O}_t^i)\right]+R(\theta'^{(p)},\mathcal{Y}_t^k). \qquad (3.42)$$

3.2.3.2 Maximization step

From (3.42), $\frac{\partial}{\partial c_i'}Q(\theta',\theta'^{(p)})=0$ becomes

$$c_i'=\frac{\pi(\mathcal{T}_t^i)}{\pi(\mathcal{O}_t^i)},\quad i=1,2,\dots,N. \qquad (3.43)$$

Observing that $\pi(H_t)=\frac{\gamma(H_t)}{\gamma(1)}=\frac{\langle\gamma(X_tH_t),\underline{1}\rangle}{\langle\gamma(X_t),\underline{1}\rangle}$ (for $H_t\equiv\mathcal{J}_t'^{ji}$, $\mathcal{O}_t^i$ or $\mathcal{T}_t^i$), update from λ_{ji} to λ'_{ji} in (3.37) and update from c_i to c_i' in (3.43) up to time T are now given by

$$\lambda'_{ji}=\frac{\langle\gamma(X_T\mathcal{J}_T^{ji}),\underline{1}\rangle}{\langle\gamma(X_T\mathcal{O}_T^i),\underline{1}\rangle},\quad 1\le j\ne i\le N, \qquad (3.44)$$

$$c_i'=\frac{\langle\gamma(X_T\mathcal{T}_T^i),\underline{1}\rangle}{\langle\gamma(X_T\mathcal{O}_T^i),\underline{1}\rangle},\quad 1\le i\le N, \qquad (3.45)$$

where the filter processes $\gamma(X_T\mathcal{O}_T^i)$, $\gamma(X_T\mathcal{T}_T^i)$ and $\gamma(X_T\mathcal{J}_T^{ji})$ are obtained from Theorem 3.3 and Theorem 3.4 in the next section.

3.2.4 Filtering equations

Let the *state process* q_t be defined as

$$q_t = \gamma(X_t) = \overline{E}(X_t\Lambda_t^{k,\theta}|\mathcal{Y}_t^k)$$

and let

$$C = \mathrm{diag}\,(\frac{c_1}{\beta_1}, \frac{c_2}{\beta_2}, \dots, \frac{c_N}{\beta_N}).$$

Theorem 3.3. *Let $\mathcal{O}_t^i$ and $\mathcal{T}_t^i$ be respectively the occupation time and the level sum defined in (3.41). The recursive equations of filtering processes $\gamma(X_t\mathcal{O}_t^i)$, $\gamma(X_t\mathcal{T}_t^i)$ and q_t are given by*

$$q_t = \pi_0 + \int_0^t A^\Delta q_s ds + \int_0^t C q_s \frac{dk_s}{\langle X_t, \beta\rangle}, \tag{3.46}$$

$$\gamma(X_t\mathcal{O}_t^i) = \int_0^t A^\Delta \gamma(X_s\mathcal{O}_s^i) ds + \int_0^t C\gamma(X_s\mathcal{O}_s^i)\frac{dk_s}{\langle X_s, \beta\rangle} + \int_0^t \langle q_s, e_i\rangle e_i ds, \tag{3.47}$$

$$\begin{aligned}\gamma(X_t\mathcal{T}_t^i) = &\int_0^t A^\Delta \gamma(X_s\mathcal{T}_s^i) ds + \int_0^t C\gamma(X_s\mathcal{T}_s^i)\frac{dk_s}{\langle X_s, \beta\rangle} \\ &+ \int_0^t c_i\langle q_s, e_i\rangle e_i ds + \int_0^t \langle q_s, e_i\rangle e_i dk_s.\end{aligned} \tag{3.48}$$

Proof. Firstly, remark that the filtering equations of Theorem 3.3 which involvethe process X_t are $\mathcal{Y}^k$-measurable, $\mathcal{Y}^k = \sigma\{k_s, 0 \le s \le t\}$, because the observation process (k_t) is determined by a SDE depending on X_t:

$$dk_t = \langle X_t, c\rangle dt + \langle X_t, \beta\rangle dW_t >$$

Consider a general scalar process defined as:

$$H_t = H_0 + \int_0^t \rho_s ds + \int_0^t \delta_s dW_s, \tag{3.49}$$

where ρ_t and δ_t are $\mathcal{G}_t$-predictable and square-integrable processes.

Ito product rule for semimartingales yields:

$$X_t H_t = X_0 H_0 + \int_0^t X_s\rho_s ds + \int_0^t X_s\delta_s dW_s + \int_0^t H_s A^\Delta X_s ds + \int_0^t H_s dV_s, \tag{3.50}$$

in which it should be noticed that $[X, H]_t = X_0 H_0$ because $\Delta X_s \Delta H_s = 0$ a.s.

Again Ito product rule applied to $X_t H_t$ and $\Lambda_t^{k,\theta}$ yields:

$$\begin{aligned}X_t H_t \Lambda_t^{k,\theta} = &X_0 H_0 + \int_0^t X_s\rho_s\Lambda_s^{k,\theta} ds + \int_0^t X_s\delta_s\Lambda_s^{k,\theta} dW_s + \int_0^t H_s A^\Delta X_s \Lambda_s^{k,\theta} ds \\ &+ \int_0^t H_s\Lambda_s^{k,\theta} dV_s + \int_0^t \Lambda_s^{k,\theta}\frac{\langle X_s, c\rangle}{\langle X_s, \beta\rangle} X_s H_s \frac{dk_s}{\langle X_s, \beta\rangle} + [XH, \Lambda^{k,\theta}]_t,\end{aligned} \tag{3.51}$$

where

$$[XH, \Lambda^{k,\theta}]_t = \int_0^t \Lambda_s^{k,\theta}\frac{\langle X_s, c\rangle}{\langle X_s, \beta\rangle} X_s\delta_s ds. \tag{3.52}$$

As $\langle X_s, e_i\rangle X_s = \langle X_s, e_i\rangle e_i$, we have

$$\int_0^t \Lambda_s^{k,\theta}\frac{\langle X_s, c\rangle}{\langle X_s,\beta\rangle}X_sH_s\frac{dk_s}{\langle X_s,\beta\rangle} = \sum_{i=1}^N\int_0^t\langle\Lambda_s^{k,\theta}X_sH_s, e_i\rangle\frac{c_i}{\beta_i}\frac{dk_s}{\beta_i}e_i, \tag{3.53}$$

$$\begin{aligned}\int_0^t X_s\delta_s\Lambda_s^{k,\theta}dB_s + \int_0^t\Lambda_s^{k,\theta}\frac{\langle X_s, c\rangle}{\langle X_s,\beta\rangle}X_s\delta_s ds &= \int_0^t X_s\delta_s\Lambda_s^{k,\theta}\frac{dk_s}{\langle X_s,\beta\rangle}\\ &= \sum_{i=1}^N\int_0^t\langle\Lambda_s^{k,\theta}\delta_sX_s, e_i\rangle\frac{dk_s}{\beta_i}e_i. \end{aligned}\tag{3.54}$$

Plugging (3.52), (3.53) and (3.54) into (3.51) we get

$$\begin{aligned}X_tH_t\Lambda_t^{k,\theta} =& X_0H_0 + \int_0^t X_s\rho_s\Lambda_s^{k,\theta}ds + \sum_{i=1}^N\int_0^t\langle\Lambda_s^{k,\theta}\delta_sX_s, e_i\rangle\frac{dk_s}{\beta_i}e_i\\ &+ \int_0^t H_sA^\Delta X_s\Lambda_s^{k,\theta}ds + \int_0^t H_s\Lambda_s^{k,\theta}dV_s + \sum_{i=1}^N\int_0^t\langle\Lambda_s^{k,\theta}X_sH_s, e_i\rangle\frac{c_i}{\beta_i}\frac{dk_s}{\beta_i}e_i.\end{aligned}\tag{3.55}$$

Appliying conditional expectation $\overline{E}$ given $\mathcal{Y}_t^k$ to both sides of (3.55) and noticing that $\overline{E}(dV_s|\mathcal{Y}_t^k) = 0$ ([Elliott 2010, Lemma 2.1, page 198]), it follows that

$$\begin{aligned}\gamma(X_tH_t) =& \gamma(X_0H_0) + \int_0^t\gamma(X_s\rho_s)ds + \sum_{i=1}^N\int_0^t\langle\gamma(\delta_sX_s), e_i\rangle\frac{dk_s}{\beta_i}e_i\\ &+ \int_0^t A^\Delta\gamma(X_sH_s)ds + \sum_{i=1}^N\int_0^t\langle\gamma(X_sH_s), e_i\rangle\frac{c_i}{\beta_i}\frac{dk_s}{\beta_i}e_i\\ =& \gamma(X_0H_0) + \int_0^t\gamma(X_s\rho_s)ds + \int_0^t\gamma(\delta_sX_s)\frac{dk_s}{\langle X_s,\beta\rangle}\\ &+ \int_0^t A^\Delta\gamma(X_sH_s)ds + \int_0^t C\gamma(X_sH_s)\frac{dk_s}{\langle X_s,\beta\rangle}.\end{aligned}\tag{3.56}$$

We now apply (3.56) to obtain the recursive equations of q_t, $\mathcal{O}_i^t$ and $\mathcal{T}_t^i$:

1. Taking $H_t = H_0 = 1$, $\rho_s = \delta_s = 0$, we obtain the recursive equation for the state process $q_t = \gamma(X_t) = \overline{E}(X_t\Lambda_t^\theta|\mathcal{Y}_t^k)$:

$$q_t = \pi_0 + \int_0^t A^\Delta q_sds + \int_0^t Cq_s\frac{dk_s}{\langle X_s,\beta\rangle}. \tag{3.57}$$

2. For the occupation time $H_t \equiv \mathcal{O}_t^i = \int_0^t\langle X_s, e_i\rangle ds$, $i = 1, 2, \ldots, N$, we take $H_t = \mathcal{O}_t^i$, $H_0 = 0$, $\rho_s = \langle X_s, e_i\rangle$, $\delta_s = 0$ and $X_t\rho_s = X_s\langle X_s, e_i\rangle = \langle X_s, e_i\rangle e_i$. Applying (3.56), we obtain

$$\gamma(X_t\mathcal{O}_t^i) = \int_0^t A^\Delta\gamma(X_s\mathcal{O}_s^i)ds + \int_0^t C\gamma(X_s\mathcal{O}_s^i)\frac{dk_s}{\langle X_s,\beta\rangle} + \int_0^t\langle q_s, e_i\rangle e_ids. \tag{3.58}$$

3. For the level sum $\mathcal{T}_t^i = \int_0^t \langle X_s, e_i \rangle dk_s$, $i = 1, 2, \ldots, N$, we first observe that

$$\mathcal{T}_t^i = \int_0^t c_i \langle X_s, e_i \rangle ds + \int_0^t \beta_i \langle X_s, e_i \rangle dB_s. \tag{3.59}$$

Then, taking $H_t = \mathcal{T}_t^i$, $H_0 = 0$, $\rho_s = c_i \langle X_s, e_i \rangle$, $\delta_s = \beta_i \langle X_s, e_i \rangle$ and applying (3.56), we get

$$\begin{aligned} \gamma(X_t \mathcal{T}_t^i) &= \int_0^t A^\Delta \gamma(X_s \mathcal{T}_s^i) ds + \int_0^t C\gamma(X_s \mathcal{T}_s^i) \frac{dk_s}{\langle X_s, \beta \rangle} \\ &+ \int_0^t c_i \langle q_s, e_i \rangle e_i ds + \int_0^t \langle q_s, e_i \rangle e_i dk_s. \end{aligned} \tag{3.60}$$

□

Theorem 3.4. *Let $\mathcal{J}_t^{ij} = \int_0^t \langle X_s, e_i \rangle \langle dX_s, e_j \rangle$ be the number of jumps from state e_i to state e_j, $1 \leq i \neq j \leq N$, $t \in [0, T]$. The filtering process $\gamma(\mathcal{J}_t^{ij})$ is obtained from the following recursive equation*

$$\gamma(X_t \mathcal{J}_t^{ij}) = \int_0^t A^\Delta \gamma(X_s \mathcal{J}_s^{ij}) ds + \int_0^t C\gamma(X_s \mathcal{J}_s^{ij}) \frac{dk_s}{\langle X_s, \beta \rangle} + \int_0^t \langle q_s, e_j \rangle \lambda_{ji} e_i ds, \tag{3.61}$$

where $C = \operatorname{diag}\left(\frac{c_1}{\beta_1}, \frac{c_2}{\beta_2}, \ldots, \frac{c_N}{\beta_N}\right)$.

Proof. Using $dX_s = AX_s ds + dV_s$, we first decompose

$$\begin{aligned} \mathcal{J}_t^{ij} &= \int_0^t \langle X_s, e_i \rangle \langle dX_s, e_j \rangle = \int_0^t \langle X_s, e_i \rangle e_j' dX_s \\ &= \int_0^t \langle X_s, e_i \rangle a_{ij} ds + \int_0^t \langle X_s, e_i \rangle e_j' dV_s. \end{aligned} \tag{3.62}$$

Applying Itô product rule and noticing that $X_s \langle X_s, e_i \rangle = \langle X_s, e_i \rangle e_i$, we have

$$\begin{aligned} X_t \mathcal{J}_t^{ij} &= \int_0^t \langle X_s, e_i \rangle a_{ji} ds e_i + \int_0^t \langle X_s, e_i \rangle e_j' dV_s e_i \\ &+ \int_0^t AX_s \mathcal{J}_s^{ij} ds + \int_0^t \mathcal{J}_s^{ij} dV_s + [X, \mathcal{J}^{ij}]_t, \end{aligned} \tag{3.63}$$

where $[X, \mathcal{J}^{ij}]_t$ is the quadratic variation computed as follows

$$
\begin{aligned}
[X, \mathcal{J}^{ij}]_t &= \sum_{0\le s\le t} \Delta \mathcal{J}^{ij}_s \Delta X_s = \sum_{0\le s\le t} (\langle X_s, e_i\rangle e_j' \Delta X_s)(\Delta X_s) \\
&= \int_0^t \left(\langle X_s, e_i\rangle e_j' dX_s - \langle X_s, e_i\rangle e_j' X_{s^-} \right) (dX_s - X_{s^-}) \\
&= \sum_{l,m=1}^N \int_0^t \left(\langle X_s, e_i\rangle e_j' e_l - \langle X_s, e_i\rangle e_j' e_m \right) \langle X_{s^-}, e_m\rangle \langle dX_s, e_l\rangle (e_l - e_m) \\
&= \sum_{l,m=1}^N \int_0^t \left(\langle X_s, e_i\rangle e_j' e_l - \langle X_s, e_i\rangle e_j' e_m \right) \langle X_{s^-}, e_m\rangle a_{lm} ds (e_l - e_m) \\
&\quad + \sum_{l,m=1}^N \int_0^t \left(\langle X_s, e_i\rangle e_j' e_l - \langle X_s, e_i\rangle e_j' e_m \right) \langle X_{s^-}, e_m\rangle \langle dV_s, e_l\rangle (e_l - e_m) \\
&= \sum_{l,m=1}^N \int_0^t \langle \langle X_s, e_i\rangle e_j' e_l e_i - \langle X_s, e_i\rangle e_j' e_m e_i, e_m\rangle a_{lm} ds (e_l - e_m) \\
&\quad + \sum_{l,m=1}^N \int_0^t \langle \langle X_s, e_i\rangle e_j' e_l e_i - \langle X_s, e_i\rangle e_j' e_m e_i, e_m\rangle \langle dV_s, e_l\rangle (e_l - e_m).
\end{aligned}
$$

Plugging into (3.63), the product $X_t \mathcal{J}^{ij}_t$ becomes:

$$
\begin{aligned}
X_t \mathcal{J}^{ij}_t &= \int_0^t \langle X_s, e_i\rangle a_{ji} ds e_i + \int_0^t \langle X_s, e_i\rangle e_j' dV_s e_i + \int_0^t A X_s \mathcal{J}^{ij}_s ds + \int_0^t \mathcal{J}^{ij}_s dV_s \\
&\quad + \sum_{l,m=1}^N \int_0^t \langle \langle X_s, e_i\rangle e_j' e_l e_i - \langle X_s, e_i\rangle e_j' e_m e_i, e_m\rangle a_{lm} ds (e_l - e_m) \\
&\quad + \sum_{l,m=1}^N \int_0^t \langle \langle X_s, e_i\rangle e_j' e_l e_i - \langle X_s, e_i\rangle e_j' e_m e_i, e_m\rangle \langle dV_s, e_l\rangle (e_l - e_m).
\end{aligned}
$$

Applying Ito product rule to $X_t \mathcal{J}^{ij}_t$ and $\Lambda^{k,\theta}_t$, we obtain

$$
\begin{aligned}
X_t \mathcal{J}^{ij}_t \Lambda^{k,\theta}_t &= \int_0^t \langle X_s \Lambda^{k,\theta}_s, e_i\rangle a_{ji} ds e_i + \int_0^t \langle X_s \Lambda^{k,\theta}_s, e_i\rangle e_j' dV_s e_i \\
&\quad + \int_0^t A^{\Delta} X_s \mathcal{J}^{ji}_s \Lambda^{k,\theta}_s ds + \int_0^t \mathcal{J}^{ij}_s \Lambda^{k,\theta}_s dV_s \\
&\quad + \sum_{l,m=1}^N \int_0^t \langle \langle X_s \Lambda^{k,\theta}_s, e_i\rangle e_j' e_l e_i - \langle X_s \Lambda^{k,\theta}_s, e_i\rangle e_j' e_m e_i, e_m\rangle a_{lm} ds (e_l - e_m) \\
&\quad + \sum_{l,m=1}^N \int_0^t \langle \langle X_s \Lambda^{k,\theta}_s, e_i\rangle e_j' e_l e_i - \langle X_s \Lambda^{k,\theta}_s, e_i\rangle e_j' e_m e_i, e_m\rangle \langle dV_s, e_l\rangle (e_l - e_m) \\
&\quad + \sum_{i=1}^N \int_0^t \langle X_s \Lambda^{k,\theta}_s, e_i\rangle \mathcal{J}^{ij}_s \frac{c_i}{\beta_i} \frac{dk_s}{\beta_i}.
\end{aligned}
\tag{3.64}
$$

The recursive equation of $\gamma(X_t\mathcal{J}_t^{ij})$ is then obtained taking conditional expectation $\overline{E}$ given $\mathcal{Y}_t^k$ of both sides of (3.64):

$$\begin{aligned}
\gamma(X_t\mathcal{J}_t^{ij}) &= \int_0^t \langle q_s, e_i\rangle a_{ji} ds e_i + \int_0^t A^\Delta \gamma(X_s\mathcal{J}_s^{ij}) ds \\
&\quad + \sum_{l,m=1}^N \int_0^t \langle\langle q_s, e_i\rangle e_j' e_l e_i - \langle q_s, e_i\rangle e_j' e_m e_i, e_m\rangle a_{lm} ds(e_l - e_m) \\
&\quad + \int_0^t C\gamma(X_s\mathcal{J}_s^{ij}) \frac{dk_s}{\langle X_s\beta\rangle} \\
&= \int_0^t \langle q_s, e_i\rangle a_{ji} ds e_i + \int_0^t A^\Delta \gamma(X_s\mathcal{J}_s^{ij}) ds + \int_0^t C\gamma(X_s\mathcal{J}_s^{ij}) \frac{dk_s}{\langle X_s\beta\rangle} \\
&\quad + \sum_{l,m=1}^N \int_0^t \langle\langle q_s, e_i\rangle e_j' e_l e_i - \langle q_s, e_i\rangle e_j' e_m e_i, e_m\rangle a_{lm} ds e_l \\
&\quad - \sum_{l,m=1}^N \int_0^t \langle\langle q_s, e_i\rangle e_j' e_l e_i - \langle q_s, e_i\rangle e_j' e_m e_i, e_m\rangle a_{lm} ds e_m \\
&= \int_0^t \langle q_s, e_i\rangle a_{ji} ds e_i + \int_0^t A^\Delta \gamma(X_s\mathcal{J}_s^{ij}) ds + \int_0^t C\gamma(X_s\mathcal{J}_s^{ij}) \frac{dk_s}{\langle X_s\beta\rangle} \\
&\quad \int_0^t \langle q_s, e_i\rangle a_{ji} ds e_j - \int_0^t \langle q_s, e_i\rangle a_{ji} ds e_i.
\end{aligned}$$

□

3.3 Discrete-Time Approximate Model DTAM-k

Since the data that we deal with have been sampled at regular time intervals, the observation range $[0, T]$ is split into a regular partition

$$0 = t_0 < t_1 < t_2 < \cdots < t_{h-1} < t_h < \ldots t_M = T, \tag{3.65}$$

where $\Delta = t_h - t_{h-1}$ is such that $M\Delta = T$ and M is the size of the observed sample $k_1, k_2, \ldots, k_M$ used to estimate the model parameter vector.

3.3.1 Components of DTAM-k

Define the following discrete-time observations

$$u_h = \frac{1}{\Delta}(k_h - k_{h-1}), \quad h = 1, 2, \ldots.$$

From (3.28) u_h verifies:

$$u_h = \langle X_h, c\rangle + \frac{1}{\sqrt{\Delta}}\langle X_h, \beta\rangle w_h, \quad h = 1, 2 \ldots, M, \tag{3.66}$$

where $c = (c_1, c_2, \ldots, c_N)'$, $\beta = (\beta_1, \beta_2, \ldots, \beta_N)'$ and $w_h \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$, $h = 1, 2 \ldots.$

Equation (3.66) determines an observation sequence $(u_h)_{h=1,2,...}$ which is driven by state process $(X_h)_{h=0,1,2,...}$, a discrete time Markov chain with state space $\mathbf{S} = \{e_1, e_2, \ldots, e_N\}$ and transition probability matrix:

$$P = I + \Delta A^{\Delta} = (\epsilon_{ji}) \in \mathbf{R}^{N \times N}.$$

Note that, for a small enough time step Δ, P is a stochastic matrix.

The parameter vector of this model is

$$\theta = (\epsilon_{ji}, 1 \leq j \neq j \leq N; c_1, c_2, \ldots, c_N; \beta_1, \beta_2, \ldots, \beta_N).$$

For $i = 1, 2, \ldots, N$, let us define the level sums of DTAM-k by

$$\mathcal{T}_h^i(u) \triangleq \sum_{l=1}^{h} \langle X_l, e_i \rangle u_l,$$
$$\mathcal{T}_h^i(u^2) \triangleq \sum_{l=1}^{h} \langle X_l, e_i \rangle u_l^2.$$

Let $\mathcal{Y}_h^u = \sigma(u_1, u_2, \ldots, u_h)$ denote the filtration of incomplete data. We will use the following filter processes to obtain ML estimates of the model parameters:

$$\tilde{\pi}(H_h) \triangleq E_\theta(H_h | \mathcal{Y}_h^u),$$
$$\tilde{\gamma}(H_h) \triangleq \overline{E}(H_h \Lambda_h^{u,\theta} | \mathcal{Y}_h^u),$$
$$\tilde{\gamma}(X_h H_h) \triangleq \overline{E}(X_h H_h \Lambda_h^{u,\theta} | \mathcal{Y}_h^u),$$

where $H_h \equiv \mathcal{J}_h^{ij}, \mathcal{O}_h^i, \mathcal{T}_h^i(u)$ or $\mathcal{T}_h^i(u^2)$.

3.3.2 Discrete-time approximate filtering equations

Our purpose here is to approximate the continuous-time filtering equations in order that the resulting equations are locally Lipschitz continuous given the observation path $\{k_1, k_2, \ldots, k_M\}$.

3.3.2.1 Approximation of state filter equation

For studying the approximate scheme of the state filtering equation, a particular case of the Duncan-Mortenson-Zakai equation, we follow Clark's *robust filter* approach [Clark 1977]. We first introduce the following processes:

$$\phi_t^{(i)} = \exp\left\{ \int_0^t \frac{c_i}{\beta_i} \frac{dk_s}{\beta_i} - \frac{1}{2} \int_0^t \left(\frac{c_i}{\beta_i} \right)^2 ds \right\}, i = 1, 2, \ldots, N, \tag{3.67}$$

$$\Phi_t = \text{diag}\left(\phi_t^{(1)}, \phi_t^{(2)}, \ldots, \phi_t^{(N)} \right), \tag{3.68}$$

$$\overline{q}_t = \Phi_t^{-1} q_t. \tag{3.69}$$

The process $\overline{q}_t$ is called the *robust filter of the state process* q_t. We will use this process to provide an approximate equation for the state equation (3.46).

Applying Ito product rule to the two processes

$$(\phi_t^{(i)})^{-1} = 1 - \int_0^t (\phi_t^{(i)})^{-1} \frac{c_i}{\beta_i} dW_s, \text{ by (3.31) and (3.36)},$$
$$q_t^{(i)} = \pi_0^{(i)} + \int_0^t (\lambda_{i1} q_s^{(1)} + \cdots + \lambda_{iN} q_s^{(N)}) ds + \int_0^t q_s^{(i)} \frac{c_i}{\beta_i} \frac{dk_s}{\beta_i}, \text{ from (3.46)},$$

for $i = 1, 2, \ldots, N$, we get

$$(\phi_t^{(i)})^{-1} q_t^{(i)} = \pi_0^{(i)} + \int_0^t (\phi_s^{(i)})^{-1} (\lambda_{i1} q_s^{(1)} + \cdots + \lambda_{iN} q_s^{(N)}) ds \tag{3.70}$$

and

$$\overline{q}_t = \pi_0 + \int_0^t (\Phi_s)^{-1} A^{\Delta} q_s ds. \tag{3.71}$$

From (3.71), we can determine $(\overline{q}_t)$ by the following ordinary differential equation:

$$\frac{d}{dt}\overline{q} = (\Phi_t)^{-1} A^{\Delta} \Phi_t \overline{q}_t, \tag{3.72}$$

with initial condition $\overline{q}_0 = \pi_0$.

Integrating (3.72) between sampling times t_{h-1} and t_h of the partition (3.65) yields

$$\overline{q}_{t_h} = \overline{q}_{t_{h-1}} + \int_{t_{h-1}}^{t_h} (\Phi_s)^{-1} A^{\Delta} \Phi_s \overline{q}_s ds, \tag{3.73}$$

and the Euler-Maruyama method [Higham 2001] shows that

$$\overline{q}_{t_h} \simeq \overline{q}_{t_{h-1}} + (\Phi_{t_{h-1}})^{-1} A^{\Delta} \Phi_{t_{h-1}} \overline{q}_{t_{h-1}} \Delta. \tag{3.74}$$

Multiplying both sides of (3.74) by Φ_{t_h}, we get

$$q_{t_h} \simeq \Phi_{t_h} (\Phi_{t_{h-1}})^{-1} [\mathbf{I} + \Delta A^{\Delta}] q_{t_{h-1}} = \Phi_{t_h} (\Phi_{t_{h-1}})^{-1} P q_{t_{h-1}}, \tag{3.75}$$

where $P = \mathbf{I} + \Delta A^{\Delta}$, with $\mathbf{I}$ is the $N \times N$ unit matrix..

We now consider the product $\Phi_{t_h} (\Phi_{t_{h-1}})^{-1}$.

The definition of $\phi_t^{(i)}$ in (3.67) implies that

$$\phi_{t_h}^{(i)} (\phi_{t_{h-1}}^{(i)})^{-1} = \exp \left\{ \int_{t_{h-1}}^{t_h} \frac{c_i}{\beta_i^2} dk_s - \frac{1}{2} \int_{h-1}^{t_h} \left(\frac{c_i}{\beta_i} \right)^2 ds, \right\}, \tag{3.76}$$

for $i = 1, 2, \ldots, N$.

As the partition (3.65) is such that $\Delta = t_h - t_{h-1}$, $k_h - k_{h-1} = \Delta u_h$, (3.76) becomes

$$\phi_{t_h}^{(i)} (\phi_{t_{h-1}}^{(i)})^{-1} = \exp \left\{ \frac{\Delta}{2\beta_i^2} (2 c_i u_h - c_i^2) \right\}, \quad i = 1, 2, \ldots, N. \tag{3.77}$$

We then have

$$\Phi_{t_h} (\Phi_{t_{h-1}})^{-1} = \text{diag} \left(\tilde{B}_h^1, \tilde{B}_h^2, \ldots \tilde{B}_h^N \right) = \tilde{B}_h, \tag{3.78}$$

where $\tilde{B}_h^{(i)} = \exp\left\{\frac{\Delta}{2\beta_i^2}(2c_i u_h - c_i^2)\right\}$, for $i = 1, 2, \ldots, N$.

Our numerical approximation of q_{t_h} will be denoted by $\tilde{q}_h$. From (3.75) and (3.78), an approximation of the Duncan-Mortenson-Zakai equation is now given by

$$\tilde{q}_h = \tilde{B}_h P \tilde{q}_{h-1}, \quad \tilde{q}_0 = \pi_0. \tag{3.79}$$

Clark [Clark 1977] has showed that the robust filter $(\overline{q}_t)$ determined by $\overline{q}_t = \Phi_t^{-1} q_t$ is a locally Lipschitz continuous function given the observed path $\{k_s, 0 \leq s \leq t\}$. Therefore, applying (3.79), it is seen that the process $(\tilde{q}_m)$ determined by equation (3.79) can be used to approximate the state filter (q_t) which inherits this continuity property.

3.3.2.2 Approximate filter equation of the number of jumps, of the occupation time and of the level sums

Using similar arguments as those used to obtain a robust filter of the state process q_s, robust versions $\overline{\gamma}(X_t H_t)$ of the filters $\gamma(X_t H_t)$ for the processes $H_t \equiv \mathcal{J}_t^{ij}, \mathcal{O}_t^i, \mathcal{T}_t^i$ can be obtained from

$$\begin{aligned}
\frac{d}{dt}\overline{\gamma}(X_t \mathcal{J}_t^{ij}) =& \Phi_t^{-1} A^{\Delta} \gamma(X_t \mathcal{J}_t^{ij}) \Phi_t \overline{\gamma}(X_t \mathcal{J}_t^{ij}) \\
&+ \langle \overline{q}, e_j \rangle \langle \Phi_t^{-1} A^{\Delta} \gamma(X_t \mathcal{J}_t^{ij}) \Phi_t e_j, e_i \rangle e_i, \qquad (3.80)\\
\frac{d}{dt}\overline{\gamma}(X_t \mathcal{O}_t^{i}) =& \Phi_t^{-1} A^{\Delta} \gamma(X_t \mathcal{O}_t^{i}) \Phi_t \overline{\gamma}(X_t \mathcal{O}_t^{i}) + \langle \overline{q}, e_i \rangle e_i, \qquad (3.81)\\
d\overline{\gamma}(X_t \mathcal{T}_t^{i}) =& \Phi_t^{-1} A^{\Delta} \gamma(X_t \mathcal{T}_t^{i}) \Phi_t \overline{\gamma}(X_t \mathcal{T}_t^{i}) dt + \langle \overline{q}, e_i \rangle e_i dk_t, \qquad (3.82)
\end{aligned}$$

with initial conditions $\overline{\gamma}(X_0 \mathcal{J}_0^{ij}) = \overline{\gamma}(X_0 \mathcal{O}_0^{i}) = \overline{\gamma}(X_0 \mathcal{T}_0^{i}(u_0)) = \underline{0}$, $1 \leq i \neq j \leq N$.

For $H_t \equiv \mathcal{J}_t^{ij}, \mathcal{O}_t^i$ or $\mathcal{T}_t^i$, let $\gamma(X_{t_h} H_{t_h})$ be an approximation of continuous-time filter processes $\gamma(X_t H_t)$ between sampling times t_{h-1} and t_h in partition (**??**), using robust filtering equations (3.80), (3.81) and (3.82).

The filter processes $\gamma(X_{t_h} \mathcal{J}_{t_h}^{ji})$, $\gamma(X_{t_h} \mathcal{O}_{t_h}^{i})$ and $\gamma(X_{t_h} \mathcal{T}_{t_h}^{i})$ will be approximated in the DTAM-k model by $\Delta-$time filters $\tilde{\gamma}(X_h \mathcal{J}_h^{ji})$, $\tilde{\gamma}(X_h \mathcal{O}_h^{i})$ and $\tilde{\gamma}(X_h \mathcal{T}_h^{i}(u_h))$, respectively, where:

$$\begin{aligned}
\tilde{\gamma}(X_h \mathcal{J}_h^{ij}) &= \tilde{B}_h P \tilde{\gamma}(X_{h-1} \mathcal{J}_{h-1}^{ij}) + \langle \tilde{q}_{h-1}, e_j \rangle \langle \tilde{B}_h P e_j, e_i \rangle e_i, \qquad (3.83)\\
\tilde{\gamma}(X_h \mathcal{O}_h^{i}) &= \tilde{B}_h P \tilde{\gamma}(X_{h-1} \mathcal{O}_{h-1}^{i}) + \langle \tilde{q}_{h-1}, e_i \rangle \tilde{B}_h P e_i, \qquad (3.84)\\
\tilde{\gamma}(X_h \mathcal{T}_h^{i}(u_h)) &= \tilde{B}_h P \tilde{\gamma}(X_{h-1} \mathcal{T}_{h-1}^{i}(u_h)) + u_h \langle \tilde{q}_{h-1}, e_i \rangle \tilde{B}_h P e_i, \qquad (3.85)
\end{aligned}$$

with initial conditions $\tilde{\gamma}(X_0 \mathcal{J}_0^{ij}) = \tilde{\gamma}(X_0 \mathcal{O}_0^{i}) = \tilde{\gamma}(X_0 \mathcal{T}_0^{i}(u_0)) = \underline{0}$, $1 \leq i \neq j \leq N$.

The error between $\langle \tilde{\gamma}(X_h H_h), \underline{1} \rangle$ (for $H_h \equiv \mathcal{J}_h^{ij}$, $\mathcal{O}_h^i$ or $\mathcal{T}_h^i$, $h = 1, 2, \ldots, M$) and $\langle \gamma(X_{t_h} H_{t_h}), \underline{1} \rangle$ (for $H_{t_h} \equiv \mathcal{J}_{t_h}^{ij}$, $\mathcal{O}_{t_h}^i$ or $\mathcal{T}_{t_h}^i$, $0 \leq t_h \leq T$) is evaluated by [James 1996, Theorem 3.3 and Corollary 3.4]:

$$|\langle \gamma(X_{t_h} H_{t_h}), \underline{1} \rangle [k] - \langle \tilde{\gamma}(X_h H_h), \underline{1} \rangle [k]| \leq C_{\|k\|} |\Delta + \omega_{\Delta}(k)|, \tag{3.86}$$

where the constant $C_{\|k\|}$ depends on $\|k\| \triangleq \sup_{0 \leq t \leq T} |k_t|$ and

$$\omega_{\Delta}(k) = \max\left\{|k_t - k_s| : 0 \leq s, t \leq T, |t - s| \leq \Delta\right\}. \tag{3.87}$$

3.3.3 Updating parameter

EM algorithm is also used to obtain numerical ML estimates for DTAM-k. At each iteration of algorithm, ML estimates of transition probabilities are updated by

$$\epsilon'_{ji} = \frac{\langle \tilde{\gamma}(X_M \mathcal{J}_M^{ij}), \underline{1} \rangle}{\langle \tilde{\gamma}(X_M \mathcal{O}_M^{i}), \underline{1} \rangle}, \quad 1 \leq j \neq i \leq N. \tag{3.88}$$

For updating remaining parameters, we consider the following conditional expectation:

$$Q(\theta', \theta) = E_\theta \left\{ \log \left[\prod_{l=1}^{h} \frac{\frac{1}{\sqrt{\Delta}} \langle X_l, \sigma \rangle \phi \left(\frac{u_l - \langle X_l, c' \rangle}{\frac{1}{\sqrt{\Delta}} \langle X_l, \sigma' \rangle} \right)}{\frac{1}{\sqrt{\Delta}} \langle X_l, \sigma' \rangle \phi \left(\frac{u_l - \langle X_l, c \rangle}{\frac{1}{\sqrt{\Delta}} \langle X_l, \sigma \rangle} \right)} \right] \middle| \mathcal{Y}_h^u \right\}, \tag{3.89}$$

where $c' = (c'_1, c'_2, \ldots, c'_N)'$, $\beta' = (\beta'_1, \beta'_2, \ldots, \beta'_N)'$ and $\phi(\cdot)$ is the $\mathcal{N}(0,1)$ density function.

At the p-th iteration of EM algorithm, vector $\theta = \theta^{(p)}$ and transform $Q(\theta', \theta^{(p)})$ as follows

$$\begin{aligned} Q(\theta', \theta^{(p)}) = &\sum_{i=1}^{N} \left\{ \tilde{\pi}(\mathcal{O}_h^i) \log \frac{1}{\beta'_i} - \frac{\Delta}{\beta_i'^2} \left[\tilde{\pi}(\mathcal{T}_h^i(u_h^2)) - 2b'_i \tilde{\pi}(\mathcal{T}_h^i(u_h)) + b_i'^2 \tilde{\pi}(\mathcal{O}_h^i) \right] \right\} \\ &+ R(\theta'^{(p)}, \mathcal{Y}_h^u), \end{aligned}$$

where the function $R(\theta'^{(p)}, \mathcal{Y}_h^u)$ does not depend on θ'.

Maximizing the function $Q(\theta', \theta^{(p)})$ above yields

$$c'_i = \frac{\tilde{\pi}(\mathcal{T}_h^i(u))}{\tilde{\pi}(\mathcal{O}_h^i)}, \quad i = 1, 2, \ldots, N, \tag{3.90}$$

$$\beta_i'^2 = \frac{\Delta}{\tilde{\pi}(\mathcal{O}_M^i)} \left[\tilde{\pi}(\mathcal{T}_M^i(u^2)) - 2c'_i \tilde{\pi}(\mathcal{T}_M^i(u)) + c_i'^2 \tilde{\pi}(\mathcal{O}_M^i) \right]. \tag{3.91}$$

For $H_h \equiv \mathcal{J}_h^{ij}, \mathcal{O}_h^i$ or $\mathcal{T}_h^i(u)$, by plugging $\tilde{\pi}(H_h) = \frac{\tilde{\gamma}(H_h)}{\tilde{\gamma}(1)} = \frac{\langle \tilde{\gamma}(X_h H_h, \underline{1}) \rangle}{\langle \tilde{q}_h, \underline{1} \rangle}$, into (3.90), (3.91), we get:

$$c'_i = \frac{\langle \tilde{\gamma}(X_M \mathcal{T}_M^i(u_M)), \underline{1} \rangle}{\langle \tilde{\gamma}(X_M \mathcal{O}_M^i), \underline{1} \rangle}, \quad i = 1, 2, \ldots, N \tag{3.92}$$

$$\begin{aligned} \beta_i'^2 = &\Delta \left(\langle \tilde{\gamma}(X_M \mathcal{O}_M^i), \underline{1} \rangle \right)^{-1} \left[\langle \tilde{\gamma}(X_M \mathcal{T}_M^i(u^2)), \underline{1} \rangle \right. \\ &\left. - 2c'_i \langle \tilde{\gamma}(X_M \mathcal{T}_M^i(u)), \underline{1} \rangle + c_i'^2 \langle \tilde{\gamma}(X_M \mathcal{O}_M^i), \underline{1} \rangle \right], \end{aligned} \tag{3.93}$$

where the discrete-time filter process $\left(\tilde{\gamma}(X_h \mathcal{T}_h^i(u^2)) \right)_{h=1,2,\ldots,M}$ is determined by:

$$\tilde{\gamma}(X_h \mathcal{T}_h^i(u_h^2)) = \Psi_h P \tilde{\gamma}(X_{h-1} \mathcal{T}_{h-1}^i(u_h^2)) + (u_h)^2 \langle p_{h-1}, e_i \rangle \Psi_h P e_i, \tag{3.94}$$

with initial condition $\tilde{\gamma}(X_0 \mathcal{T}_0^i(u^2)) = \underline{0}$, $i = 1, 2, \ldots, N$.

3.4 Experiments with real data

The parameters will be estimated from the observed data, the number of states being chosen after examining the data histograms. We deal with data coming from some tropical humid areas, but our method can also be tested on other types of climate.

3.4.1 Real data

Using standard formulas of the extraterrestrial radiation (Chapter 1), we compute the CISs from the total solar radiation measurements performed in two French islands: Guadeloupe and La Réunion which are tropical areas (as mentioned in Section 1.6.2).

To apply DTM-K model, we have used one month length samples of daily clearness index (Table 3.1) from measurements sampled at 1-minute in La Réunion in 2009, 2010 and 2011. Measurements corresponding to March 2009 are denoted K-DATA-0309. Similar notations hold for other measurements.

We have applied DTAM-k model of CTM-k with total radiations sampled at 1−s (Section 1.6.2) in Guadeloupe in 2006 (see Table 3.3) by Dr. T. Soubdhan, University of Antilles-Guyane. Time interval $[0, T]$ with $T = 1$ hour was split using (3.65) with a small enough $\Delta = 1/3600$.

Month	2009		2010		2011	
	Average	Variance	Average	Variance	Average	Variance
March	0.4569	0.0372	0.4187	0.0484	0.5290	0.0108
April	0.2044	0.0394	0.5534	0.0089	0.4604	0.0304
May	0.3486	0.0467	0.5127	0.0172	0.4463	0.0353
August	0.4949	0.0080	0.5043	0.0140	0.5170	0.0074
September	0.5368	0.0116	0.4385	0.0269	0.5122	0.0180
October	0.5484	0.0113	0.5554	0.0171	0.4811	0.0268

Table 3.1: Average and variance of clearness index sequences in some months in 2009, 2010 and 2011, La Réunion - France.

Day type		K_h
I	Clear	$0.7 \leq K_h < 0.9$
II	Partially cloudy	$0.3 \leq K_h < 0.7$
III	Cloudy	$0.0 \leq K_h < 0.3$

Table 3.2: Classification of days according to clearness index (Table 4.2 in [Ianetz 2008, page 94]).

Iqbal 1983 proposed utilizing the magnitude of the daily clearness index K_h to define sky conditions, see Table 3.2. Note that there are several criterias for solar-day classification which depend on geographical location and climatic conditions. For expample, 4 classes were identified in Guadeloupe [Soubdhan 2009]: clear sky days, intermittent clear sky days, intermittent cloudy sky days and cloudy sky days while in La Réunion 5 classes were found [Delsaut 2013]: cloudy days, intermittent bad days, disturbed days, intermittent good days and clear sky days. For simplicity, we here use the classification according to Iqbal 1983.

3.4.2 Estimations

CTM-k will be considered by estimating DTAM-k parameters with $\Delta = 1/3600$, using1-s mean CISs: k-DATA-I.1, k-DATA-II.2 and k-DATA-III.3.

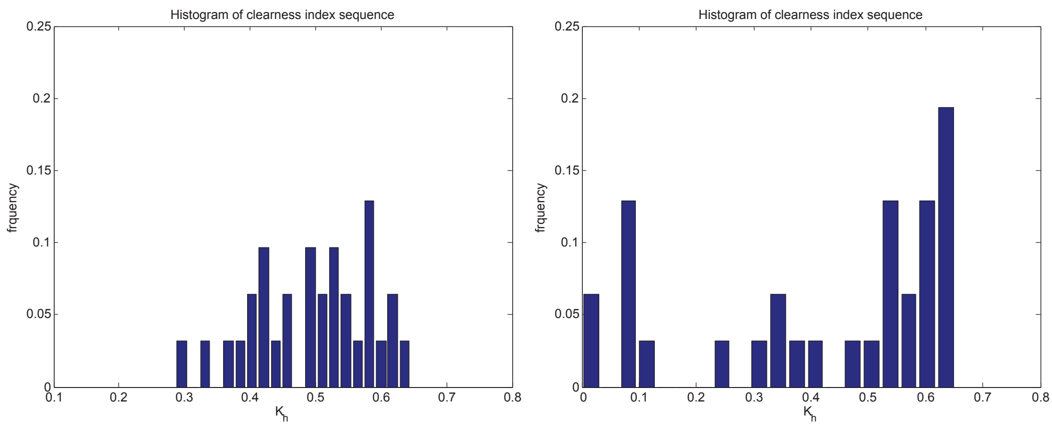
Observing the histograms of these data Figure 3.3, Figure 3.4 and Figure 3.5, we have decided that $N = 4$ states.

3.4.2.1 DTM-K parameter estimations

DTM-K parameter are estimated when using three daily CISs observed in August 2009 (K-DATA-0809, *variance* 0.0080, Figure 3.6(a)), March 2010 (K-DATA-0310, *variance* 0.0484, Figure 3.7(a)) and October 2011 (K-DATA-1011, *variance* 0.0268, Figure 3.8(a)).

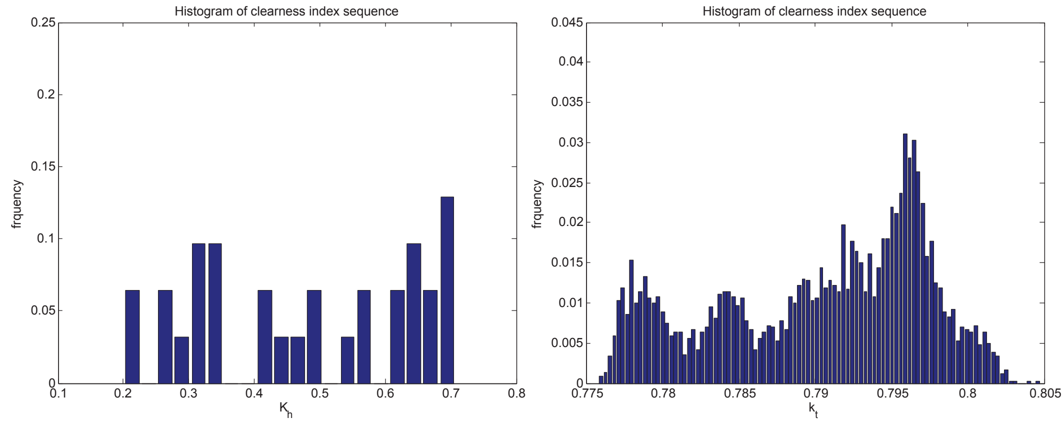
Data set	Mean of CIS	Variance of CIS	time $-$ dayth (in 2006)
DATA-I.1	0.7904	0.0069	$09h10h - 262$th
DATA-I.2	0.7536	0.0062	$13h14h - 109$th
DATA-I.3	0.7205	0.0097	$13h14h - 139$th
DATA-II.1	0.6162	0.2715	$09h10h - 118$th
DATA-II.2	0.4172	0.2234	$13h14h - 234$th
DATA-II.3	0.6178	0.1730	$09h10h - 234$th
DATA-III.1	0.2975	0.0308	$13h14h - 339$th
DATA-III.2	0.1052	0.0201	$09h10h - 285$th
DATA-III.3	0.1229	0.0269	$13h14h - 180$th

Table 3.3: Mean clearness index (k) sequences computed from measurements of total solar radiation sample at 1-s in Guadeloupe-France in 2006. DATA-I.1 stands for sequences belonging to type I in Table 3.2, and similarly for others.



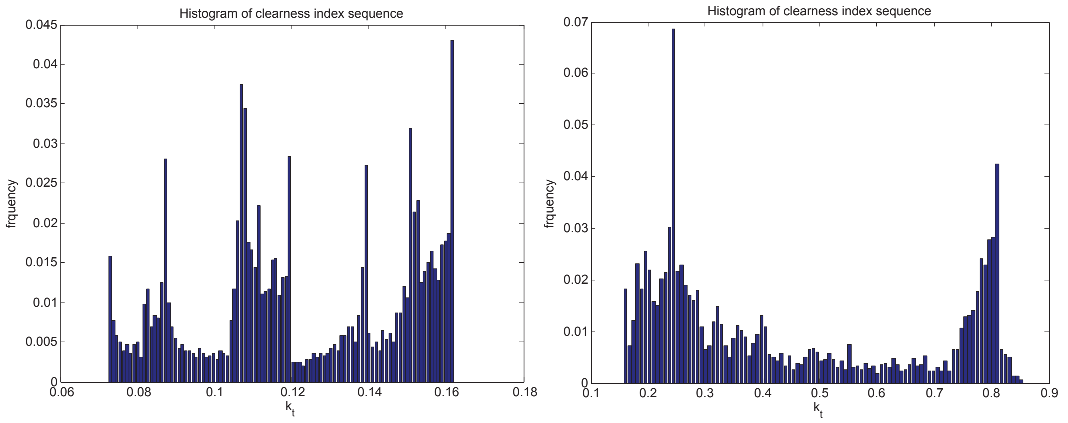
(a) (b)

Figure 3.3: Histogram of daily CIS measured in: (a) August 2009, (b) March 2010, La Réunion.



(a) (b)

Figure 3.4: Histogram of observed data: (a) daily CIS measured in October 2011, La Réunion; (b) k-DATA-I.1.



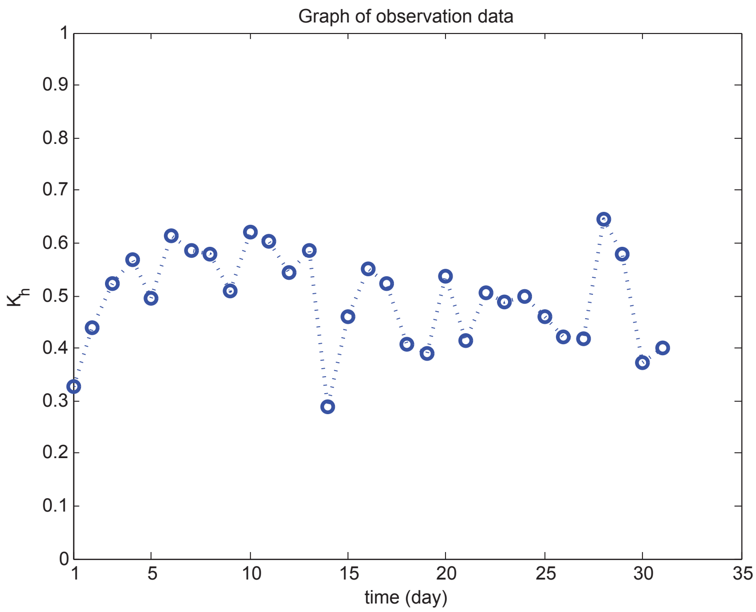
(a) (b)

Figure 3.5: Histogram of observed data: (a) k-DATA-III.3; (b) k-DATA-II.2.

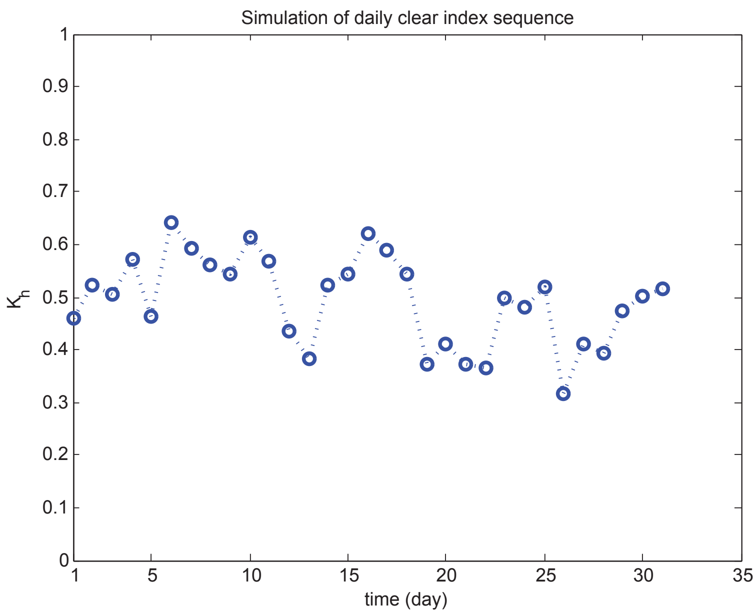
The histograms have incited to take $N = 4$ states, the parameter vector is $\theta = \{(a_{ji}), 1 \leq j \neq i \leq 4; b, \alpha\}$, with $b = (b_1, b_2, b_3, b_4)'$ and $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)'$. The transition probability matrix is

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}.$$

Note that $a_{ii} = 1 - \sum_{j \neq i} a_{ji}$, for $i = 1, 2, 3, 4$.

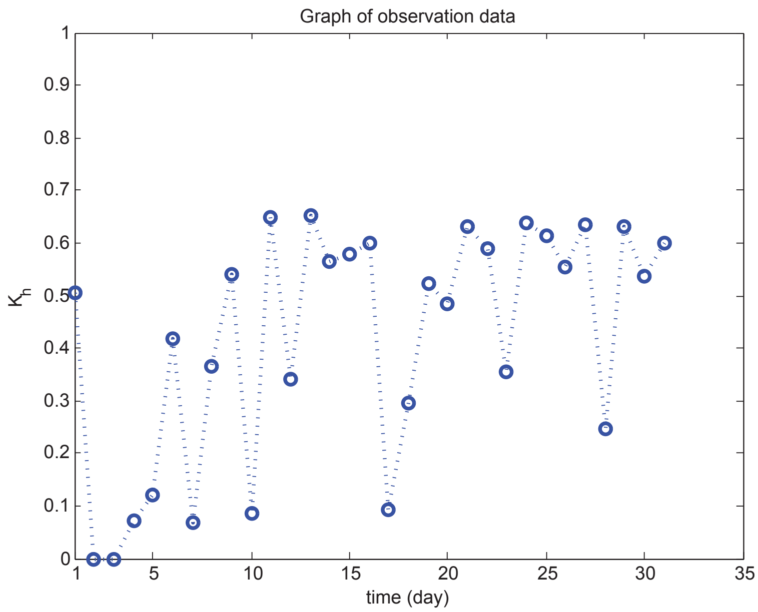


(a)

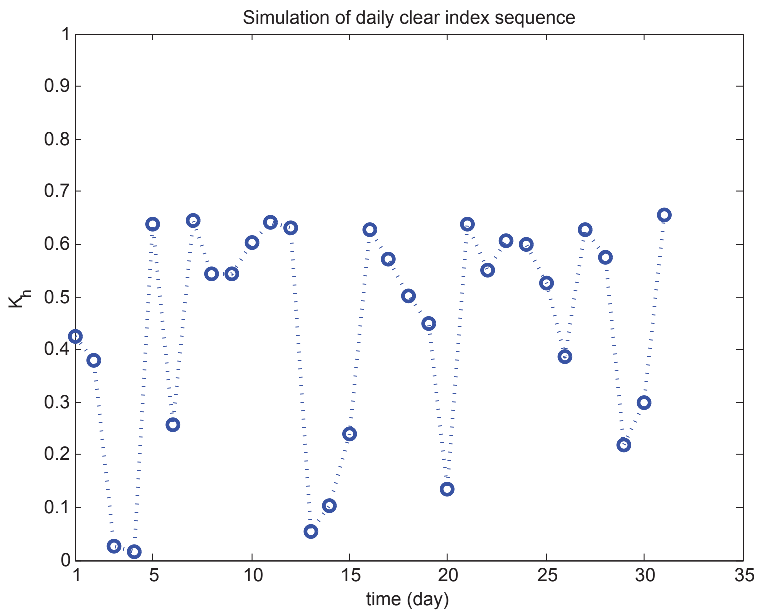


(b)

Figure 3.6: Graph of daily CIS measured in August 2011, La Réunion-France, (a), and its simulation, (b).

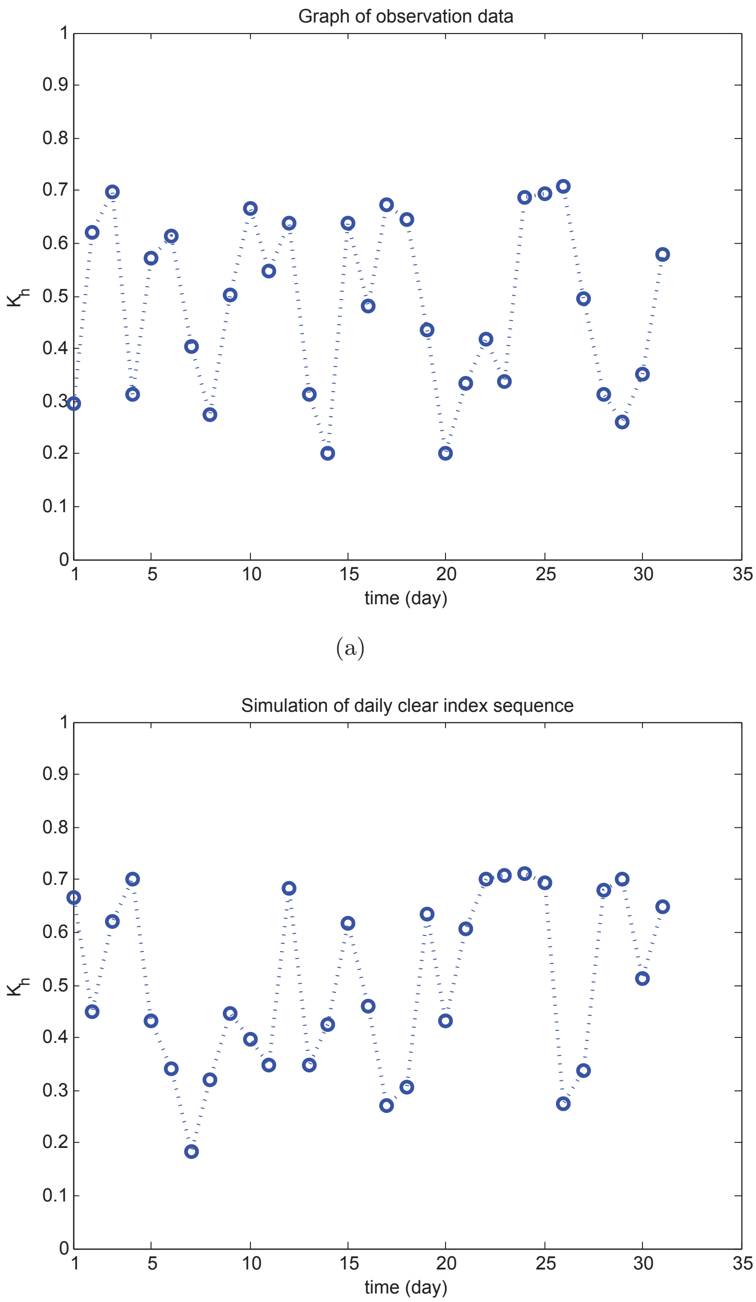


(a)



(b)

Figure 3.7: Graph of daily CIS measured in March 2010, La Réunion-France, (a), and its simulation, (b).



(a)

(b)

Figure 3.8: Graph of daily CIS measured in October 2011, La Réunion-France, (a), and its simulation, (b).

Initial parameter vector was:

$$
\begin{aligned}
A^{(0)} &= \begin{pmatrix} 0.3 & 0.3 & 0.2 & 0.3 \\ 0.4 & 0.4 & 0.4 & 0.4 \\ 0.2 & 0.1 & 0.3 & 0.1 \\ 0.1 & 0.2 & 0.1 & 0.2 \end{pmatrix}, \\
b^{(0)} &= (0.25, 0.5, 0.7, 0.8)', \\
\alpha^{(0)} &= (0.076, 0.076, 0.076, 0.076)'.
\end{aligned}
$$

After 50 iterations, we obtain the following estimates:

1. From K-DATA-0809:

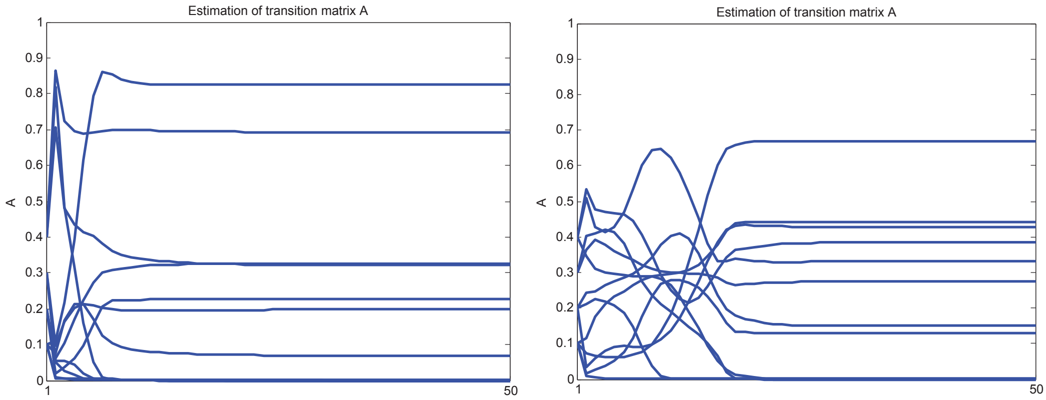
$$
\begin{aligned}
A^{(50)} &= \begin{pmatrix} 0.3075 & 0.1978 & 0.3271 & 0.0000 \\ 0.6925 & 0.5078 & 0.3210 & 0.0000 \\ 0.0000 & 0.0688 & 0.3519 & 0.8259 \\ 0.0000 & 0.2256 & 0.0000 & 0.1741 \end{pmatrix}, \\
b^{(50)} &= (0.3874, 0.4872, 0.5749, 0.6223)', \\
\beta^{(50)} &= (0.0586, 0.0473, 0.0167, 0.0154)'.
\end{aligned}
$$

2. From K-DATA-0310:

$$
\begin{aligned}
A^{(50)} &= \begin{pmatrix} 0.4274 & 0.2736 & 0.1504 & 0.0000 \\ 0.4420 & 0.2979 & 0.3835 & 0.3299 \\ 0.0000 & 0.0000 & 0.4660 & 0.6687 \\ 0.1306 & 0.4285 & 0.0001 & 0.0014 \end{pmatrix}, \\
b^{(50)} &= (0.0620, 0.4268, 0.5816, 0.6396)', \\
\alpha^{(50)} &= (0.0425, 0.1097, 0.0250, 0.0084)'.
\end{aligned}
$$

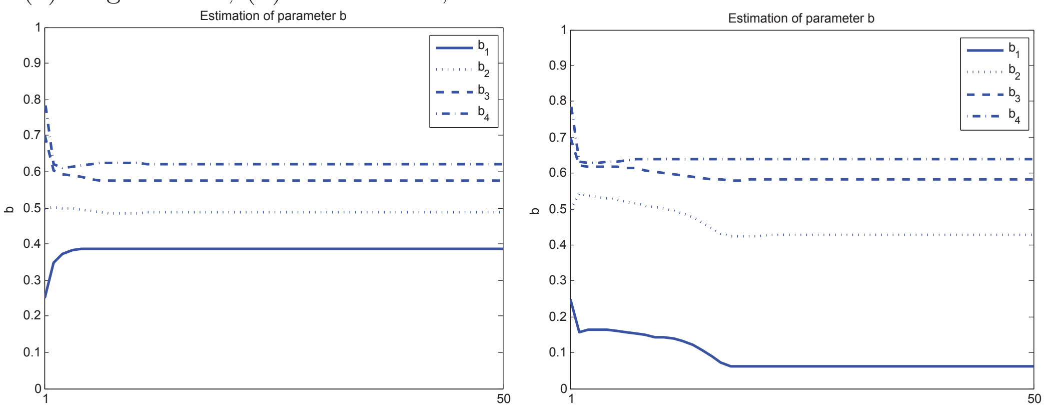
The graphs in Figure 3.9, Figure 3.10 and Figure 3.11 show the evolution of these estimates.

Models with their estimated parameters were used to simulated some paths:Figure 3.6(b) and Figure 3.7(b).



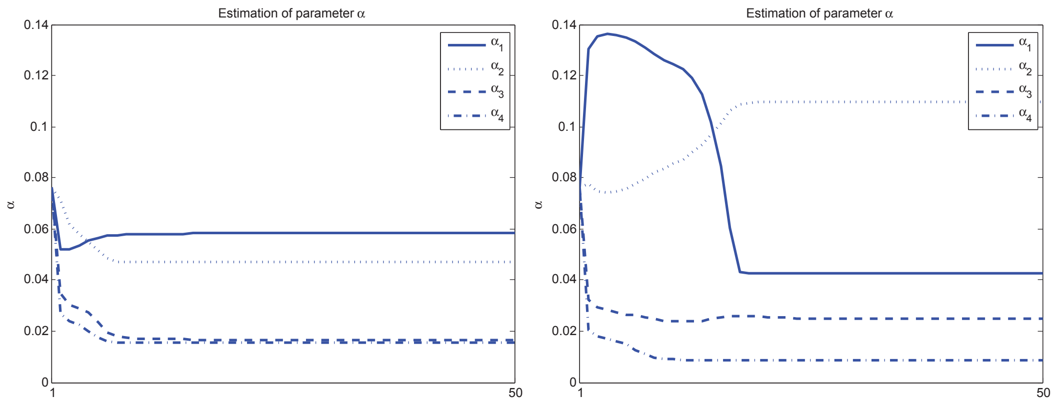
(a) (b)

Figure 3.9: Estimation of probability transition matrix A from daily CIS measured in: (a) August 2009, (b) March 2010, La Réunion-France.



(a) (b)

Figure 3.10: Estimation of parameter vector b from daily CIS measured in: (a) August 2009, (b) March 2010, La Réunion-France.



(a) (b)

Figure 3.11: Estimation of parameter vector α from daily CIS measured in: (a) August 2009, (b) March 2010, La Réunion-France.

Different limit estimations ($\theta^{1,*}$ and $\theta^{2,*}$) are obtained from the same observation data when using different initial parameters ($\theta^{(1,0)}$ and $\theta^{(2,0)}$, respectively). However, simulated data obtained from the models with these limit vectors do not have many differences. Simulated values generated by these models have the same distribution as observation data. The mean of simulated paths converge to the mean of observation data. These remarks confirm the properties of EM algorithm used in our estimatiion setting.

We checked this problem in several cases. For instance, to estimate the parameters of DTM-K with $N = 4$ states by using the daily CISs observed in October 2011 at La Réunion (Figure 3.8(a)), we started with two initial parameter vectors $\theta^{(1,0)}$, $\theta^{(2,0)}$ as follows:

The vector $\theta^{(1,0)}$ includes :

$$A^{(1,0)} = \begin{pmatrix} 0.3 & 0.3 & 0.2 & 0.3 \\ 0.4 & 0.4 & 0.4 & 0.4 \\ 0.2 & 0.1 & 0.3 & 0.1 \\ 0.1 & 0.2 & 0.1 & 0.2 \end{pmatrix},$$

$$b^{(1,0)} = (0.25, 0.5, 0.7, 0.8)',$$

$$\alpha^{(1,0)} = (0.076, 0.076, 0.076, 0.076)'.$$

The vector $\theta^{(2,0)}$ includes :

$$A^{(2,0)} = \begin{pmatrix} 0.18 & 0.08 & 0.10 & 0.20 \\ 0.21 & 0.11 & 0.60 & 0.20 \\ 0.31 & 0.20 & 0.19 & 0.40 \\ 0.30 & 0.61 & 0.11 & 0.40 \end{pmatrix},$$

$$b^{(2,0)} = (0.35, 0.45, 0.55, 0.65)',$$

$$\alpha^{(2,0)} = (0.1, 0.2, 0.3, 0.4)'.$$

Results after 100 iterations:

1. The limit vector $\theta^{(1,*)}$ obtained from $\theta^{(1,0)}$:

$$A^{(1,100)} = \begin{pmatrix} 0.2192 & 0.3714 & 0.0000 & 0.1646 \\ 0.4008 & 0.2282 & 0.5722 & 0.5197 \\ 0.3800 & 0.4005 & 0.1956 & 0.0000 \\ 0.0000 & 0.0000 & 0.2322 & 0.3158 \end{pmatrix},$$

$$b^{(1,100)} = (0.2793, 0.4312, 0.6392, 0.7004)',$$

$$\alpha^{(1,100)} = (0.0522, 0.0985, 0.0343, 0.0050)'.$$

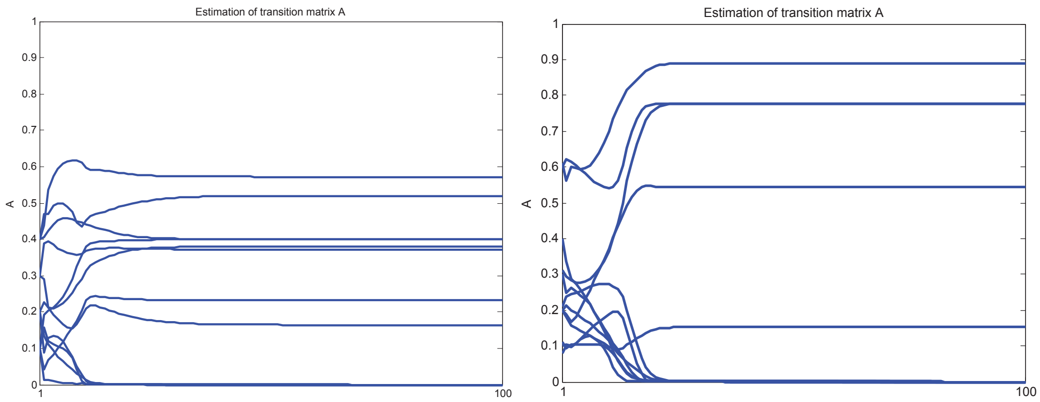
2. The limit vector $\theta^{(2,*)}$ obtained from $\theta^{(2,0)}$:

$$A^{(2,100)} = \begin{pmatrix} 0.2223 & 0.0000 & 0.0000 & 0.5443 \\ 0.0000 & 0.1095 & 0.7774 & 0.1528 \\ 0.7777 & 0.0000 & 0.2226 & 0.0000 \\ 0.0000 & 0.8905 & 0.0000 & 0.3029 \end{pmatrix},$$

$$b^{(2,100)} = (0.2860, 0.6459, 0.3818, 0.5567)',$$

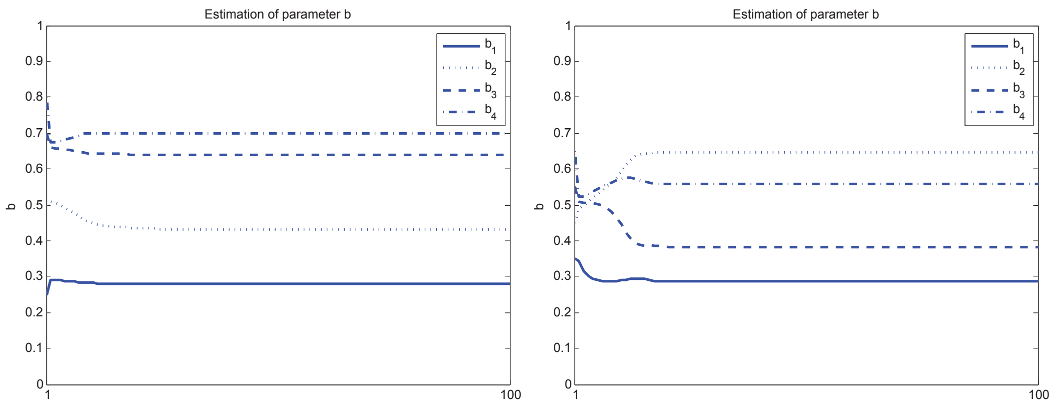
$$\alpha^{(2,100)} = (0.0428, 0.0381, 0.1173, 0.1139)'.$$

The graphs in Figure 3.12, Figure 3.13 and Figure 3.14 show the evolution of estimates from $\theta^{(1,0)}$ and $\theta^{(2,0)}$. Figure 3.8(b) shows a simulated path from the model with limit vector parameter $\theta^{(1,*)}$. Convergence of averages and estimated cumulative distribution function of simulated values generated by the models with these limit vector parameter $\theta^{(1,*)}$, $\theta^{(2,*)}$ are shown in Figure 3.15.



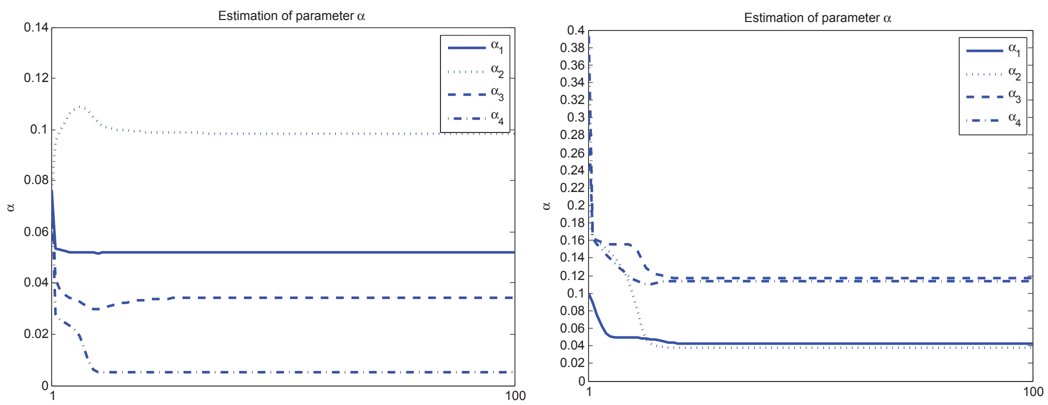
(a) (b)

Figure 3.12: Estimation of the transition matrix A from K-DATA-1011 and the initial parameter vector: (a) $\theta^{(1,0)}$; (b) $\theta^{(2,0)}$.



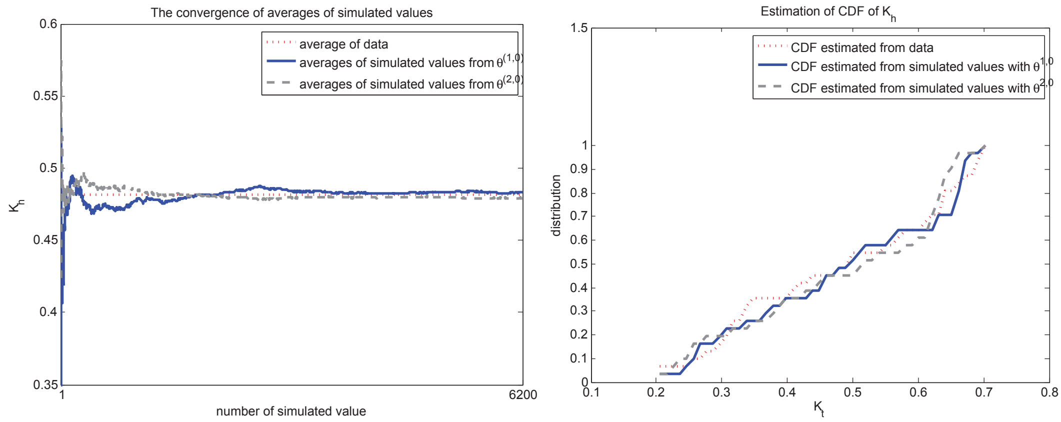
(a) (b)

Figure 3.13: Estimation of parameter vector b from K-DATA-1011 and the initial parameter vector: (a) $\theta^{(1,0)}$; (b) $\theta^{(2,0)}$.



(a) (b)

Figure 3.14: Estimation of parameter vector α from K-DATA-1011 and the initial parameter vector: (a) $\theta^{(1,0)}$; (b) $\theta^{(2,0)}$.



(a) (b)

Figure 3.15: (a) Convergence of average of simulated values; (b) Estimated CDF from simulated values.

3.4.2.2 Some illustrations for DTAM-k

By using the three 1-s mean CISs: k-DATA-I.1 (Figure 3.19a), k-DATA-II.2 (Figure 3.21a) and k-DATA-III.3 (Figure 3.20a), the parameter vector $\theta = \{A^{\Delta} = (\lambda_{ji}), 1 \le j \ne i \le 4; c, \beta\}$ of CTM-k will be approximated by estimated parameter vector $\theta = \{P = (\epsilon_{ji}), 1 \le j \ne i \le 4; c, \beta\}$ of its DTAM-k with $\Delta = 1/3600$, here $c = (c_1, c_2, c_3, c_4)'$, $\beta = (\beta_1, \beta_2, \beta_3, \beta_4)'$ and $P = \mathbf{I} + \Delta A^{\Delta}$.

For CTM-k, we start with

$$\begin{aligned} A^{\Delta(0)} &= \begin{pmatrix} -13 & 3 & 2 & 3 \\ 4 & -9 & 4 & 3 \\ 2 & 3 & -10 & 1 \\ 7 & 3 & 4 & -7 \end{pmatrix}, \\ c^{(0)} &= (0.25, 0.35, 0.55, 0.85)', \\ \beta^{(0)} &= (1, 2, 3, 4)'. \end{aligned}$$

then estimates of transition matrix in its DTAM-k will be started with:

$$P = \mathbf{I} + \Delta A^{\Delta} = \begin{pmatrix} 0.9964 & 0.0008 & 0.0006 & 0.0008 \\ 0.0011 & 0.9975 & 0.0011 & 0.0008 \\ 0.0006 & 0.0008 & 0.9972 & 0.0003 \\ 0.0019 & 0.0008 & 0.0011 & 0.9981 \end{pmatrix}.$$

After 50 iterations of EM algorithm, we obtain

1. From k-DATA-I.1:

$$
\begin{aligned}
P^{(50)} &= \begin{pmatrix} 0.9964 & 0.0008 & 0.0006 & 0.0008 \\ 0.0011 & 0.9975 & 0.0011 & 0.0008 \\ 0.0005 & 0.0008 & 0.9972 & 0.0003 \\ 0.0019 & 0.0008 & 0.0011 & 0.9981 \end{pmatrix}, \\
c^{(50)} &= (0.0252, 0.0252, 0.0253, 0.0252)', \\
\beta^{(50)} &= (0.0615, 0.0618, 0.0600, 0.0622)'.
\end{aligned}
$$

2. From k-DATA-II.2:

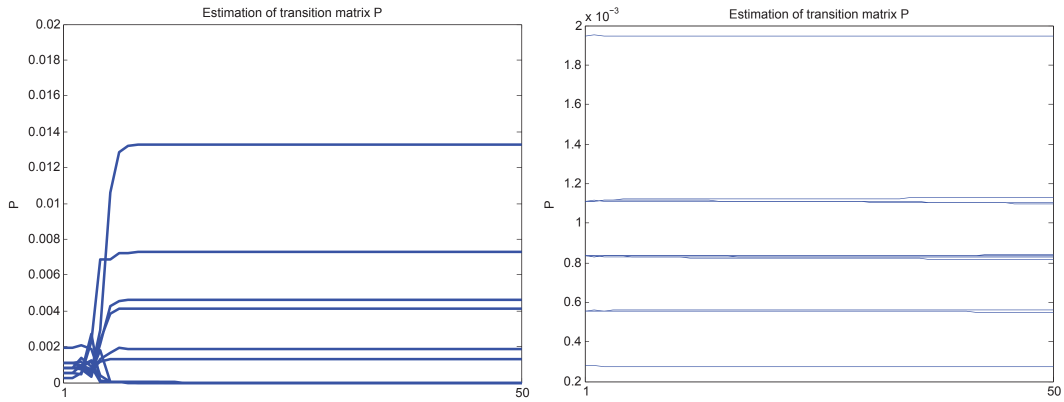
$$
\begin{aligned}
P^{(50)} &= \begin{pmatrix} 0.9964 & 0.0008 & 0.0006 & 0.0008 \\ 0.0011 & 0.9975 & 0.0011 & 0.0008 \\ 0.0006 & 0.0008 & 0.9972 & 0.0003 \\ 0.0019 & 0.0008 & 0.0011 & 0.9981 \end{pmatrix}, \\
c^{(50)} &= (-0.6084, -0.4740, -0.9695, -0.4315)', \\
\beta^{(50)} &= (1.3048, 1.3032, 1.3242, 1.2984)'.
\end{aligned}
$$

3. From k-DATA-III.3:

$$
\begin{aligned}
P^{(50)} &= \begin{pmatrix} 0.9941 & 0.0041 & 0.0073 & 0.0000 \\ 0.0046 & 0.9940 & 0.0000 & 0.0133 \\ 0.0013 & 0.0000 & 0.9927 & 0.0000 \\ 0.0000 & 0.0019 & 0.0000 & 0.9867 \end{pmatrix}, \\
c^{(50)} &= (0.2189, -0.1635, 1.1028, -1.0794)', \\
\beta^{(50)} &= (0.0068, 0.0068, 0.0096, 0.0100)'.
\end{aligned}
$$

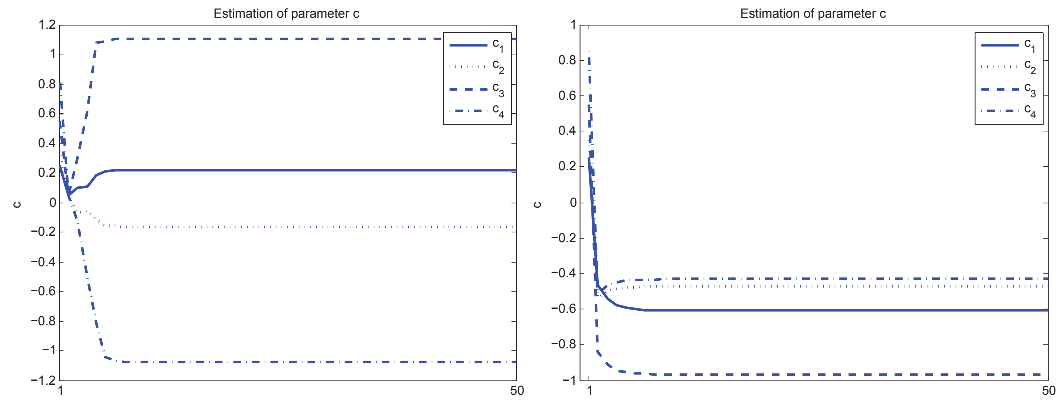
The graphs in Figure 3.16, Figure 3.17 and Figure 3.18 show the evolution of estimates from k-DATA-III.3 and k-DATA-II.2.

The models with estimated parameters are used to simulate some paths: Figure 3.19(b), Figure 3.20(b) and Figure 3.21(b).



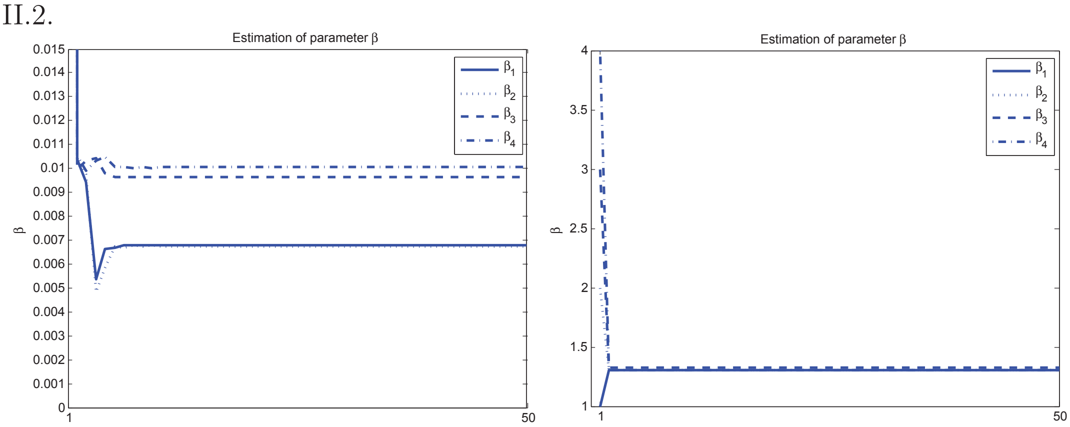
(a) (b)

Figure 3.16: Estimation of transition probability matrix P from: (a) k-DATA-III.3; (b) k-DATA-II.2.



(a) (b)

Figure 3.17: Estimation of parameter vector c from: (a)k-DATA-III.3; (b) k-DATA-II.2.



(a) (b)

Figure 3.18: Estimation of parameter vector β from: (a) k-DATA-III.3; (b) k-DATA-II.2.

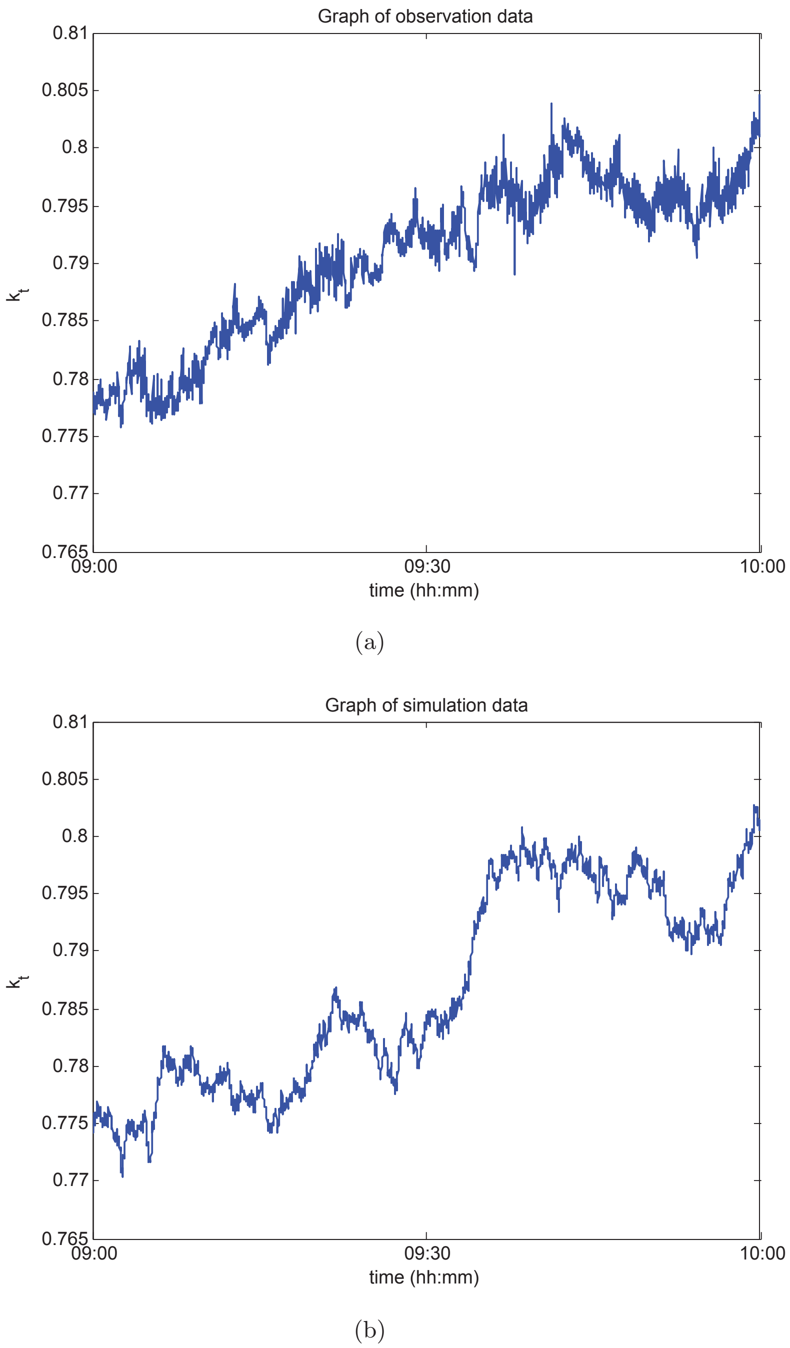
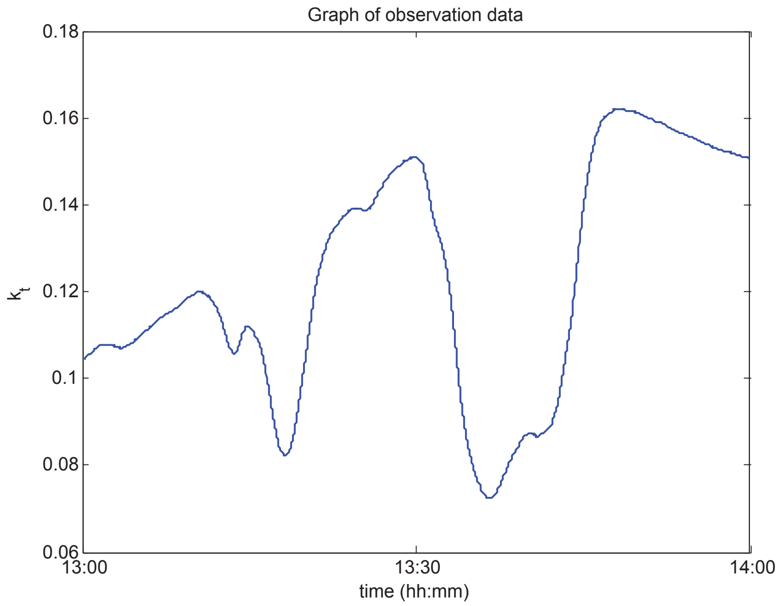
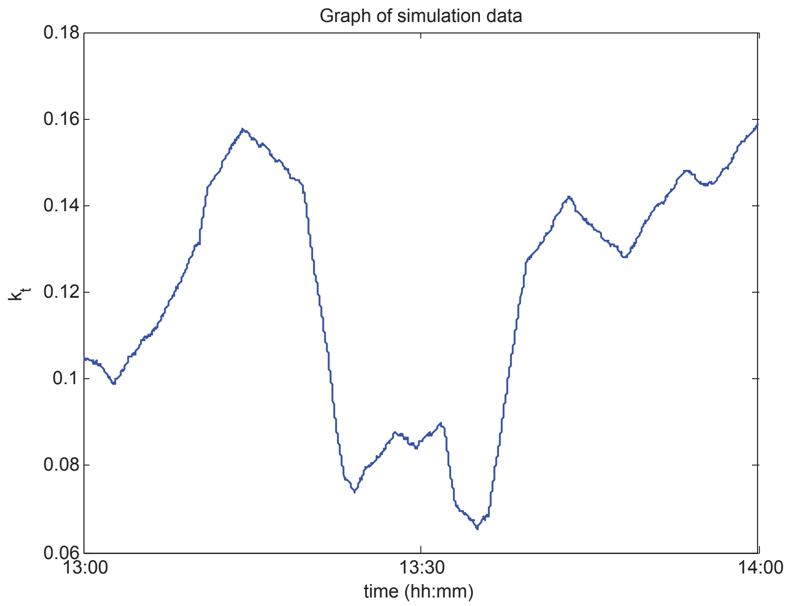


Figure 3.19: Graph of observed data k-DATA-I.1, (a), and its simulation, (b).



(a)



(b)

Figure 3.20: Graph of observed data k-DATA-III.3, (a), and its simulation, (b).

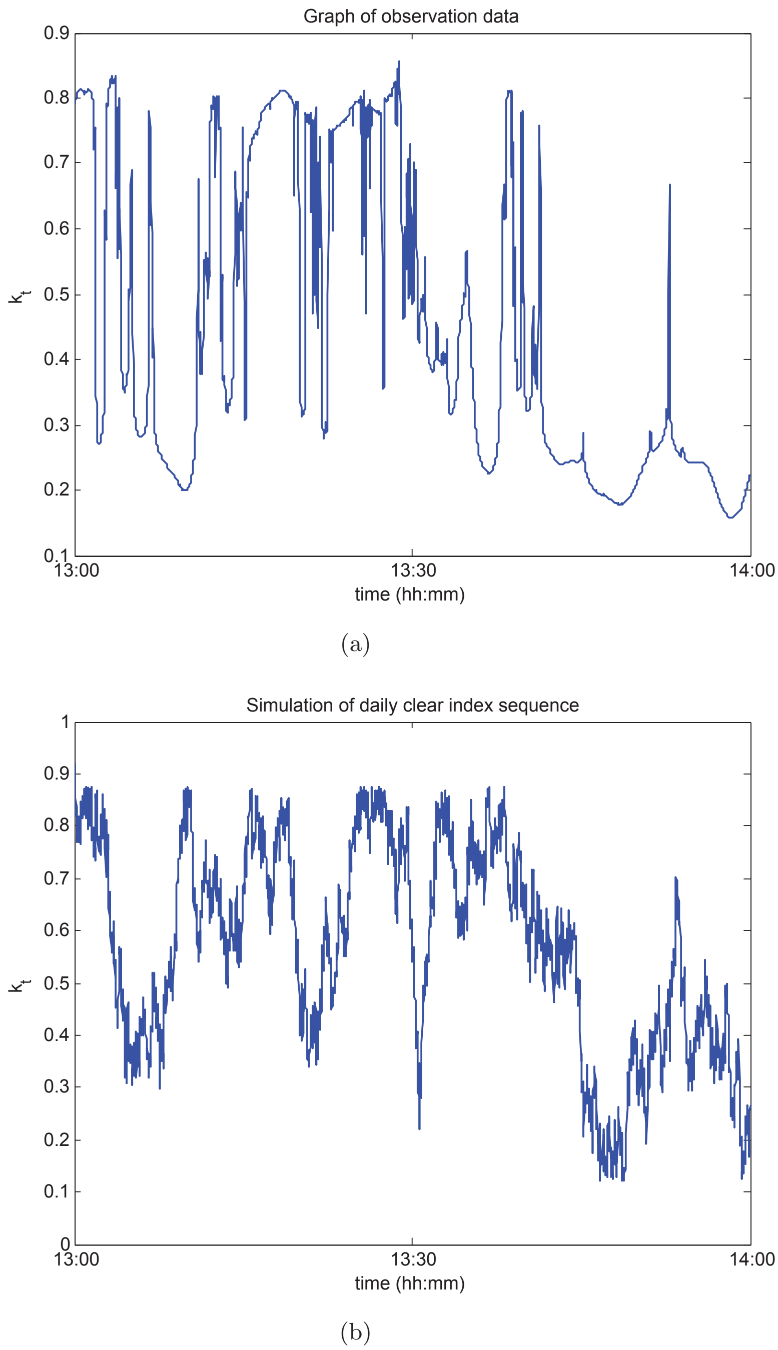


Figure 3.21: Graph of observed data k-DATA-II.2, (a), and its simulation, (b).

Chapter 4

A Stochastic model for the total solar radiation process

Contents

Résumé

Dans ce chapitre, nous proposons un HMM à temps continu, noté CTM-y, pour modéliser le rayonnement solaire total $(y_t)_{t\geq 0}$ considéré comme un processus aléatoire en milieu aléatoire.

Comme pour CTM-k, le processus détat (X_t) est une chaîne de Markov à temps continu dont les états représentent les divers régimes météorologiques mais maintenant, le processus d'observation (y_t) des mesures du rayonnement solaire total est solution de l'équation différentielle stochastique en milieu aléatoire suivante :

$$dy_t = [g(X_t)I_t - y_t]dt + \sigma(X_t)y_t dW_t,$$

où I_t désigne le rayonnement extraterrestre à l'instant t et (W_t) est un mouvement Brownien standard.

Nous détaillons la méthode de probabilité de référence avec la technique de changement de mesure et nous établissons les équations de filtrage pour la mise à jour de l'estimation du vecteur de paramètres dans l'algorithme EM. Nous faisons de même pour DTAM-y, un modèle approché à temps discret.

Les applications numériques, faites à partir de mesures échantillonnées à 1Hz, permettent la. Les travaux de ce chapitre ont été présentés à [Tran 2011] et dans [Tran b].

Abstract

In the present chapter, we propose a continuous-time HMM, denoted CTM-y, for modelling the total solar radiation $(y_t)_{t\geq 0}$ considered as a stochastic process in random medium.

As for CTM-k, the state process (X_t) is a continuous-time Markov chain whose states represent various meteorological regimes but now, the total solar radiation observation process (y_t) is solution of the following stochastic differential equation in random medium:

$$dy_t = [g(X_t)I_t - y_t]dt + \sigma(X_t)y_t dW_t,$$

where I_t denotes the extraterrestrial radiation at time t and (W_t) a standard Brownian motion.

We detail the reference probability method with the change of measure technic and we establish the filter equations for updating the parameter vector estimation in EM algorithm. We do the same for DTAM-y, an approximating discrete-time model.

Experiments are done using measurements sampled at 1Hz.

The works of this chapter have been presented in [Tran 2011] and in [Tran b].

4.1 CTM-y

4.1.1 State process

Assuming that there are $N \geq 1$ *regimes* (states) in meteorological dynamics, we will model the state dynamics by a continuous-time homogeneous Markov chain $(X_t)_{0 \leq t \leq T}$ with state space $\mathbf{S} = \{e_1, e_2, ..., e_N\}$, transition rate matrix $A^{\Delta} = (\lambda_{ji}) \in \mathbf{R}^{N \times N}$ and verifying the following equation

$$X_t = X_0 + \int_0^t A^{\Delta} X_s ds + V_t, \tag{4.1}$$

where (V_t) is a $(\mathcal{F}_t)$-martingale.

4.1.2 Pseudo-clearness index

Recall that clearness index k_t is a dimensionless indicator of total solar radiation G_t: $k_t = \frac{G_t}{I_t}$, where I_t is the extraterrestrial radiation.

We now assume that each meteorological *regime* $i = 1, \ldots, N$ induces a level of *pseudo-clearness index*, say g_i. Let $g = (g_1, g_2, \ldots, g_N)'$ and assume that the g_i's are distinct.

The dynamics of such levels is descibed through the *pseudo-clearness index function*:

$$g(X_t) = \sum_{i=1}^{N} \langle X_t, e_i \rangle g_i = \langle X_t, g \rangle. \tag{4.2}$$

4.1.3 Observation process

During regime i, we assume that the process (y_t) is solution of a SDE with an almost affine drift depending on both pseudo-clearness index and extraterrestrial radiation, the change dy_t generally being likely to be negative but not lower than a threshold:

$$\langle X_t, e_i \rangle dy_t = \langle X_t, e_i \rangle [(g_i I_t - y_t)dt + \sigma_i y_t dW_t], \tag{4.3}$$

where I_t is the extraterrestrial radiation received at time t, (W_t) is a standard Brownian motion and σ_i^2 is a constant noise variance during regime i.

We then have

$$\sum_{i=1}^{N} \langle X_t, e_i \rangle dy_t = \sum_{i=1}^{N} \langle X_t, e_i \rangle [(g_i I_t - y_t)dt + \sigma_i y_t dW_t]. \tag{4.4}$$

and since $\sum_{i=1}^{N} \langle X_t, e_i \rangle = 1$, (4.4) can be written as

$$dy_t = [g(X_t) I_t - y_t]dt + \sigma(X_t) y_t dW_t, \tag{4.5}$$

where $\sigma(X_t) = \sum_{i=1}^{N} \langle X_t, e_i \rangle \sigma_i = \langle X_t, \sigma \rangle$ with $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_N)'$.

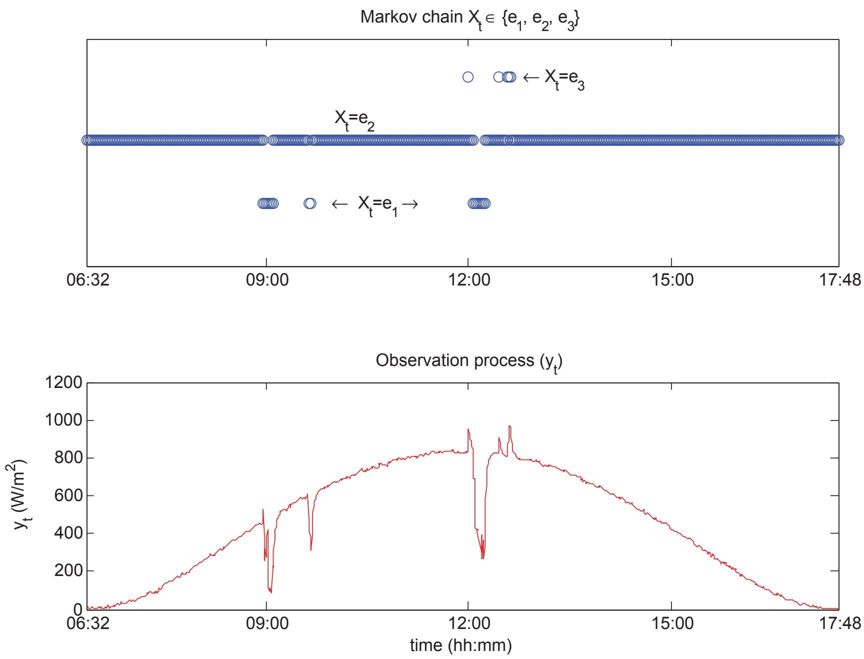


Figure 4.1: An illustration for the CTM-y: $(X_t, y_t)_{t\in[0,T]}$.

The proposed model CTM-y that we propose for the total solar radiation process is the pair of processes (X_t, y_t) (Figure 4.1) defined by equations (4.1) and (4.5), with parameter vector:

$$\theta := \{\lambda_{ji}, 1 \leq i \neq j \leq N;\ g_1, g_2, \ldots, g_N;\ \sigma_1, \sigma_2, \ldots, \sigma_N\}. \tag{4.6}$$

The number of states N will be suggested by the user and our aim is to estimate the parameter vector θ given the observed data.

4.1.4 Filtrations

Let $\mathcal{G}_t^y = \sigma\{X_s, y_s; 0 \leq s \leq t\}$ and $\mathcal{Y}_t^y = \sigma\{y_s; 0 \leq s \leq t\}$ denote the filtrations of the complete data and the incomplete data, respectively. These filtrations contain all the available information of observation history up to time t and $\mathcal{Y}_t^y$ will be used to estimate the parameters of the model.

4.1.5 Change of measure

Let us now detail the reference probability method with the change-of-measure technique.

By Ito formula

$$d\left(\frac{\log y_t}{\sigma(X_t)}\right) = \frac{g(X_t)I_t - y_t}{\sigma(X_t)y_t}dt + dW_t. \tag{4.7}$$

Working on the reference probability space $(\Omega, (\mathcal{F}_t), \overline{P})$ with a fix probability measure $\overline{P}$, we will frst establish the equations of filter processes. Then we will translate these results in the "real world" $(\Omega, (\mathcal{F}_t), P_\theta)$ with a change of probability measure using Girsanov theorem.

Using Ito formula in the "real world" $(\Omega, (\mathcal{F}_t), P_\theta)$ the observation process equation (4.5) has the equivalent form:

$$d\left(\frac{\log y_t}{\sigma(X_t)}\right) = \frac{g(X_t)I_t - y_t}{\sigma(X_t)y_t}dt + dW_t. \tag{4.8}$$

Consider a change of measure from P_θ to $\overline{P}$ such that

$$\left.\frac{\overline{P}}{P_\theta}\right|_{\mathcal{G}_t^y} = \overline{\Lambda}_t^{y,\theta} = \exp\left\{-\int_0^t \frac{g(X_s)I_s - y_s}{\sigma(X_s)y_s}dW_s - \frac{1}{2}\int_0^t \left[\frac{g(X_s)I_s - y_s}{\sigma(X_s)y_s}\right]^2 ds\right\}. \tag{4.9}$$

By Girsanov theorem and (4.8), $\left(\frac{\log y_t}{\sigma(X_t)}\right)$ is a standard Brownian motion under $\overline{P}$, a more convenient measure to work with.

We can come back to the "real world" with the initial probability measure P_θ by inverse change:

$$\left.\frac{P_\theta}{\overline{P}}\right|_{\mathcal{G}_t^y} = \Lambda_t^{y,\theta} = \exp\left\{\int_0^t \frac{g(X_s)I_s - y_s}{\sigma(X_s)y_s}d\left(\frac{\log y_s}{\sigma(X_s)}\right) - \frac{1}{2}\int_0^t \left[\frac{g(X_s)I_s - y_s}{\sigma(X_s)y_s}\right]^2 ds\right\}. \tag{4.10}$$

Girsanov theorem implies that under P_θ, $(\tilde{W}_t)$ is a standard Brownian motion if we define

$$d\tilde{W}_t = d\left(\frac{\log y_t}{\sigma(X_t)}\right) - \frac{g(X_t)I_t - y_t}{\sigma(X_t)y_t}dt. \tag{4.11}$$

It follows that under P_θ,

$$dy_t = [g(X_t)I_t - y_t]dt + \sigma(X_t)y_t d\tilde{W}_t, \tag{4.12}$$

where $(\tilde{W}_t)$ is a standard Brownian motion. That is, P_θ is the distribution law of the observation process (y_t) in the "real world".

Note that:

1. By Ito formula and (4.10), the process $\overline{\Lambda}_t^\theta$ is solution of the SDE:

$$\overline{\Lambda}_t^{y,\theta} = 1 - \int_0^t \overline{\Lambda}_s \frac{g(X_s)I_s - y_s}{\sigma(X_s)y_s}dW_s. \tag{4.13}$$

2. Similarly,

$$\Lambda_t^{y,\theta} = 1 + \int_0^t \Lambda_s^\theta \frac{g(X_s)I_s - y_s}{\sigma(X_s)y_s} d\left(\frac{\log y_s}{\sigma(X_s)}\right) \tag{4.14}$$

$$= 1 + \int_0^t \Lambda_s^\theta \left(\frac{g(X_s)I_s - y_s}{\sigma(X_s)y_s}\right)^2 ds + \int_0^t \Lambda_s^\theta \frac{g(X_s)I_s - y_s}{\sigma(X_s)y_s}dW_s. \tag{4.15}$$

3. We have $\Lambda_t^{y,\theta}\overline{\Lambda}_t^{y,\theta} = 1$.

4.2 Parameter estimations in continuous time

As for CTM-k, the estimates of transition probabilities λ_{ji} $(1 \leq i \neq j \leq N)$ are given by [Elliott 1995, chap. 3, pg. 68-70] and [James 1996]:

$$\lambda'_{ji} = \frac{\pi(\mathcal{J}_t^{ji})}{\pi(\mathcal{O}_t^i)}, \quad 1 \leq j \neq i \leq N \tag{4.16}$$

where $\pi(\mathcal{J}_t^{ji})$ and $\pi(\mathcal{O}_t^i)$ are the normalized filter (see Section 2.8.2).

We now proceed to the computations in EM algorithm for getting ML estimates of the parameter vector $g = (g_1, g_2, \ldots, g_N)'$.

4.2.1 Expectation Step

Update from parameter θ to new parameter θ' will be obtained by maximizing the pseudo log-likelihood function defined as

$$Q(\theta', \theta) = E_\theta \left(\log \Lambda_t^{y,\theta'\theta} | \mathcal{Y}_t^y \right), \tag{4.17}$$

where

$$\Lambda_t^{y,\theta'\theta} = \left. \frac{dP_{\theta'}}{dP_\theta} \right|_{\mathcal{G}_t^y}.$$

First, we have:

$$\begin{aligned} Q(\theta', \theta) = & E_\theta \left\{ \int_0^t \left[\frac{\langle X_s, g' \rangle I_s - y_s}{(\langle X_s, \sigma' \rangle y_s)^2} - \frac{\langle X_s, g \rangle I_s - y_s}{(\langle X_s, \sigma \rangle y_s)^2} \right] dy_s \right. \\ & \left. - \frac{1}{2} \int_0^t \left[\left(\frac{\langle X_s, g' \rangle I_s - y_s}{\langle X_s, \sigma' \rangle y_s} \right)^2 - \left(\frac{\langle X_s, g \rangle I_s - y_s}{\langle X_s, \sigma \rangle y_s} \right)^2 \right] ds \middle| \mathcal{Y}_t^y \right\}, \end{aligned} \tag{4.18}$$

where $g' = (g'_1, g'_2, \ldots, g'_N)'$, $\sigma' = (\sigma'_1, \sigma'_2, \ldots, \sigma'_N)'$.

At the p-th iteration of the algorithm Expectation Step, we set $\theta = \theta'^{(p)}$ in (4.18) to obtain

$$\begin{aligned} Q(\theta', \theta'^{(p)}) = E_{\theta'^{(p)}} & \left\{ \sum_{i=1}^N \frac{1}{\sigma_i'^2} \left[g'_i \int_0^t \frac{I_s}{y_s^2} \langle X_s, e_i \rangle dy_s - \int_0^t \frac{1}{y_s} \langle X_s, e_i \rangle dy_s \right. \right. \\ & - \frac{1}{2} g_i'^2 \int_0^t \frac{I_s^2}{y_s^2} \langle X_s, e_i \rangle ds + g'_i \int_0^t \frac{I_s}{y_s} \langle X_s, e_i \rangle ds \\ & \left. \left. - \frac{1}{2} \int_0^t \langle X_s, e_i \rangle ds \right] \middle| \mathcal{Y}_t^y \right\} + R(\theta'^{(p)}, \mathcal{Y}_t^y), \end{aligned} \tag{4.19}$$

where $R(\theta'^{(p)}, \mathcal{Y}_t^y)$ does not depend on θ'.

4.2.2 Maximization Step

The maximum of $Q(\theta', \theta'^{(p)})$ is achieved by solving

$$\frac{\partial}{\partial \theta'} Q(\theta', \theta'^{(p)}) = 0.$$

We will then find a ML estimate g'_i for updating g_i, $i = 1, 2, \ldots, N$ as follows. We first see that

$$\begin{aligned}\frac{\partial}{\partial g'_i} Q(\theta', \theta'^{(p)}) =& \frac{1}{\sigma_i'^2}\left[E_{\theta'^{(p)}}\left(\int_0^t \frac{I_s}{y_s^2} \langle X_s, e_i \rangle dy_s \middle| \mathcal{Y}_t^y \right) - g'_i E_{\theta'^{(p)}}\left(\int_0^t \frac{I_s^2}{y_s^2} \langle X_s, e_i \rangle ds \middle| \mathcal{Y}_t^y \right) \right. \\ & \left. + E_{\theta'^{(p)}}\left(\int_0^t \frac{I_s}{y_s} \langle X_s, e_i \rangle ds \middle| \mathcal{Y}_t^y \right) \right]. \end{aligned} \tag{4.20}$$

Let

$$\mathcal{M}_t^i = \int_0^t \frac{I_s}{y_s} \langle X_s, e_i \rangle ds, \tag{4.21}$$

$$\mathcal{N}_t^i = \int_0^t \frac{I_s^2}{y_s^2} \langle X_s, e_i \rangle ds, \tag{4.22}$$

$$\mathcal{T}_t^i = \int_0^t \frac{I_s}{y_s^2} \langle X_s, e_i \rangle dy_s. \tag{4.23}$$

Then

$$\frac{\partial}{\partial g'_i} Q(\theta', \theta'^{(p)}) = 0 \Longleftrightarrow g'_i = \frac{E_{\theta'^{(p)}}(\mathcal{M}_t^i | \mathcal{Y}_t^y) + E_{\theta'^{(p)}}(\mathcal{T}_t^i | \mathcal{Y}_t^y)}{E_{\theta'^{(p)}}(\mathcal{N}_t^i | \mathcal{Y}_t^y)}. \tag{4.24}$$

For $H_t \equiv \mathcal{M}_t^i$, $\mathcal{N}_t^i$ and $\mathcal{T}_t^i$, we use the unnormalized filters $\pi(H_t) = E_{\theta'^{(p)}}(H_t | \mathcal{Y}_t^y)$, to see that the new estimate g'_i of g_i up to time T is

$$g'_i = \frac{\pi(\mathcal{M}_T^i) + \pi(\mathcal{T}_T^i)}{\pi(\mathcal{N}_T^i)}, \quad i = 1, 2, \ldots, N, \tag{4.25}$$

Now, we have to compute $\pi(H_t) = E_{\theta'^{(p)}}(H_t | \mathcal{Y}_t^y)$, for $H_t \equiv \mathcal{J}_t^{ij}$, $\mathcal{O}_t^i$, $\mathcal{M}_t^i$, $\mathcal{N}_t^i$ and $\mathcal{T}_t^i$.

4.3 Equation of continuous time filters

As mentioned briefly in Chapter 2 and as applied in Chapter 3, we can obtain the normalized filter $\pi(H_t) = E_{\theta'^{(p)}}(H_t | \mathcal{Y}_t^y)$ from the unnormalized filter $\gamma(H_t) = \overline{E}(H_t \Lambda_t^{y,\theta} | \mathcal{Y}_t^y)$ by noticing that

$$\pi(H_t) = \frac{\gamma(H_t)}{\gamma(1)} \tag{4.26}$$

Now our aim is to compute $\gamma(H_t)$ for $H_t \equiv 1$, $\mathcal{J}_t^{ij}$, $\mathcal{O}_t^i$, $\mathcal{M}_t^i$, $\mathcal{N}_t^i$ or $\mathcal{T}_t^i$.

Recall that it is not generally possible to obtain directly the equations for $\gamma(H_t)$, see [James 1996], but, it is possible to obtain the equations for $\gamma(H_t X_t)$ and then get $\gamma(H_t)$ as $\gamma(H_t) = \langle \gamma(H_t X_t), \underline{1} \rangle$, where $\underline{1}$ denotes the column N-vector of ones.

In particular, for $H_t \equiv 1$, $p_t = \gamma(X_t)$, and $\gamma(1) = \langle \gamma(X_t), \underline{1} \rangle = \langle p_t, \underline{1} \rangle$ so that (4.26) yields

$$\pi(H_t) = \frac{\langle \gamma(H_t X_t), \underline{1} \rangle}{\langle p_t, \underline{1} \rangle} \tag{4.27}$$

A recursive equation for $\gamma(X_t H_t)$ for $H_t \equiv 1$, $\mathcal{J}_t^{ij}$, $\mathcal{O}_t^i$, $\mathcal{M}_t^i$, $\mathcal{N}_t^i$ and $\mathcal{T}_t^i$ is stated in the following theorem.

Theorem 4.1. *Let H_t a scalar process of the form*

$$H_t = H_0 + \int_0^t \mu_s ds + \int_0^t \rho_s' dV_s + \int_0^t \delta_s dW_s, \tag{4.28}$$

where the real μ_s, δ_s and N-dimensional vector ρ_s are $(\mathcal{F}_t)$-predictable, square-integrable processes.

The unnormalized estimate $\gamma(H_t X_t)$ is given by

$$\begin{aligned}\gamma(H_t X_t) =& \gamma(H_0 X_0) + \int_0^t A^{\Delta} \gamma(H_s X_s) ds + \int_0^t \gamma(\mu_s X_s) ds \\ &+ \sum_{i,j=1}^{N} \int_0^t \langle \gamma(\rho_s^j X_s - \rho_s^i X_s), e_i \rangle a_{ji} ds (e_j - e_i) \\ &+ \int_0^t C_s \gamma(X_s H_s) \frac{dy_s}{\sigma(X_s) y_s} + \int_0^t \gamma(\delta_s X_s) \frac{dy_s}{\sigma(X_s) y_s}.\end{aligned} \tag{4.29}$$

Proof. Applying Ito product rule for semimartingales, compute the product $H_t X_t$ (where the process X_t is given by (2.36)) as follows:

$$\begin{aligned}H_t X_t =& H_0 X_0 + \int_0^t H_s A^{\Delta} X_s ds + \int_0^t H_s dV_s \\ &+ \int_0^t \mu_s X_s ds + \int_0^t \rho_s' X_s dV_s + \int_0^t \delta_s X_s dB_s + [X, H]_t\end{aligned} \tag{4.30}$$

Since $\Delta H_t = \rho_t' \Delta V_t + \rho_t' \Delta W_t$, $\Delta X_t = \Delta V_t$ and $\Delta X_t \Delta W_t = 0$ a.s., we have:

$$\begin{aligned}[X, H]_t &= \sum_{0 < s \leq t} (\rho_s' \Delta X_s) \Delta X_s \\ &= \int_0^t \left(\rho_s' dX_s - \rho_s' X_s \right) (dX_s - X_s) \\ &= \sum_{i,j=1}^{N} \int_0^t (\rho_s^j - \rho_s^i) \langle X_s, e_i \rangle \langle dX_s, e_j \rangle (e_j - e_i).\end{aligned}$$

Then, using $dX_s = AX_s ds + dV_s$, we get:

$$\begin{aligned}[X, H]_t = & \sum_{i,j=1}^{N} \int_0^t (\rho_s^j - \rho_s^i) \langle X_s, e_i \rangle \langle dV_s, e_j \rangle (e_j - e_i) \\ &+ \sum_{i,j=1}^{N} \int_0^t \langle \rho_s^j X_s - \rho_s^i X_s \rangle a_{ji} ds (e_j - e_i)\end{aligned}$$

Plugging in (4.30) yields

$$\begin{aligned}
H_tX_t =& H_0X_0 + \int_0^t H_sA^\Delta X_s ds + \int_0^t H_s dV_s \\
&+ \int_0^t \mu_s X_s ds + \int_0^t \rho'_s X_s dV_s + \int_0^t \delta_s X_s dW_s \qquad (4.31)\\
&+ \sum_{i,j=1}^N \int_0^t (\rho_s^j - \rho_s^i)\langle X_s, e_i\rangle \langle dV_s, e_j\rangle (e_j - e_i) \\
&+ \sum_{i,j=1}^N \int_0^t \langle \rho_s^j X_s - \rho_s^i X_s, e_i\rangle a_{ji} ds (e_j - e_i) \qquad (4.32)
\end{aligned}$$

Now, Ito product applied to the processes $\Lambda_t^{y,\theta}$ and H_tX_t yields:

$$\begin{aligned}
\Lambda_t^{y,\theta} H_tX_t =& H_0X_0 + \int_0^t \Lambda_s^{y,\theta} H_sA^\Delta X_s ds + \int_0^t \Lambda_s^{y,\theta} H_s dV_s \\
&+ \int_0^t \mu_s \Lambda_s^{y,\theta} X_s ds + \int_0^t \rho'_s \Lambda_s^{y,\theta} X_s dV_s + \int_0^t \delta_s \Lambda_s^{y,\theta} X_s dW_s \qquad (4.33)\\
&+ \sum_{i,j=1}^N \int_0^t (\rho_s^j - \rho_s^i)\langle \Lambda_s^{y,\theta} X_s, e_i\rangle \langle dV_s, e_j\rangle (e_j - e_i) \\
&+ \sum_{i,j=1}^N \int_0^t \langle \rho_s^j \Lambda_s^{y,\theta} X_s - \rho_s^i \Lambda_s^{y,\theta} X_s, e_i\rangle a_{ji} ds (e_j - e_i) \\
&+ \int_0^t \Lambda_s^{y,\theta} X_s H_s \frac{g(X_s)I_s - y_s}{\sigma(X_s)y_s} \frac{dy_s}{\sigma(X_s)y_s} + [HX, \Lambda^{y,\theta}]_t \qquad (4.34)
\end{aligned}$$

where the quadratic variation is given by

$$[XH, \Lambda^{y,\theta}]_t = \int_0^t \delta_s \Lambda_s^{y,\theta} X_s \frac{g(X_s)I_s - y_s}{\sigma(X_s)y_s} ds \tag{4.35}$$

Plugging (4.35) into (4.34), we get

$$\int_0^t \Lambda_s^{y,\theta} X_s H_s \frac{g(X_s)I_s - y_s}{\sigma(X_s)y_s} \frac{dy_s}{\sigma(X_s)y_s} = \sum_{i=1}^N \int_0^t \langle \Lambda_s^{y,\theta} X_s H_s, e_i\rangle \frac{g_iI_s - y_s}{\sigma_i y_s} \frac{dy_s}{\sigma_i y_s} e_i, \tag{4.36}$$

$$\begin{aligned}
\int_0^t \delta_s \Lambda_s^{y,\theta} X_s dB_s + \int_0^t \delta_s \Lambda_s^{y,\theta} X_s \frac{g(X_s)I_s - y_s}{\sigma(X_s)y_s} ds &= \int_0^t \delta_s \Lambda_s^{y,\theta} X_s \frac{dy_s}{\sigma(X_s)y_s} \\
&= \sum_{i=1}^N \int_0^t \langle \Lambda_s^{y,\theta} \delta_s X_s, e_i\rangle \frac{dy_s}{\sigma_i y_s} e_i.
\end{aligned} \tag{4.37}$$

Hence

$$\begin{aligned}\Lambda_t^{y,\theta} H_t X_t =& H_0 X_0 + \int_0^t \Lambda_s^{y,\theta} H_s A^{\Delta} X_s ds + \int_0^t \Lambda_s^{y,\theta} H_s dV_s + \int_0^t \mu_s \Lambda_s^{y,\theta} X_s ds \\ &+ \int_0^t \rho_s' \Lambda_s^{y,\theta} X_s dV_s + \sum_{i,j=1}^N \int_0^t (\rho_s^j - \rho_s^i) \langle \Lambda_s^{y,\theta} X_s, e_i \rangle \langle dV_s, e_j \rangle (e_j - e_i) \\ &+ \sum_{i,j=1}^N \int_0^t \langle \rho_s^j \Lambda_s^{y,\theta} X_s - \rho_s^i \Lambda_s^{y,\theta} X_s, e_i \rangle a_{ji} ds (e_j - e_i) \\ &+ \sum_{i=1}^N \int_0^t \langle \Lambda_s^{y,\theta} X_s H_s, e_i \rangle \frac{g_i I_s - y_s}{\sigma_i y_s} \frac{dy_s}{\sigma_i y_s} e_i \\ &+ \sum_{i=1}^N \int_0^t \langle \Lambda_s^{y,\theta} \delta_s X_s, e_i \rangle \frac{dy_s}{\sigma_i y_s} e_i. \end{aligned} \tag{4.38}$$

Taking conditional expectation $\overline{E}$ given $\mathcal{Y}_t^y$ of both sides of (4.39) we obtain

$$\begin{aligned}\gamma(H_t X_t) =& \gamma(H_0 X_0) + \int_0^t A^{\Delta} \gamma(H_s X_s) ds + \int_0^t \gamma(\mu_s X_s) ds \\ &+ \sum_{i,j=1}^N \int_0^t \langle \gamma(\rho_s^j X_s - \rho_s^i X_s), e_i \rangle a_{ji} ds (e_j - e_i) \\ &+ \sum_{i=1}^N \int_0^t \langle \gamma(X_s H_s), e_i \rangle \frac{g_i I_s - y_s}{\sigma_i y_s} \frac{dy_s}{\sigma_i y_s} e_i \\ &+ \sum_{i=1}^N \int_0^t \langle \gamma(\delta_s X_s), e_i \rangle \frac{dy_s}{\sigma_i y_s} e_i. \end{aligned} \tag{4.39}$$

As

$$\sum_{i=1}^N \int_0^t \langle \gamma(\delta_s X_s), e_i \rangle \frac{dy_s}{\sigma_i y_s} e_i = \int_0^t \gamma(\delta_s X_s) \frac{dy_s}{\sigma(X_s) y_s},$$

(4.39) completes the proof. □

Let us specify Theorem 4.1.

Take $H_t = H_0 = 1$, $\mu_s = \delta_s = 0$, $\rho = \underline{0}$. Then

Corollary 4.1. *The recursive equation for the state process* $p_t = \gamma(X_t) = \overline{E}(X_t \Lambda_t^{y,\theta} | \mathcal{Y}_t^y)$ *is:*

$$p_t = \pi_0 + \int_0^t A^{\Delta} p_s ds + \int_0^t C_s p_s \frac{dy_s}{\sigma(X_s) y_s} \tag{4.40}$$

As in (3.62), we decompose the number of jumps from e_i to e_j in the time interval $[0, t]$ as follows:

$$\mathcal{J}_t^{ij} = \int_0^t \langle X_s, e_i \rangle a_{ji} ds + \int_0^t \langle X_s, e_i \rangle e_j' dV_s$$

Then, applying Theorem 4.1 with $H_t = \mathcal{J}_t^{ij}$, $H_0 = 0$, $\mu_s = \langle X_s, e_i\rangle a_{ji}$, $\rho = \langle X_s, e_i\rangle e_j$, $\delta_s = 0$, we get:

Corollary 4.2.

$$\gamma(\mathcal{J}_t^{ij} X_t) = \int_0^t A^\Delta \gamma(\mathcal{J}_t^{ij} X_s)ds + \int_0^t a_{ji}\langle p_s, e_i\rangle e_j ds + \int_0^t C_s \gamma(\mathcal{J}_s^{ij} X_s)\frac{dy_s}{\sigma(X_s)y_s} \tag{4.41}$$

Note here that

$$\sum_{i,j=1}^N \int_0^t \langle \gamma(\rho_s^j X_s - \rho_s^i X_s), e_i\rangle a_{ji} ds (e_j - e_i) = \int_0^t a_{ji}\langle p_s, e_i\rangle e_j ds - \int_0^t a_{ji}\langle p_s, e_i\rangle e_i ds.$$

because $\rho_s^j = \langle X_s, e_i\rangle$, $\rho_s^i = 0$.

Recalling that $X_s\langle X_s, e_i\rangle = \langle X_s, e_i\rangle e_i$ and applying Theorem 4.1 to the processes $\mathcal{O}_t^i = \int_0^t \langle X_s, e_i\rangle ds$, $\mathcal{M}_t^i = \int_0^t \frac{I_s}{y_s}\langle X_s, e_i\rangle ds$ and $\mathcal{N}_t^i = \int_0^t \frac{I_s^2}{y_s^2}\langle X_s, e_i\rangle ds$, $i = 1, 2, \ldots, N$, we obtain the recursive equations for $\gamma(\mathcal{O}_t^i X_t)$, $\gamma(\mathcal{M}_t^i X_t)$ and $\gamma(\mathcal{N}_t^i X_t)$, respectively, as stated in Corollary 4.3, Corollary 4.4 and Corollary 4.5 below.

Take $H_t = \mathcal{O}_t^i$, $H_0 = 0$, $\mu_s = \langle X_s, e_i\rangle$, $\delta_s = 0$, $\rho = \underline{0}$, to get

Corollary 4.3.

$$\gamma(\mathcal{O}_t^i X_t) = \int_0^t A^\Delta \gamma(\mathcal{O}_s^i X_s)ds + \int_0^t C_s \gamma(\mathcal{O}_s^i X_s)\frac{dy_s}{\sigma(X_s)y_s} + \int_0^t \langle p_s, e_i\rangle e_i ds \tag{4.42}$$

Take $H_t = \mathcal{M}_t^i$, $H_0 = 0$, $\mu_s = \frac{I_s}{y_s}\langle X_s, e_i\rangle$, $\delta_s = 0$, $\rho = \underline{0}$, to get

Corollary 4.4.

$$\gamma(\mathcal{M}_t^i X_t) = \int_0^t A^\Delta \gamma(\mathcal{M}_s^i X_s)ds + \int_0^t C_s \gamma(\mathcal{M}_s^i X_s)\frac{dy_s}{\sigma(X_s)y_s} + \int_0^t \frac{I_s}{y_s}\langle p_s, e_i\rangle e_i ds \tag{4.43}$$

Take $H_t = \mathcal{N}_t^i$, $H_0 = 0$, $\mu_s = \frac{I_s^2}{y_s^2}\langle X_s, e_i\rangle$, $\delta_s = 0$, $\rho = \underline{0}$, to get

Corollary 4.5.

$$\gamma(\mathcal{N}_t^i X_t) = \int_0^t A^\Delta \gamma(\mathcal{N}_s^i X_s)ds + \int_0^t C_s \gamma(\mathcal{N}_s^i X_s)\frac{dy_s}{\sigma(X_s)y_s} + \int_0^t \frac{I_s^2}{y_s^2}\langle p_s, e_i\rangle e_i ds \tag{4.44}$$

Consider now $\mathcal{T}_t^i = \int_0^t \frac{I_s}{y_s^2}\langle X_s, e_i\rangle dy_s$, $i = 1, 2, \ldots, N$ and observe that

$$\mathcal{T}_t^i = \int_0^t \frac{I_s}{y_s^2}(I_s g_i - y_s)\langle X_s, e_i\rangle ds + \int_0^t \frac{I_s}{y_s}\sigma_i\langle X_s, e_i\rangle dB_s. \tag{4.45}$$

Applying Theorem 4.1 with $H_t = \mathcal{T}_t^i$, $H_0 = 0$, $\mu_s = \frac{I_s}{y_s^2}(I_s g_i - y_s)\langle X_s, e_i\rangle$, $\rho = \underline{0}$, $\delta_s = \frac{I_s}{y_s}\sigma_i\langle X_s, e_i\rangle$, we obtain

$$\begin{aligned}\gamma(\mathcal{T}_t^i X_t) = &\int_0^t A^\Delta \gamma(\mathcal{T}_s^i X_s)ds + \int_0^t C_s \gamma(\mathcal{T}_s^i X_s)\frac{dy_s}{\sigma(X_s)y_s} \\ &+ \int_0^t \frac{I_s}{y_s^2}(I_s g_i - y_s)\langle p_s, e_i\rangle e_i ds + \int_0^t \frac{I_s}{y_s}\sigma_i\langle p_s, e_i\rangle e_i \frac{dy_s}{\sigma(X_s)y_s}\end{aligned} \tag{4.46}$$

Noticing that

$$\int_0^t \frac{I_s}{y_s}\sigma_i\langle p_s, e_i\rangle e_i \frac{dy_s}{\sigma(X_s)y_s} = \int_0^t \frac{I_s}{y_s}\sigma_i\langle p_s, e_i\rangle e_i \frac{dy_s}{\sigma_i y_s}$$

we then get

Corollary 4.6.

$$\begin{aligned}\gamma(\mathcal{T}_t^i X_t) = &\int_0^t A^\Delta \gamma(\mathcal{T}_s^i X_s)ds + \int_0^t C_s \gamma(\mathcal{T}_s^i X_s)\frac{dy_s}{\sigma(X_s)y_s} \\ &+ \int_0^t \frac{I_s}{y_s^2}(I_s g_i - y_s)\langle p_s, e_i\rangle e_i ds + \int_0^t \frac{I_s}{y_s^2}\langle p_s, e_i\rangle e_i dy_s. \end{aligned} \tag{4.47}$$

4.4 Discrete-time approximating model DTAM

In this section we consider partition (3.65) of time interval $[0, T]$ ($0 = t_0 < t_1 < t_2 < \cdots < t_{m-1} < t_m < \cdots < t_M \leq T$ with equal width $\Delta = t_m - t_{m-1}$) to obtain an approximating model of Continuous-Time Model (CTM) and the computable approximations of the continuous-time filter equations described in Section 4.2.

4.4.1 Components of the model

From partition (**??**), we first define the following discrete-time sampled observations:

$$z_m = \frac{1}{\Delta}(y_m - y_{m-1}), \quad m = 1, 2 \ldots \tag{4.48}$$

The CTM-y discrete-time approximating model is the pair of processes $(X_m, z_m)_{m=1,2,\ldots}$ defined as follows:

1. the *state process* is a discrete time Markov chain $(X_h)_{h=0,1,2,\ldots}$ with state space $\mathbf{S} = \{e_1, e_2, \ldots, e_N\}$ and transition probability matrix:

$$P = I + \Delta A^\Delta = (\epsilon_{ji}) \in \mathbf{R}^{N\times N};$$

2. the state process (X_m) is related to the observation sequence $(z_m)_{m=1,2,\ldots,}$ through the equation:

$$z_m = [\langle X_m, g\rangle I_m - y_m] + \frac{1}{\sqrt{\Delta}}\langle X_m, \sigma\rangle y_m w_m, \quad m = 1, 2 \ldots, \tag{4.49}$$

where $w_m \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$, $m = 1, 2 \ldots$;

3. the parameter vector of the model is:

$$\theta = \{\epsilon_{ji}, 1 \leq j \neq i \leq N; g_1, g_2, \ldots, g_N; \sigma_1, \sigma_2, \ldots, \sigma_N\}.$$

4.4.2 Robust approximation of filter equations

We propose now some computable equations approximating the continuous-time equations described in Section 4.2.

For the approximating schema of the state filter equation, we follow again Clark *robust filter* approach [Clark 1977] using the robust forward state filter density introduced by [James 1996, Krishnamurthy 2002].

We first introduce the following processes:

$$\phi_t^{(i)} = \exp\left\{ \int_0^t \frac{g_i I_s - y_s}{\sigma_i y_s} \frac{dy_s}{\sigma_i y_s} - \frac{1}{2} \int_0^t \left(\frac{g_i I_s - y_s}{\sigma_i y_s} \right)^2 ds \right\}, i = 1, 2, \ldots, N, \quad (4.50)$$

$$\Phi_t = \operatorname{diag}\left(\phi_t^{(1)}, \phi_t^{(2)}, \ldots, \phi_t^{(N)} \right). \quad (4.51)$$

Definition 4.1. *For $t \in [0, T]$, define*

$$\overline{p}_t \triangleq (\Phi_t)^{-1} p_t. \quad (4.52)$$

The process $(\overline{p}_t)$ is called the robust filter of the state filter process (p_t).

Considering similarly as in (4.13) and (4.14), we have

$$(\phi_t^{(i)})^{-1} = 1 - \int_0^t (\phi_t^{(i)})^{-1} \frac{g_i I_s - y_s}{\sigma_i y_s} dW_s, \quad i = 1, 2, \ldots, N. \quad (4.53)$$

From the state filter equation,

$$p_t^{(i)} = \pi_0^{(i)} + \int_0^t (\lambda_{i1} p_s^{(1)} + \cdots + \lambda_{iN} p_s^{(N)}) ds + \int_0^t p_s^{(i)} \frac{g_i I_s - y_s}{\sigma_i y_s} \frac{dy_s}{\sigma_i y_s},$$
$$i = 1, 2, \ldots, N. \quad (4.54)$$

By Itô rule, we have

$$(\phi_t^{(i)})^{-1} p_t^{(i)} = \pi_0^{(i)} + \int_0^t (\phi_s^{(i)})^{-1} (\lambda_{i1} p_s^{(1)} + \cdots + \lambda_{iN} p_s^{(N)}) ds, \quad i = 1, 2, \ldots, N. \quad (4.55)$$

An equivalent form of (4.52) is therefore:

$$\overline{p}_t = (\Phi_t)^{-1} p_t = \pi_0 + \int_0^t (\Phi_s)^{-1} A^{\Delta} p_s ds. \quad (4.56)$$

By (4.56), the process $(\overline{p}_t)$ will be determined by the ordinary differential equation:

$$\frac{d}{dt} \overline{p}_t = (\Phi_t)^{-1} A^{\Delta} \Phi_t \overline{p}_t, \quad \overline{p}_0 = \pi_0. \quad (4.57)$$

An approximating schema to get the equation of (p_t) from $(\overline{p}_t)$ will be considered below.

From (4.57), an approximation of $\overline{p}_t$ between the sampling times t_{m-1} and t_m of the partition (3.65) is given by

$$\overline{p}_{t_m} = \overline{p}_{t_{m-1}} + \int_{t_{m-1}}^{t_m} (\Phi_s)^{-1} A^{\Delta} \Phi_s \overline{p}_s ds. \tag{4.58}$$

Euler-Maruyama method (see e.g. [Higham 2001]) yields

$$\overline{p}_{t_m} \simeq \overline{p}_{t_{m-1}} + (\Phi_{t_{m-1}})^{-1} A^{\Delta} \Phi_{t_{m-1}} \overline{p}_{t_{m-1}} \Delta. \tag{4.59}$$

Multiplying both sides of (4.59) by Φ_{t_m}, we obtain

$$p_{t_m} \simeq \Phi_{t_m} (\Phi_{t_{m-1}})^{-1} [\mathbf{I} + \Delta A^{\Delta}] p_{t_{m-1}}, \tag{4.60}$$

where $\mathbf{I}$ is the $N \times N$ unit matrix.

Plugging $\tilde{A} = \mathbf{I} + \Delta A^{\Delta}$ into (4.60) yields

$$p_{t_m} \simeq \Phi_{t_m} (\Phi_{t_{m-1}})^{-1} \tilde{A} p_{t_{m-1}}. \tag{4.61}$$

For $i = 1, 2, \dots, N$, the definition of $\phi_t^{(i)}$ in (4.50) implies that

$$\phi_{t_m}^{(i)} (\phi_{t_{m-1}}^{(i)})^{-1} = \exp\left\{ \int_{t_{h-1}}^{t_m} \frac{g_i I_s - y_s}{(\sigma_i y_s)^2} dy_s - \frac{1}{2} \int_{m-1}^{t_m} \left(\frac{g_i I_s - y_s}{\sigma_i y_s} \right)^2 ds \right\}. \tag{4.62}$$

As the partition (3.65) is such that $\Delta = t_m - t_{m-1}$, $y_m - y_{m-1} = \Delta z_m$, (4.62) becomes

$$\begin{aligned} \phi_{t_m}^{(i)} (\phi_{t_{m-1}}^{(i)})^{-1} &= \exp\left\{ \Delta \left[\frac{(g_i I_m - y_m) z_m}{(\sigma_i y_m)^2} - \frac{1}{2} \left(\frac{g_i I_h - y_m}{\sigma_i y_m} \right)^2 \right] \right\} \\ &= \psi_m^{(i)} \quad i = 1, 2, \dots, N. \end{aligned} \tag{4.63}$$

We then have

$$\Phi_{t_m} (\Phi_{t_{m-1}})^{-1} = \operatorname{diag}\left(\psi_m^{(1)}, \psi_m^{(2)} \dots \psi_m^{(N)} \right) = \Psi_m, \tag{4.64}$$

where

$$\psi_m^{(i)} = \exp\left\{ \Delta \left[\frac{(g_i I_m - y_m) z_m}{(\sigma_i y_m)^2} - \frac{1}{2} \left(\frac{g_i I_m - y_m}{\sigma_i y_m} \right)^2 \right] \right\}, \quad i = 1, 2, \dots, N.$$

From (4.61) and (4.64), it is seen that

$$q_m = \Psi_m \tilde{A} q_{m-1}, \quad q_0 = \pi_0, \quad m = 1, 2, \dots, M, \tag{4.65}$$

defines a numerical approximation q_m of p_{t_m}.

Let $\mathcal{Y}_m^z$ denote the σ-algebra generated by $\{z_1, z_2, \dots, z_m\}$ and let $\tilde{\gamma}(H_m X_m) = \overline{E}(\Lambda_m^{z,\theta} H_m X_m | \mathcal{Y}_m^z)$ for $H_m \equiv \mathcal{J}_m^{ij}$, $\mathcal{O}_m^i$, $\mathcal{M}_m^i$, $\mathcal{N}_m^i$ and $\mathcal{T}_m^i$ $(i = 1, 2, \dots, N)$ with:

$$\mathcal{J}_m^{ij} = \sum_{h=1}^{m} \langle X_{h-1}, e_i \rangle \langle X_h, e_j \rangle, \quad \mathcal{O}_m^i = \sum_{h=1}^{m} \langle X_h, e_i \rangle, \quad \mathcal{M}_m^i = \sum_{h=1}^{m} \frac{I_h}{y_h} \langle X_h, e_i \rangle$$

$$\mathcal{N}_m^i = \sum_{h=1}^{m} \frac{I_h^2}{y_h^2} \langle X_h, e_i \rangle, \quad \mathcal{T}_m^i = \sum_{h=1}^{m} \frac{z_h I_h}{y_h^2} \langle X_h, e_i \rangle.$$

Similarly to Section 3.3.2.2, the continuous-time filtering processes $\gamma(\mathcal{J}_t^{ji} X_t)$, $\gamma(\mathcal{O}_t^i X_t)$, $\gamma(\mathcal{M}_t^i X_t)$, $\gamma(\mathcal{N}_t^i X_t)$ and $\gamma(\mathcal{T}_t^i X_t)$ in CTM-y will be approximated by the processes $\tilde{\gamma}(\mathcal{J}_m^{ji} X_m)$, $\tilde{\gamma}(\mathcal{O}_m^i X_m)$, $\tilde{\gamma}(\mathcal{M}_m^i X_m)$, $\tilde{\gamma}(\mathcal{N}_m^i X_m)$ and $\tilde{\gamma}(\mathcal{T}_m^i X_m)$, respectively, if they are determined as follows:

$$\tilde{\gamma}(\mathcal{J}_m^{ji} X_m) = \Psi_m \tilde{A} \tilde{\gamma}(\mathcal{J}_{m-1}^{ji} X_{m-1}) + \alpha_{ji} \langle q_{m-1}, e_i \rangle \psi_m^j e_j, \tag{4.66}$$
$$\tilde{\gamma}(\mathcal{O}_m^i X_m) = \Psi_m \tilde{A} \tilde{\gamma}(\mathcal{O}_{m-1}^i X_{m-1}) + \langle q_{m-1}, e_i \rangle \Psi_m \tilde{A} e_i, \tag{4.67}$$
$$\tilde{\gamma}(\mathcal{M}_m^i X_m) = \Psi_m \tilde{A} \tilde{\gamma}(\mathcal{M}_{m-1}^i X_{m-1}) + \frac{I_m}{y_m} \langle q_{m-1}, e_i \rangle \Psi_m \tilde{A} e_i, \tag{4.68}$$
$$\tilde{\gamma}(\mathcal{N}_m^i X_m) = \Psi_m \tilde{A} \tilde{\gamma}(\mathcal{N}_{m-1}^i X_{m-1}) + \left(\frac{I_m}{y_m}\right)^2 \langle q_{m-1}, e_i \rangle \Psi_m \tilde{A} e_i, \tag{4.69}$$
$$\tilde{\gamma}(\mathcal{T}_m^i X_m) = \Psi_m \tilde{A} \tilde{\gamma}(\mathcal{T}_{m-1}^i X_{m-1}) + \frac{z_m I_m}{y_m^2} \langle q_{m-1}, e_i \rangle \Psi_m \tilde{A} e_i, \tag{4.70}$$

with initial conditions $\tilde{\gamma}(\mathcal{J}_0^{ji} X_0) = \tilde{\gamma}(\mathcal{O}_0^i X_0) = \tilde{\gamma}(\mathcal{M}_0^i X_0) = \tilde{\gamma}(\mathcal{N}_0^i X_0) = \tilde{\gamma}(\mathcal{T}_0^i X_0) = \underline{0}$, where $\underline{0}$ denotes the column-vector of N zeros and (q_m) is the state filter approximation obtained from (4.65).

By using the discrete-time approximate filter processes above, updates, in our DTAM-y model, from ϵ_{ji} to ϵ'_{ji} $(1 \le j \ne i \le N)$ and from g_i to g_i $(i = 1, 2, \dots, N)$ are given by

$$\epsilon'_{ji} = \frac{\langle \tilde{\gamma}(\mathcal{J}_M^{ji} X_M), \underline{1} \rangle}{\langle \tilde{\gamma}(\mathcal{O}_M^i X_M), \underline{1} \rangle}, \tag{4.71}$$
$$g'_i = \frac{\langle \tilde{\gamma}(\mathcal{T}_M^i X_M), \underline{1} \rangle + \langle \tilde{\gamma}(\mathcal{M}_M^i X_M), \underline{1} \rangle}{\langle \tilde{\gamma}(\mathcal{N}_M^i X_M), \underline{1} \rangle}. \tag{4.72}$$

4.4.3 Estimation of the noise variance

In DTAM-k model, ML estimation of the noise variance can be obtained by using Radon-Nikodym derivative [James 1996]:

$$\frac{dP_{\theta'}}{dP_\theta} = \prod_{h=1}^{M} \frac{1}{\frac{1}{\sqrt{\Delta}} y_h \langle X_h, \sigma' \rangle \sqrt{2\pi}} \exp\left\{ -\frac{(z_h - I_h \langle X_h, g' \rangle + y_h)^2}{2 \frac{(y_h \langle X_h, \sigma' \rangle)^2}{\Delta}} \right\}$$
$$\times \left[\frac{1}{\frac{1}{\sqrt{\Delta}} y_h \langle X_h, \sigma \rangle \sqrt{2\pi}} \exp\left\{ -\frac{(z_h - I_h \langle X_h, g \rangle + y_h)^2}{2 \frac{(y_h \langle X_h, \sigma \rangle)^2}{\Delta}} \right\} \right]^{-1}. \tag{4.73}$$

The conditional expectation of the log-likelihood function is determined by

$$\begin{aligned}
Q(\theta',\theta) =& E_\theta\left\{\log\left(\frac{dP_{\theta'}}{dP_\theta}\right)|\mathcal{Y}^z_m\right\}\\
=& E_\theta\Bigg\{\sum_{h=1}^m\Bigg[-\log\langle X_h,\sigma'\rangle-\frac{1}{2}\frac{\Delta}{(y_h\langle X_h,\sigma'\rangle)^2}\big(z_h^2+I_h^2\langle X_h,g'\rangle^2\\
&+y_h^2-2z_hI_h\langle X_h,g'\rangle+2z_hy_h-2y_hI_h\langle X_h,g'\rangle\big)\Bigg]\Bigg|\mathcal{Y}^z_m\Bigg\}+R(\theta,\mathcal{Y}^z_m)\\
=& E_\theta\Bigg\{\sum_{h=1}^m\Bigg[-\sum_{i=1}^N\langle X_h,e_i\rangle\log\sigma'_i-\frac{\Delta}{2}\sum_{i=1}^N\langle X_h,e_i\rangle\frac{1}{\sigma_i'^2}\Big(\frac{z_h^2}{y_h^2}+\frac{I_h^2}{y_h^2}g_i'^2\\
&+1-2\frac{z_hI_h}{y_h^2}g'_i+2\frac{z_h}{y_h}-2\frac{I_h}{y_h}g'_i\Big)\Bigg]\Bigg|\mathcal{Y}^z_m\Bigg\}+R(\theta,\mathcal{Y}^z_m)\\
=& \sum_{i=1}^N\Bigg[-\tilde{\pi}(\mathcal{O}^i_m)\log\sigma'_i-\frac{\Delta}{2}\frac{1}{\sigma_i'^2}\Big(\tilde{\pi}(\mathcal{L}^i_m)+\tilde{\pi}(\mathcal{N}^i_m)g_i'^2\\
&+\tilde{\pi}(\mathcal{O}^i_m)-2\tilde{\pi}(\mathcal{T}^i_m)g'_i+2\tilde{\pi}(\mathcal{Q}^i_m)-2\tilde{\pi}(\mathcal{M}^i_m)g'_i\Big)\Bigg]+R(\theta,\mathcal{Y}^z_M)
\end{aligned}$$

where the function $R(\theta,\mathcal{Y}^z_m)$ does not depend on θ' and $\tilde{\pi}(H_m)=E_\theta(H_m|\mathcal{Y}^z_m)$ for $H_m\equiv\mathcal{L}^i_m$, $\mathcal{N}^i_m$, $\mathcal{O}^i_m$, $\mathcal{T}^i_m$, $\mathcal{Q}^i_m$ and $\mathcal{M}^i_m$ $(i=1,2,\ldots,N, m=1,2,\ldots,M)$ where:

$$\mathcal{L}^i_m=\sum_{h=1}^m\frac{z_h^2}{y_h^2}\langle X_h,e_i\rangle,\quad \mathcal{Q}^i_m=\sum_{h=1}^m\frac{z_h}{y_h}\langle X_h,e_i\rangle.$$

Taking the derivative of $Q(\theta',\theta)$ with respect to σ'_i $(i=1,2,\ldots,N)$, we obtain

$$\begin{aligned}
\frac{\partial}{\partial\sigma'_i}Q(\theta',\theta)=&-\frac{1}{\sigma'_i}\tilde{\pi}(\mathcal{O}^i_m)+\frac{\Delta}{\sigma_i'^3}\Big[\tilde{\pi}(\mathcal{L}^i_m)+\tilde{\pi}(\mathcal{N}^i_m)g_i'^2\\
&+\tilde{\pi}(\mathcal{O}^i_m)-2\tilde{\pi}(\mathcal{T}^i_m)g'_i+2\tilde{\pi}(\mathcal{Q}^i_m)-2\tilde{\pi}(\mathcal{M}^i_m)g'_i\Big].
\end{aligned}$$

Now $\frac{\partial}{\partial\sigma'_i}Q(\theta',\theta)=0$ becomes

$$\begin{aligned}
\sigma_i'^2=&\frac{\Delta}{\tilde{\pi}(\mathcal{O}^i_m)}\Big[\tilde{\pi}(\mathcal{L}^i_m)+\tilde{\pi}(\mathcal{N}^i_m)g_i'^2\\
&+\tilde{\pi}(\mathcal{O}^i_m)-2\tilde{\pi}(\mathcal{T}^i_m)g'_i+2\tilde{\pi}(\mathcal{Q}^i_m)-2\tilde{\pi}(\mathcal{M}^i_m)g'_i\Big],\quad i=1,2,\ldots,N.
\end{aligned}\tag{4.74}$$

Note that

$$\tilde{\pi}(H_m)=\frac{\langle\tilde{\gamma}(H_mX_m),\underline{1}\rangle}{\langle p_m,\underline{1}\rangle},\quad \text{for } H_m\equiv\mathcal{L}^i_m,\mathcal{N}^i_m,\mathcal{O}^i_m,\mathcal{T}^i_m,\mathcal{Q}^i_m \text{ and } \mathcal{M}^i_m,\tag{4.75}$$

From (4.75), it is seen that (4.74) (up to time M) becomes:

$$\begin{aligned}
\sigma_i'^2=&\frac{\Delta}{\langle\tilde{\gamma}(\mathcal{O}^i_MX_M),\underline{1}\rangle}\Big[\langle\tilde{\gamma}(\mathcal{L}^i_MX_M),\underline{1}\rangle+\langle\tilde{\gamma}(\mathcal{N}^i_MX_M),\underline{1}\rangle g_i'^2+\langle\tilde{\gamma}(\mathcal{O}^i_MX_M),\underline{1}\rangle\\
&-2\langle\tilde{\gamma}(\mathcal{T}^i_MX_M),\underline{1}\rangle g'_i+2\langle\tilde{\gamma}(\mathcal{Q}^i_MX_M),\underline{1}\rangle-2\langle\tilde{\gamma}(\mathcal{M}^i_MX_M),\underline{1}\rangle g'_i\Big],\ i=1,2,\ldots,N.
\end{aligned}\tag{4.76}$$

where the discrete-time filter processes $\tilde{\gamma}(\mathcal{O}^i_M X_M)$, $\tilde{\gamma}(\mathcal{M}^i_M X_M)$, $\tilde{\gamma}(\mathcal{N}^i_M X_M)$, $\tilde{\gamma}(\mathcal{T}^i_M X_M)$ are obtained from recursive equations (4.67), (4.68), (4.69), (4.70), respectively, and the processes $\tilde{\gamma}(\mathcal{L}^i_M X_M)$, $\tilde{\gamma}(\mathcal{Q}^i_M X_M)$ are computed from

$$\tilde{\gamma}(\mathcal{L}^i_m X_m) = \Psi_m \tilde{A} \tilde{\gamma} \mathcal{L}^i_{m-1} X_{m-1}) + \left(\frac{z_m}{y_m}\right)^2 \langle q_{m-1}, e_i \rangle \Psi_m \tilde{A} e_i, \tag{4.77}$$

$$\tilde{\gamma}(\mathcal{Q}^i_m X_m) = \Psi_m \tilde{A} \tilde{\gamma}(\mathcal{Q}^i_{m-1} X_{m-1}) + \frac{z_m}{y_m} \langle q_{m-1}, e_i \rangle \Psi_m \tilde{A} e_i, \; m = 1, 2, \ldots, M, \tag{4.78}$$

with initial condition $\tilde{\gamma}(\mathcal{L}^i_0 X_0) = \tilde{\gamma}(\mathcal{Q}^i_0 X_0) = \underline{0}$.

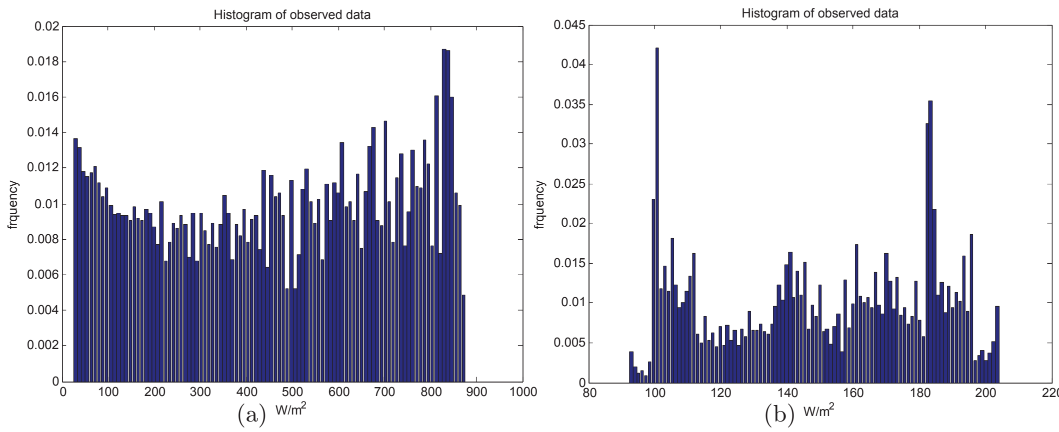


Figure 4.2: Histograms of observed data: (a) y-DATA-I.4; (b) y-DATA-III.4.

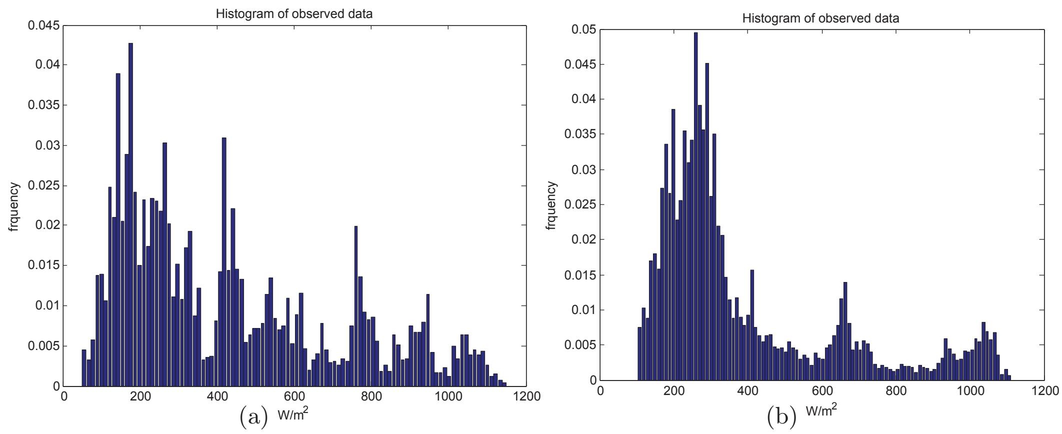


Figure 4.3: Histograms of observed data: (a) y-DATA-II.1; (b) y-DATA-II.2.

4.5 Experiments with real data

We proceed now to parameter estimations with all the total solar radiation sequences mentioned in Table 3.3. Various shapes of data histograms being observed, we tried different numbers of states: $N = 2, 4, 6, \ldots$.

The sequence of ML estimates $\{\theta^{(p)}, p = 1, 2, \ldots\}$ generated by EM algorithm fastly converges to a stability point when choosing reasonable initial parameter vectors and correct number of states.

Here are some illustrations when dealing with two 4 hours-long data sequences:

- y-DATA-I.4, a total solar radiation sequence of type I (clear day) recorded from 06:00 to 10:00 during the 262th-day, 2006, in Guadeloupe. Clearness index mean and variance are 0.7564, 0.0014, repectively

- y-DATA-III.4, a total solar radiation sequence of type III (cloudy day) recorded from 13:00 to 17:00 during the 180th-day, 2006, in Guadeloupe. Clearness index mean and variance are 0.1456, 0.0018, respectively.

Similarly, we proceed with two 1 hour-long sequences of type II (partially cloudy day): y-DATA-II.1, y-DATA-II.2 (see Table 3.3).

Observing these data histograms (Figure 4.2 and Figure 4.3), we have decided that the number of states is $N = 4$.

Consider the two data sequences: y-DATA-I.4 (Figure 4.7(a)), y-DATA-III.4

(Figure 4.8(a)) and start with initial parameter vector $\theta^{(0)}$:

$$A^{(0)} = \begin{pmatrix} -30 & 13 & 7 & 8 \\ 11 & -39 & 7 & 10 \\ 9 & 13 & -20 & 9 \\ 10 & 13 & 6 & -27 \end{pmatrix},$$
$$g^{(0)} = (0.1, 0.5, 0.7, 0.9)',$$
$$\sigma^{(0)} = (0.2, 2, 4, 6)'.$$

The parameters of the CTM-y will be approximated by the estimates of its DTAM-y from the initial parameter vector $\{P^{(0)}, g^{(0)}, \sigma^{(0)}\}$, where

$$P^{(0)} = \mathbf{I} + \Delta A^{(0)} = \begin{pmatrix} 0.9917 & 0.0036 & 0.0019 & 0.0022 \\ 0.0031 & 0.9892 & 0.0019 & 0.0028 \\ 0.0025 & 0.0036 & 0.9944 & 0.0025 \\ 0.0027 & 0.0036 & 0.0017 & 0.9925 \end{pmatrix},$$

and $\Delta = 1/3600$.

The estimated parameters after 100 iterations are (from the 30-th iteration, the estimated parameters are quite similar, see Figure 4.4, Figure 4.5 and Figure 4.6):

1. For y-DATA-I.4:

$$P^{(100)} = \begin{pmatrix} 0.9661 & 0.0000 & 0.0000 & 0.0000 \\ 0.0339 & 0.9718 & 0.0000 & 0.0000 \\ 0.0000 & 0.0282 & 0.9812 & 0.0000 \\ 0.0000 & 0.0000 & 0.0188 & 1.0000 \end{pmatrix},$$
$$g^{(100)} = (20.5595, 15.2412, 12.6386, 1.5645)',$$
$$\sigma^{(100)} = (0.5107, 0.4960, 0.4993, 0.1346)'.$$

2. For y-DATA-III.4:

$$P^{(100)} = \begin{pmatrix} 0.9905 & 0.0000 & 0.0000 & 0.0007 \\ 0.0000 & 0.9833 & 0.0013 & 0.0000 \\ 0.0000 & 0.0167 & 0.9934 & 0.0013 \\ 0.0095 & 0.0000 & 0.0053 & 0.9980 \end{pmatrix},$$
$$g^{(100)} = (-0.7715, 1.7319, 0.6673, 0.0638)',$$
$$\sigma^{(100)} = (0.0923, 0.0881, 0.0782, 0.0551)'.$$

The graphs in Figure 4.4, Figure 4.5 and Figure 4.6 show the estimates evolution for y-DATA-I.4 and y-DATA-III.4.

Using the estimated parameters, we generate some simulations for y-DATA-I.4, y-DATA-III.4 which are shown in Figure 4.7(b) and Figure 4.8(b), respectively.

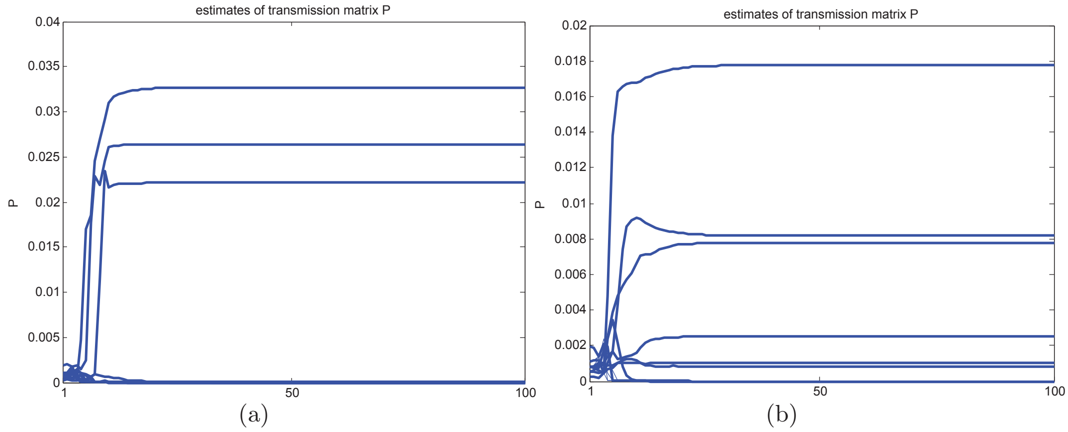


Figure 4.4: Estimation of transition probability matrix P: (a) from y-DATA-I.4; (b) from y-DATA-III.4.

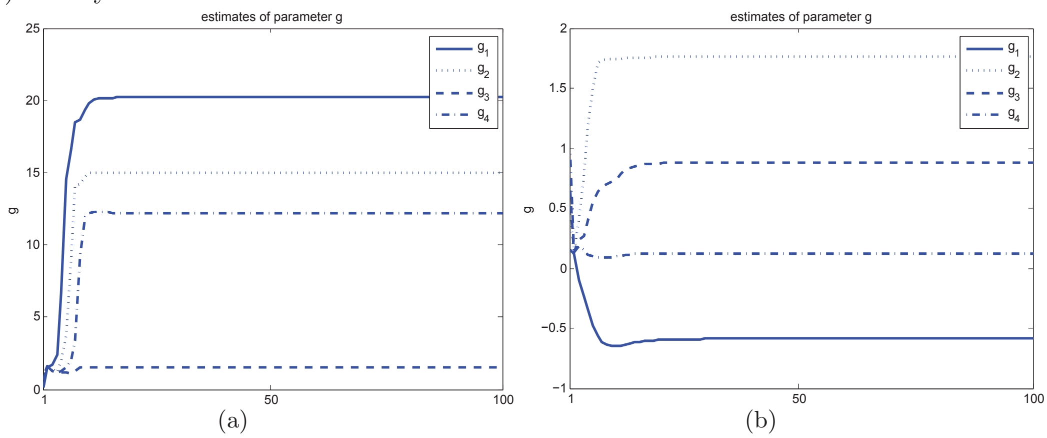


Figure 4.5: Estimation of parameter vector g: (a) from y-DATA-I.4; (b) from y-DATA-III.4.

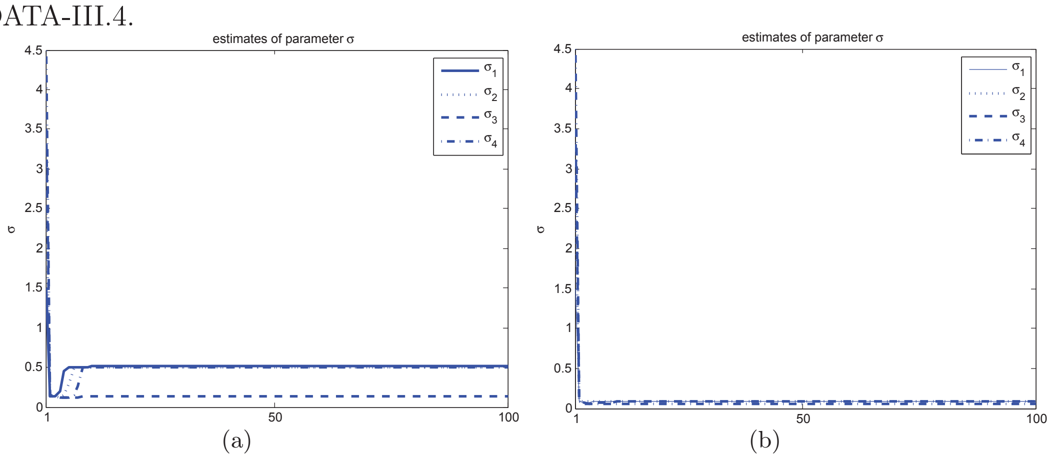
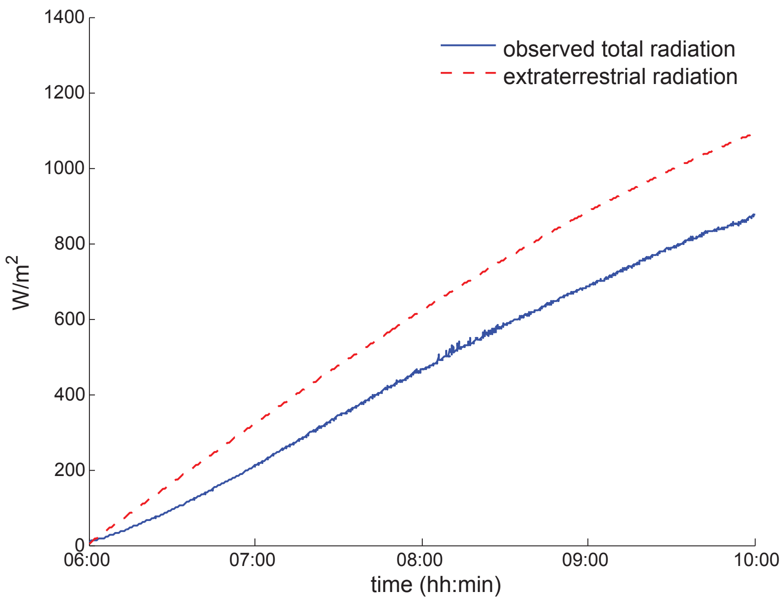
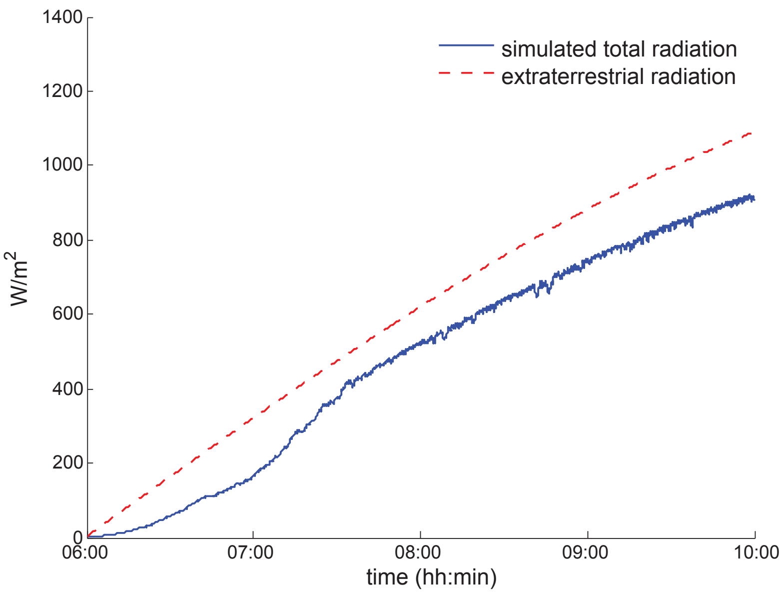


Figure 4.6: Estimation of parameter vector σ: (a) from y-DATA-I.4; (b) from y-DATA-III.4.

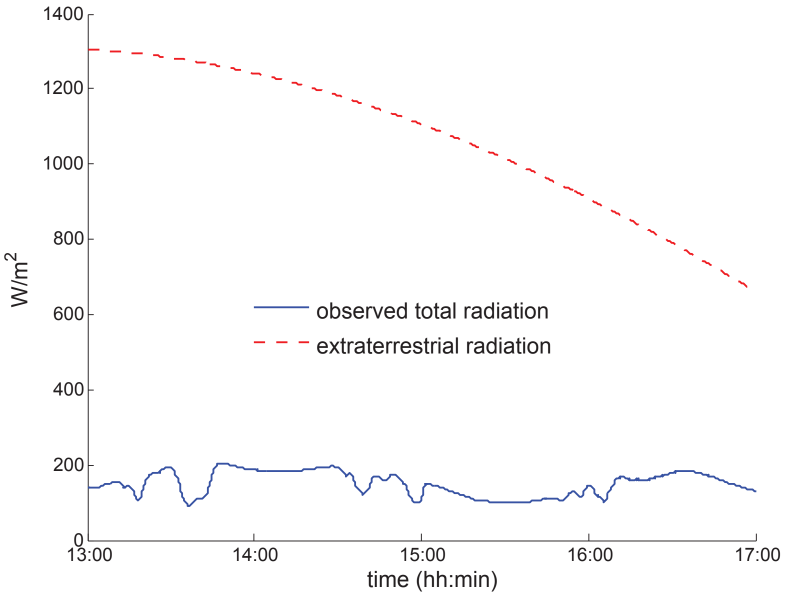


(a)

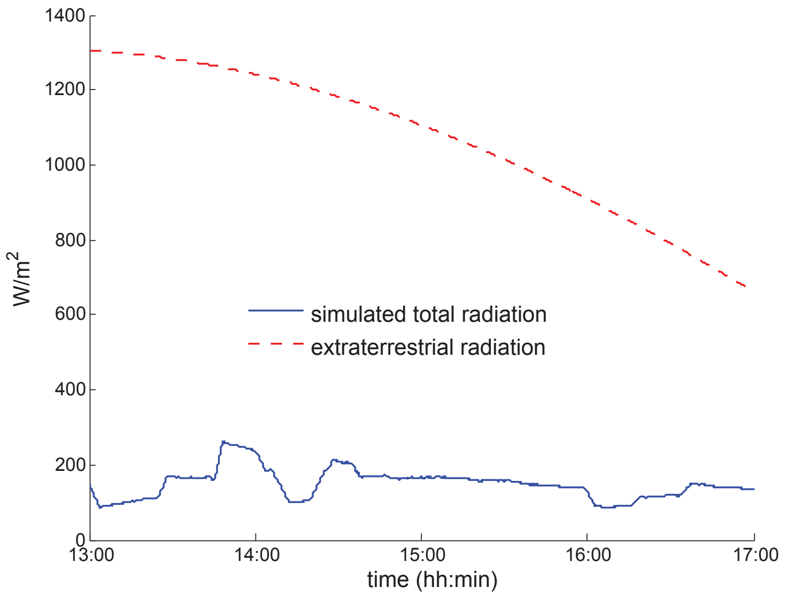


(b)

Figure 4.7: (a) Graph of observed data y-DATA-I.4; (b) A simulation for y-DATA-I.4.

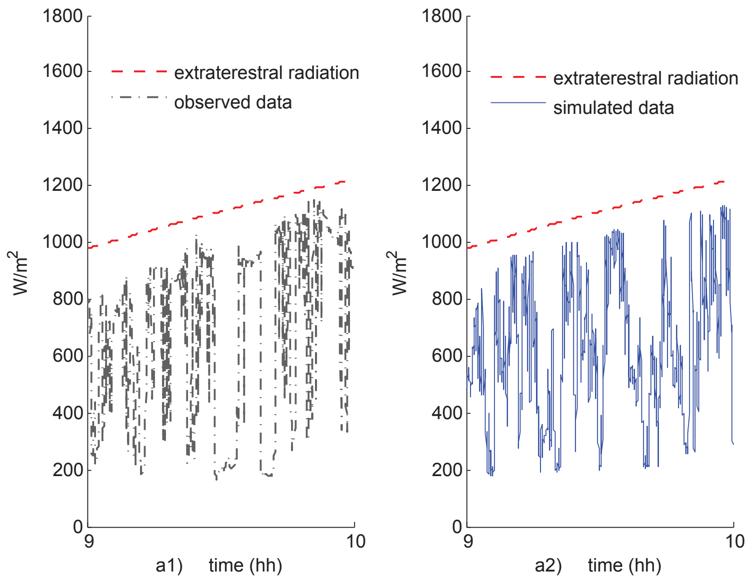


(a)

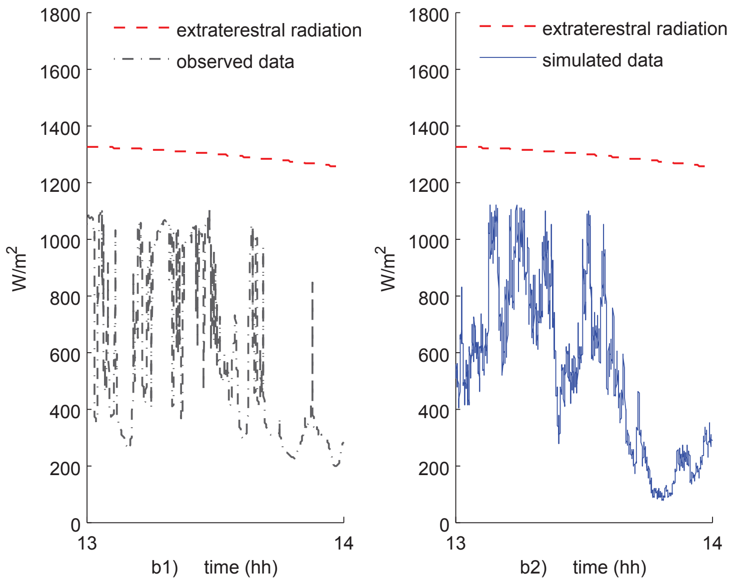


(b)

Figure 4.8: (a) Graph of observed data y-DATA-III.4; (b) A simulation for y-DATA-III.4.



(a)



(b)

Figure 4.9: (a) Graph of observed data y-DATA-II.1, a1), and its simulation, a2); (b) Graph of observed data y-DATA-II.2, b1), and its simulation, b2).

We now consider numerical estimations based on the data sequences y-DATA-II.1 (Figure 4.9a1)) and y-DATA-II.2 (Figure 4.9b1)). We started with the following initial parameters:

$$A^{(0)} = \begin{pmatrix} -7 & 4 & 3 & 2 \\ 3 & -13 & 3 & 4 \\ 3 & 2 & -9 & 4 \\ 1 & 7 & 3 & -10 \end{pmatrix},$$
$$g^{(0)} = (0.15, 0.45, 0.65, 0.95)',$$
$$\sigma^{(0)} = (1.5, 2.5, 3.5, 4.5)'.$$

and the matrix:

$$P^{(0)} = \mathbf{I} + \Delta A^{(0)} = \begin{pmatrix} 0.9981 & 0.0011 & 0.0008 & 0.0006 \\ 0.0008 & 0.9964 & 0.0008 & 0.0011 \\ 0.0008 & 0.0006 & 0.9975 & 0.0011 \\ 0.0003 & 0.0019 & 0.0008 & 0.9972 \end{pmatrix},$$

where $\Delta = 1/3600$.

The estimates obtained after 100 iterations:

1. From y-DATA-II.1:

$$P^{(100)} = \begin{pmatrix} 0.9418 & 0.1219 & 0.0000 & 0.1186 \\ 0.0212 & 0.8781 & 0.0060 & 0.0000 \\ 0.0070 & 0.0000 & 0.9843 & 0.0000 \\ 0.0300 & 0.0000 & 0.0096 & 0.8814 \end{pmatrix},$$
$$g^{(100)} = (-30.7737, 92.3213, 3.5591, 94.4960)',$$
$$\sigma^{(100)} = (3.4343, 3.6536, 1.3474, 3.6932).$$

2. From y-DATA-II.2:

$$P^{(100)} = \begin{pmatrix} 0.9980 & 0.0011 & 0.0008 & 0.0006 \\ 0.0008 & 0.9964 & 0.0008 & 0.0011 \\ 0.0009 & 0.0006 & 0.9975 & 0.0011 \\ 0.0003 & 0.0019 & 0.0008 & 0.9972 \end{pmatrix},$$
$$g^{(100)} = (0.9527, 0.9378, 0.9022, 0.9583)',$$
$$\sigma^{(100)} = (2.4123, 2.4157, 2.4168, 2.4193)'.$$

The evolutions of these estimates are shown in Figure 4.10, Figure 4.11 and Figure 4.12.

Using the estimated parameters, we generate some simulations for y-DATA-II.1, y-DATA-II.2 shown in Figure 4.9a2) and Figure 4.9b2), respectively.

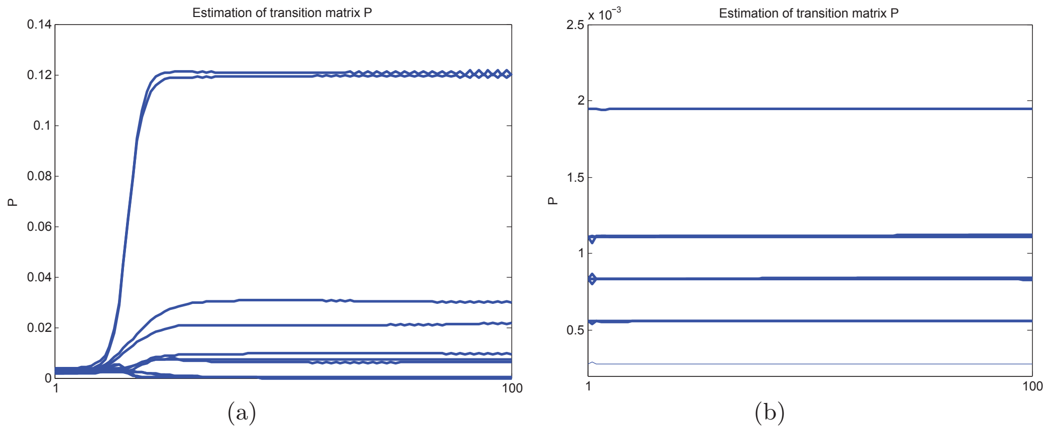


Figure 4.10: Estimation of transition probability matrix P: (a) from y-DATA-II.1; (b) from y-DATA-II.2.

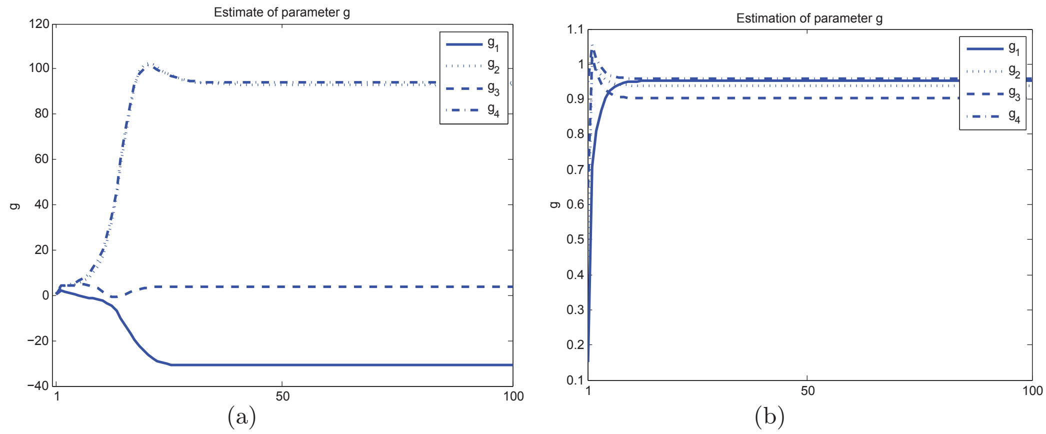


Figure 4.11: Estimation of parameter vector g: (a) from y-DATA-II.1; (b) from y-DATA-II.2.

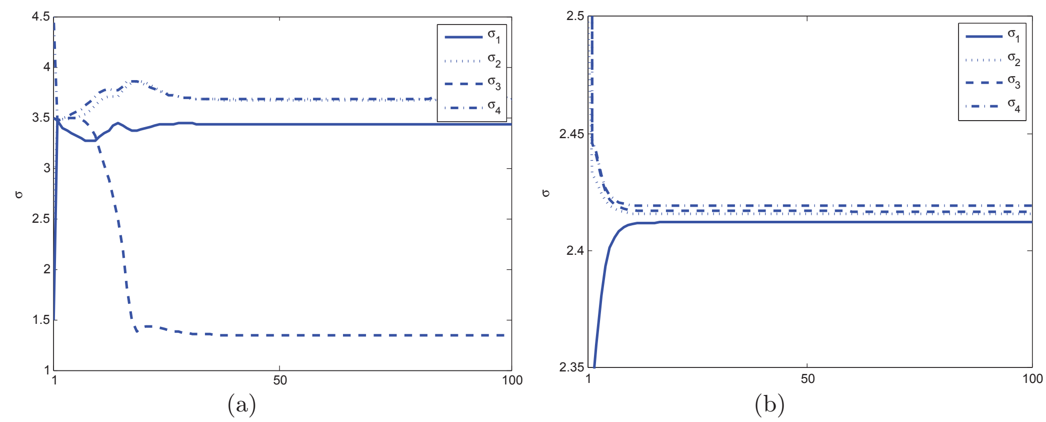
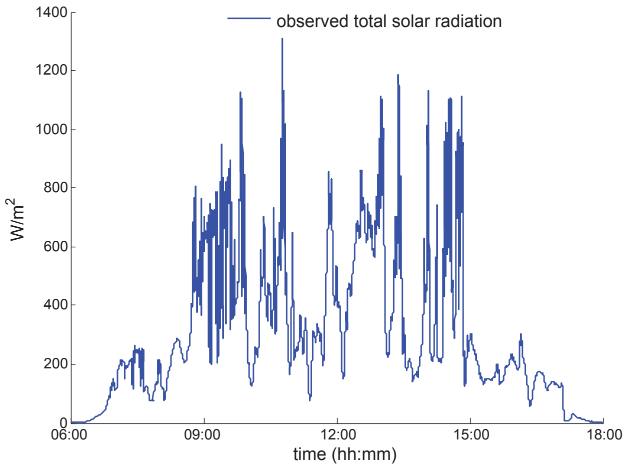


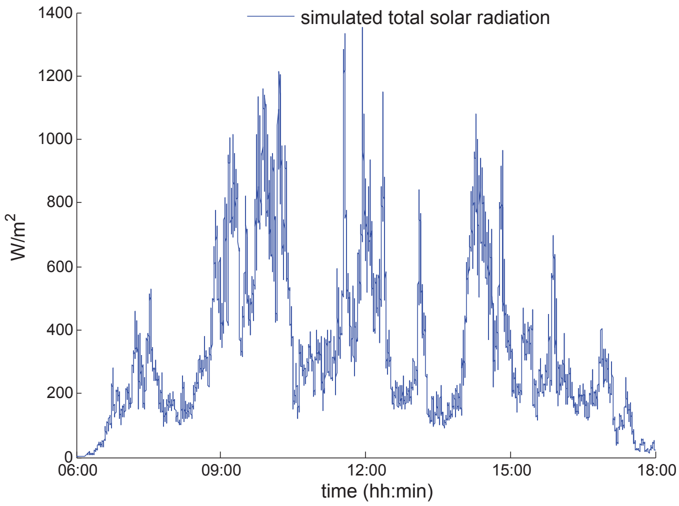
Figure 4.12: Estimation of parameter vector σ: (a) from y-DATA-II.1; (b) from y-DATA-II.2.

4.6 Simulations of total solar radiation day

Applying similar methods, we have obtained the simulations shown in Figure 4.13(b), Figure 4.14(b) and Figure 4.15(b) for observed data y-DATA-day-I, y-DATA-day-II and y-DATA-day-III whose graphs are plotted in Figure 4.13(a), Figure 4.14(a) and Figure 4.15(a), respectively.



(a)



(b)

Figure 4.13: (a) Graph of observed data y-DATA-day-I; (b) A simulation for y-DATA-day-I.

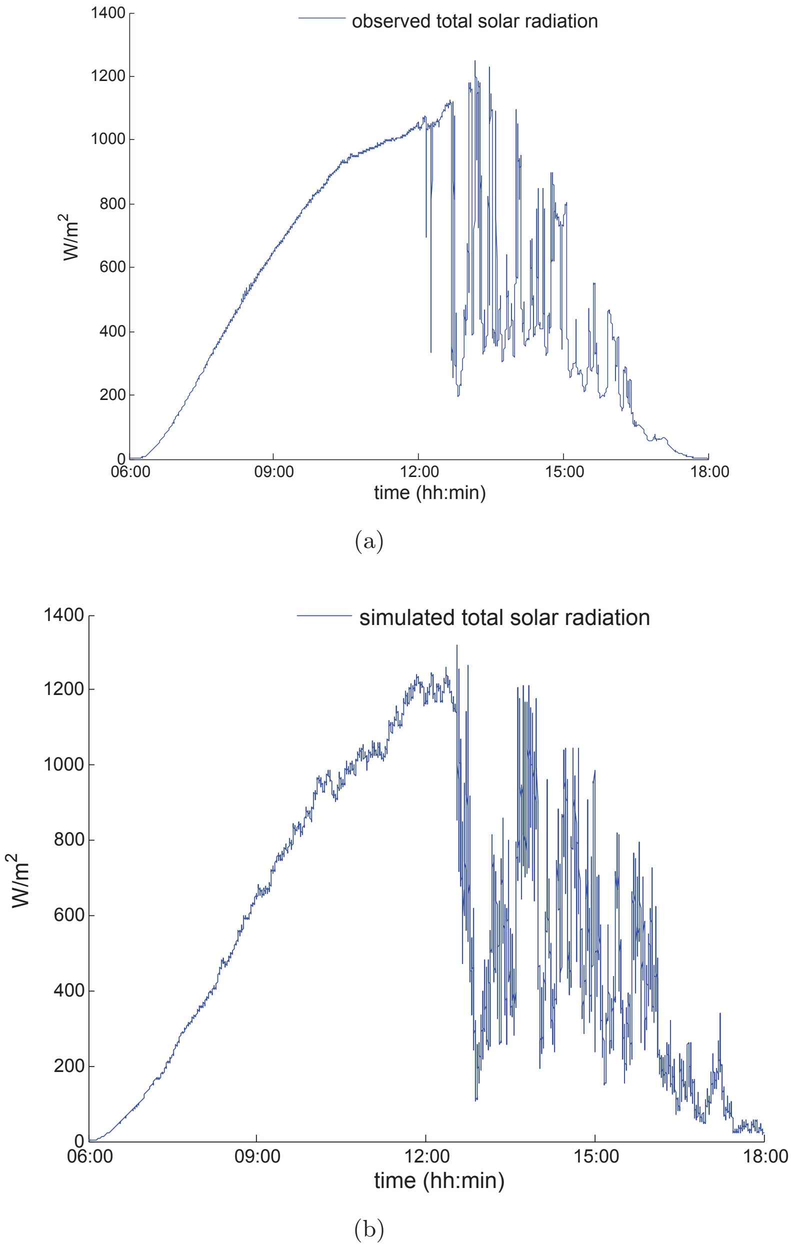
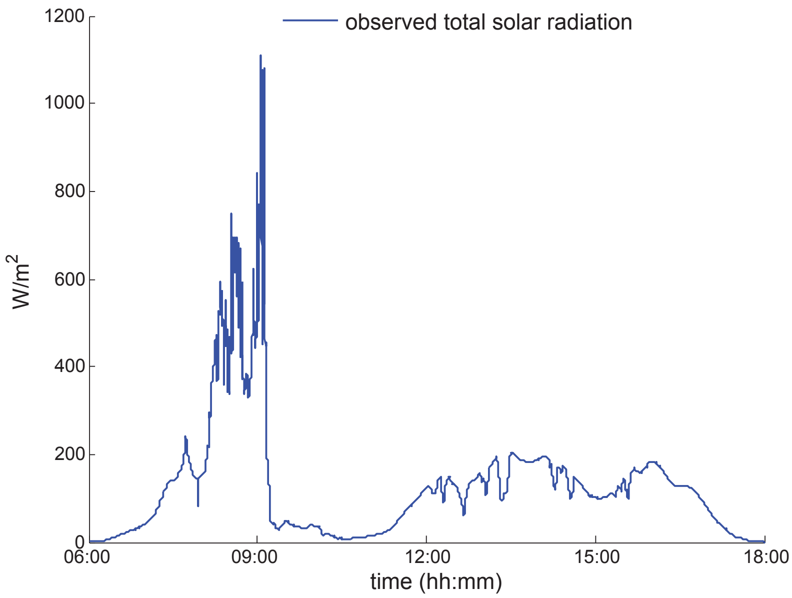
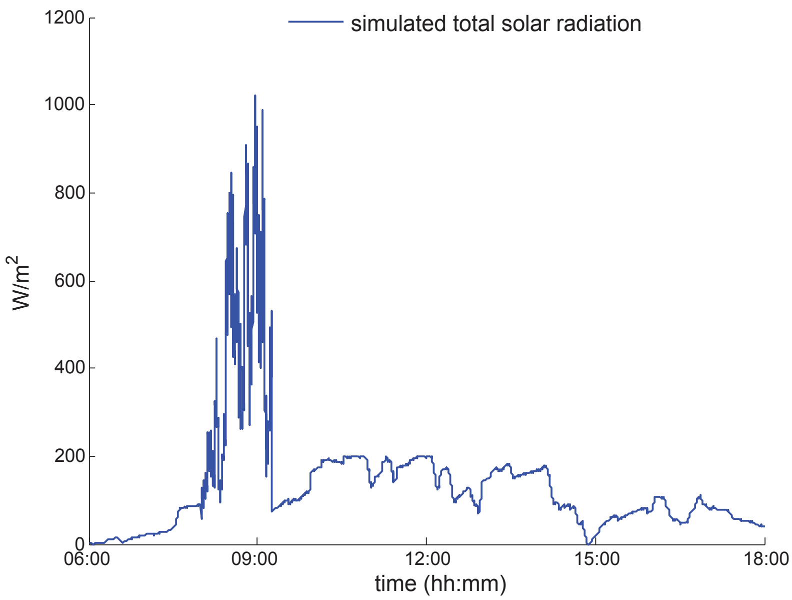


Figure 4.14: (a) Graph of observed data y-DATA-day-II; (b) A simulation for y-DATA-day-II.



(a)



(b)

Figure 4.15: (a) Graph of observed data y-DATA-day-III; (b) A simulation for y-DATA-day-III.

Chapter 5

Some applications using our models

Contents

Résumé

Ce chapitre présente deux applications utilisant des trajectoires simulées par nos modèles : amélioration de l'estimation de la distribution de l'indice journalier de clarté et prédiction du rayonnement solaire total.

Pour la première application, nous utilisons DTM-K estimé avec les données mesurées à l'île de La R éunion, et npEM un algorithme d'estimation de mélanges.

Pour la seconde application, nous utilisons CTM-k et CTM-y estimés par des données mesurées à 1Hz à l'île de La Guadeloupe et nous proposons une prédiction à un horizon d'une heure.

Abstract

This chapter presents two applications using paths simulated by our models: improvement of the daily clearness index distribution estimation and prediction of total solar radiation.

For the first application, we use DTM-K estimated from La R éunion island data and npEM a mixture estimation algorithm.

For the second application, we use CTM-k and CTM-y estimated from data sampled at 1Hz in Guadeloupe and we propose a prediction with horizon one hour.

5.1 Estimating the distribution of daily clearness index

Estimating the distribution of daily clearness index over a month or over a specific period can be of interest for deciding wether our model estimated over this period still works on a longer period or not. It can also be used for clustering daily CISs observed on various periods.

Indeed, using DTM-K with its parameter estimated from a sample of daily CISs, of one-month-length say, (Table 3.1), we can simulate a much larger n-sample of K_h, say $K'_1, \dots, K'_n$, over this period and get a smooth estimation of the pdf over this month. Doing the same with another month and getting another n-sample, a KS (Kolmogorov-Smirnov) test can be performed to reject or not the hypothesis that both pdf are the same. If the hypothesis is rejected (w.r.t. a p-value), we can reject the hypothesis that both models are the same. On the other hand, KS distance between two sequences, computed from the two n-samples, can be used for clustering CISs by performing some standard clustering methods.

5.1.1 Kernel estimators

The Gaussian kernel estimator of the density is the function $\widehat{f}_\delta$ defined as

$$\widehat{f}_\delta(x) = \frac{1}{\delta n} \sum_{h=1}^{n} \phi\left(\frac{x - K'_h}{\delta}\right), \tag{5.1}$$

where $\delta > 0$ is a *bandwidth* (a smoothing parameter) and $\phi(\cdot)$ denotes the $\mathcal{N}(0,1)$ density function kernel.

This estimator is of course much smoother than the uniform kernel estimator, that is the empirical pdf $\widehat{f}$, defined as follows: divide $[0,1]$ interval (the range of K_h) into L sub-intervals $]x_{l-1}, x_l]$ of equal length $\Delta x = \frac{1}{L}$ with $x_0 = 0$ and $x_l = l\Delta x, l = 1, 2, \ldots, L$, then

$$\widehat{f}(x) = \frac{n_l}{n}\frac{1}{\Delta x}, \tag{5.2}$$

where n_l is the number of observed values in the interval $]x_{l-1}, x_l], l = 1, 2, \ldots, L$.

5.1.2 Mixtures of nonparametric densities

The estimator of the preceding section can be improved by assuming that the density is a finite mixture of nonparametric densities, which means that the components of the mixture are completely unspecified. While such a mixture is identifiable and can be estimated by an EM-type algorithm in the parametric case (for example in the case of a Gaussian mixture, Figure 5.1), this model, in the nonparametric case, is not necessarily identifiable but can be estimated by the npEM (nonparametric EM) algorithm [Benaglia 2009] which is implemented in the *npEM* fuction of the R package *mixtools*.

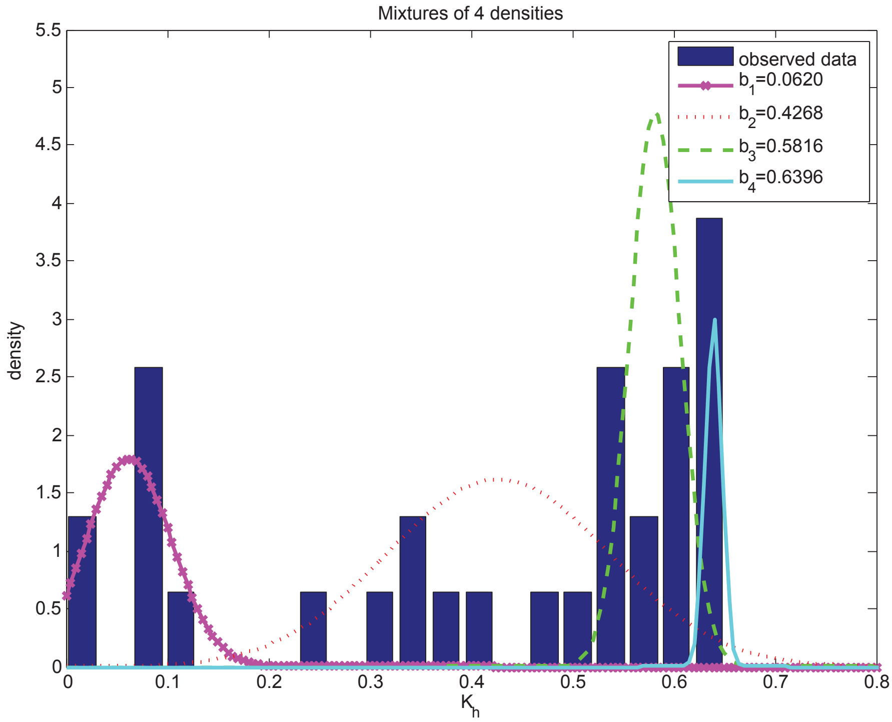
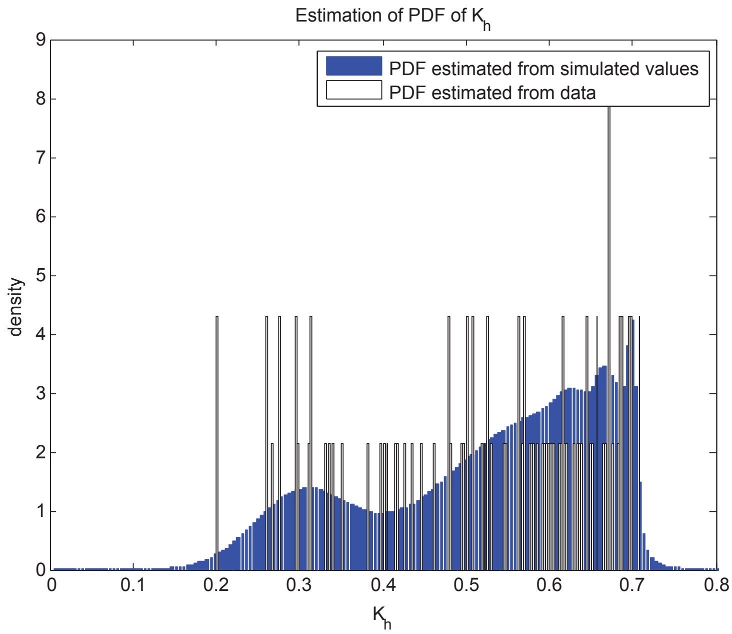


Figure 5.1: pdf of La Réunion K_h in March, estimated as a mixture by npEM algorithm

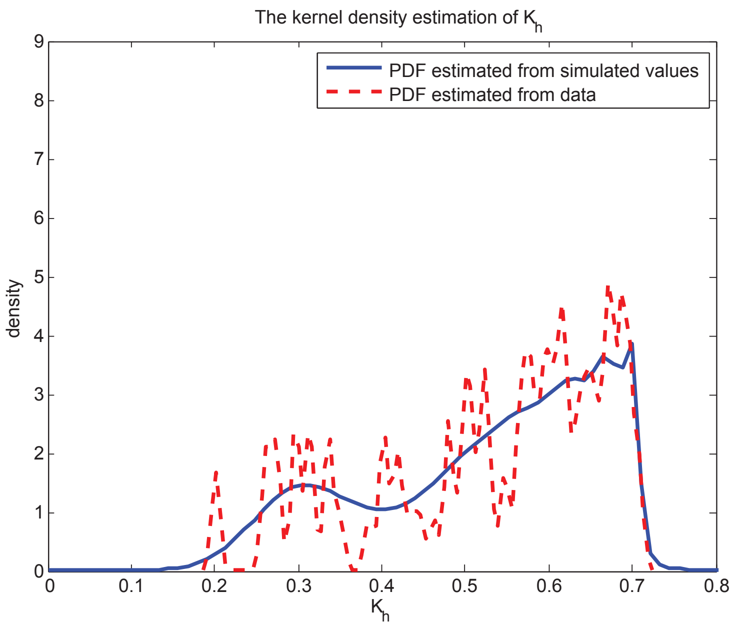
5.1.3 Experiments

From each of the three daily CISs observed in La Réunion island during October 2009, 2010 and 2011, respectively, we have estimated the parameters of the three corresponding DTM-k models. We have generated 50000 simulated paths of 31 values from each of these three estimated model. Then, from these $n = 3\times 50000\times 31$ values, we have estimated the pdf as shown in Figure 5.2 with different levels of smoothing parameter, *bandwidth* δ and Δt, see Figure 5.3 and Figure 5.2.

pdf of K_h were obtained similarly in March, April, May, August and September, see Figure 5.4, Figure 5.5 and Figure 5.6, respectively.

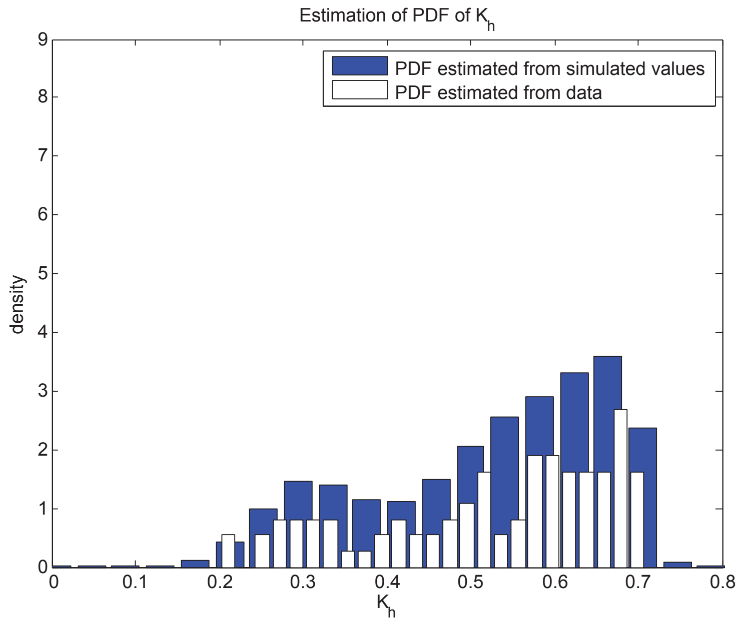


(a)

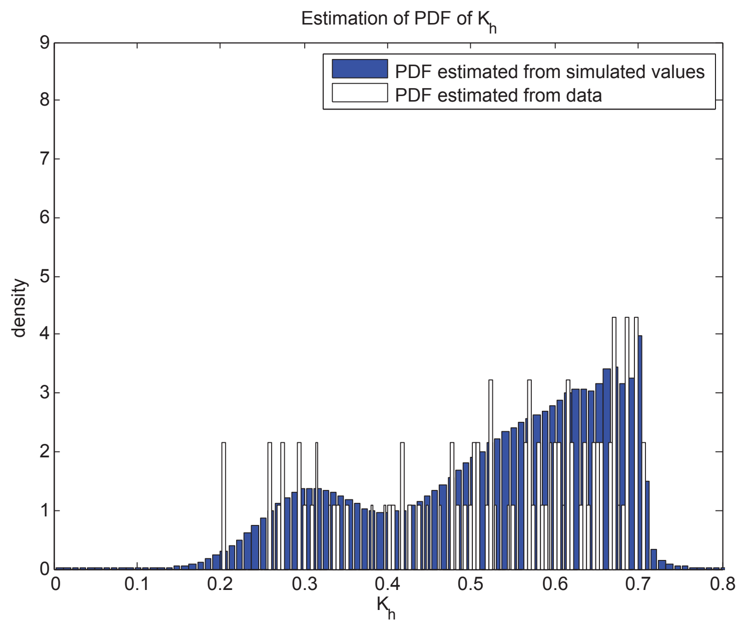


(b)

Figure 5.2: pdf of La Réunion K_h in October: a) observed data histogram with $\Delta x = 1/200$, (b) kernel estimation with bandwidth $\delta = 1/200$.

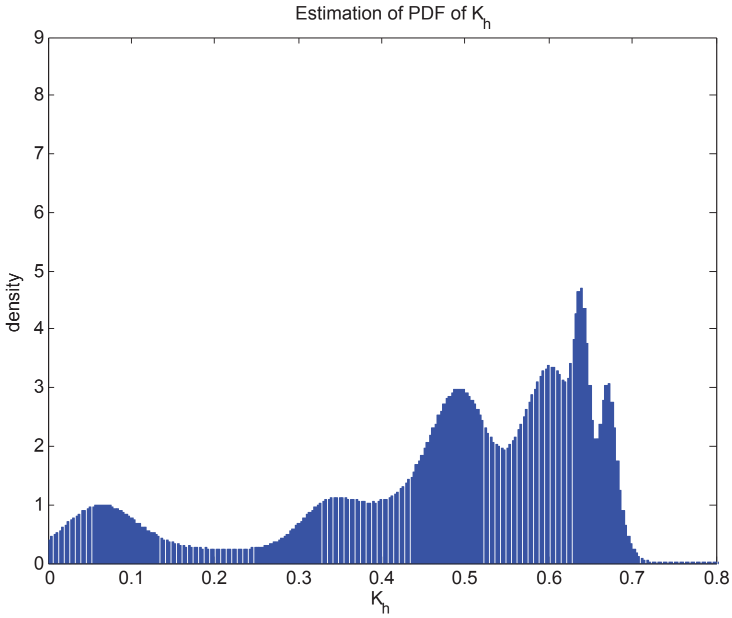


(a)

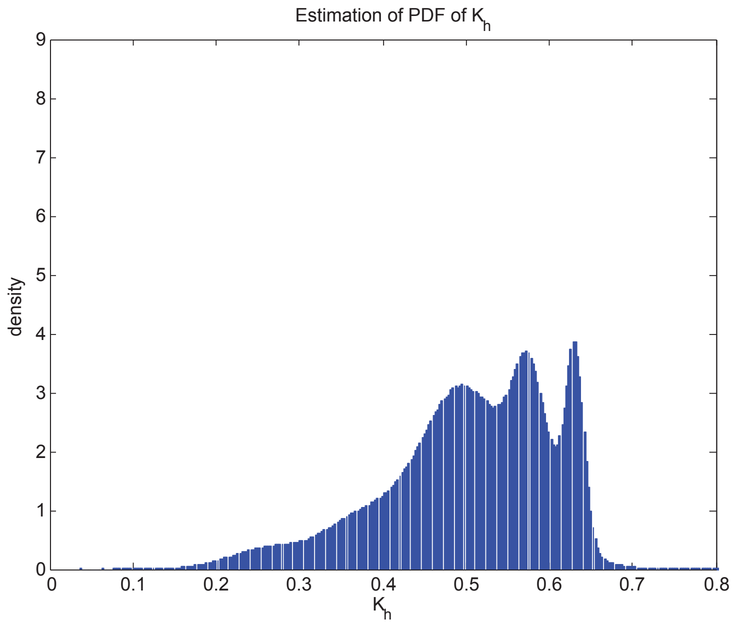


(b)

Figure 5.3: pdf of La Réunion K_h in October: (a) observed data with $\Delta x = 1/25$ (b) $\Delta x = 1/100$.



(a)



(b)

Figure 5.4: pdf of La Réunion K_h from simulated values (with $\Delta x = 1/300$) : (a) in March, (b) in August.

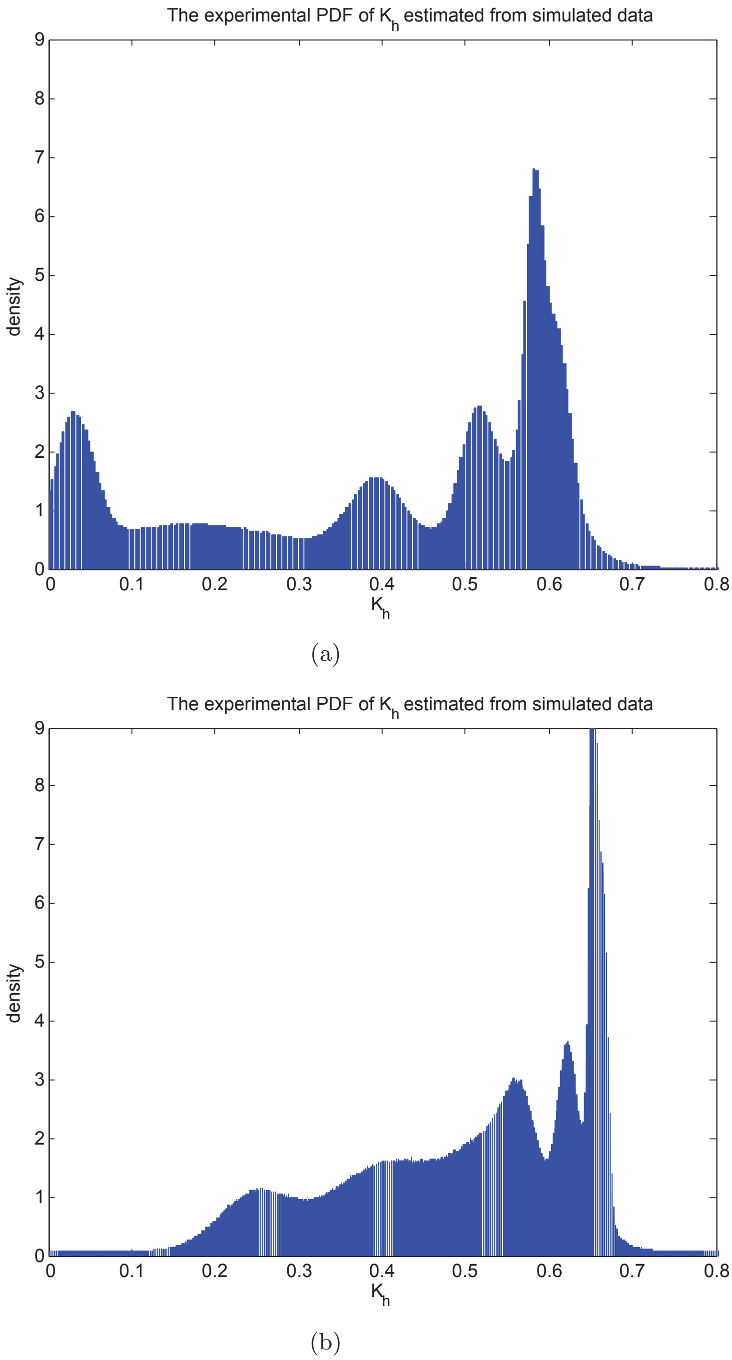
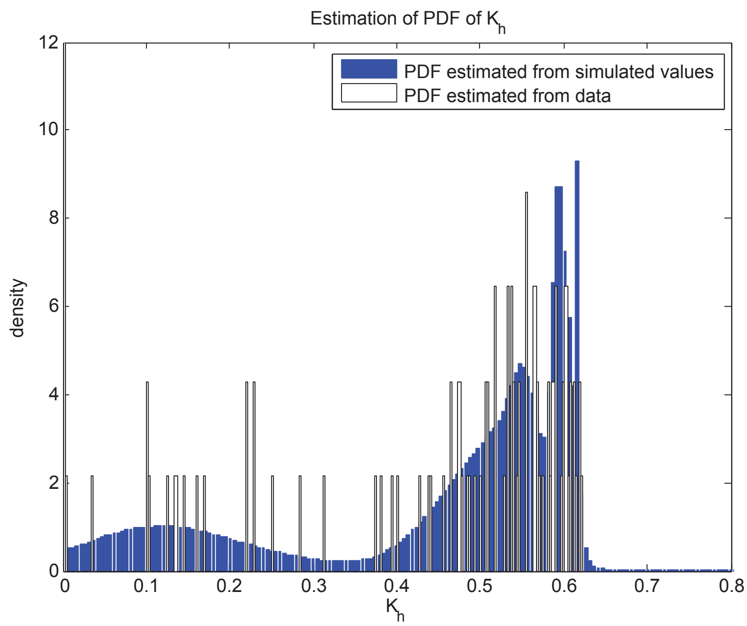
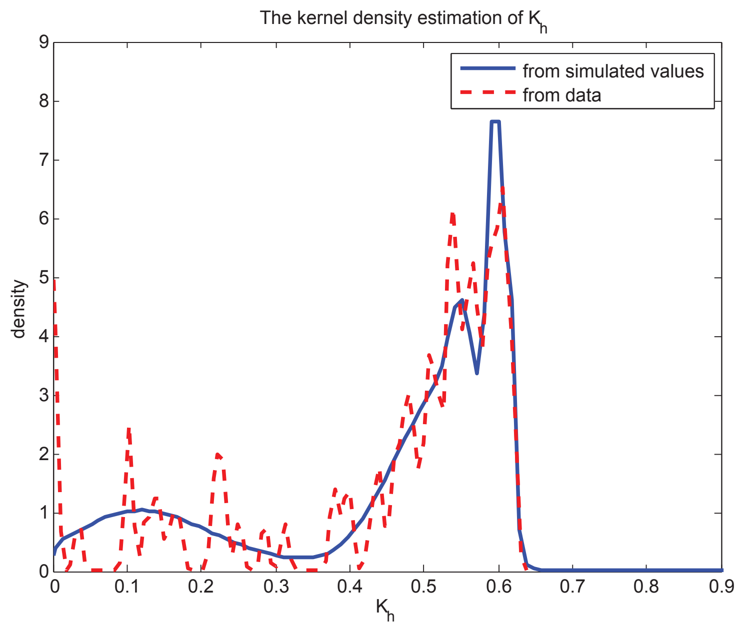


Figure 5.5: pdf of La Réunion K_h estimated from simulated values (with $\Delta x = 1/300$): (a) in April; (b) in September.



(a)



(b)

Figure 5.6: pdf of K_h for May in La Réunion: (a) observed values with $\Delta x = 1/200$; (b) kernel estimation with $\delta = 1/200$.

5.2 Prediction

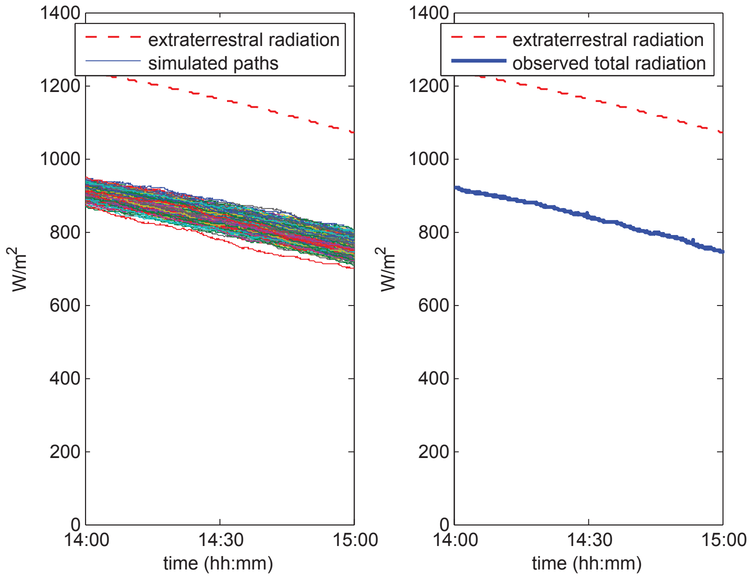
In this Section we implicitely assume that the model estimated on a time interval $[0, T]$ still works on the time interval $[T, T']$, with $T' > T$. Using the model with its estimated parameters and the extraterrestrial radiation computed on $[T, T']$, we generate a large number of paths in the period $[T, T']$. Prediction of total solar radiation on $[T, T']$ then follows.

5.2.1 Confidence region and prediction error for hourly total solar radiation

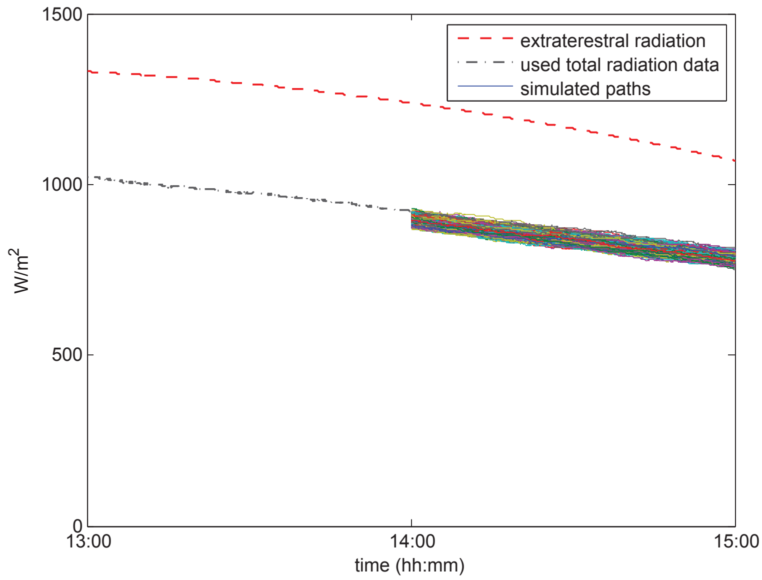
First, CTM-k and CTM-y parameters have been estimated from the 1 Hz measurements of total solar radiation observed during $T = 1$h. Extraterrestrial radiation has been computed on $[T, T']$ with $T' = 1$h. Using this, $L = 500$ paths have been simulated on $[T, T']$, providing a confidence region of prediction for the total radiation.

The error of prediction is usually evaluated by the MSPE (Mean Square Prediction Error), that is the expectation (estimated by a mean) of the squared difference between predicted values and observed values. MSPE has also to be averaged over the L paths, that can also provide a confidence interval. We prefer here to simplfy the prediction evaluation by computing, for each simulated path, the relative error between the mean of predicted values and the mean of obseved values and then provide the mean of these relative errors.

Observed datasets are mentioned in Table 3.3. The results are shown in Table 5.1. The results obtained from CTM-y seem better than that from CTM-k. Some cases of simulated paths are shown in Figure 5.7, Figure 5.8, Figure 5.9, Figure 5.11, Figure 5.13, Figure 5.15 and Figure 5.17.

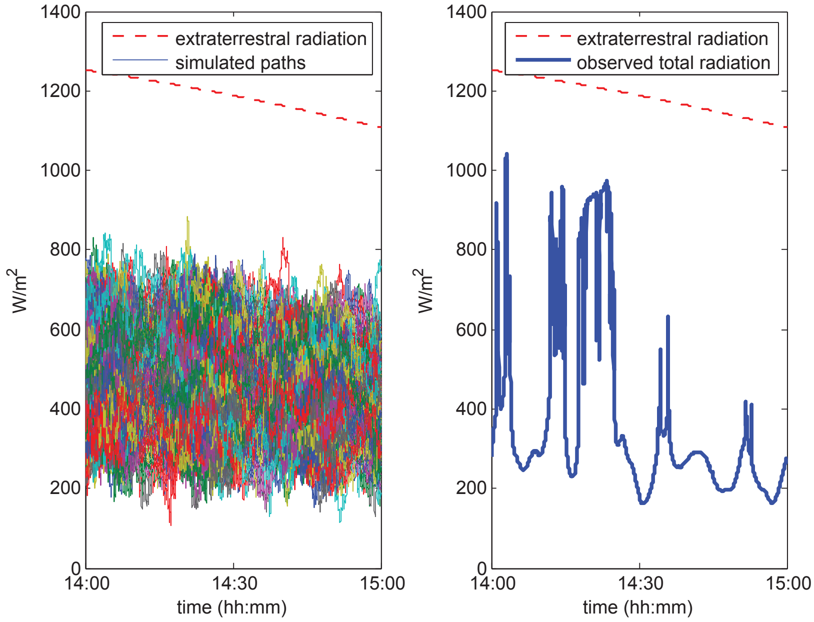


(a)

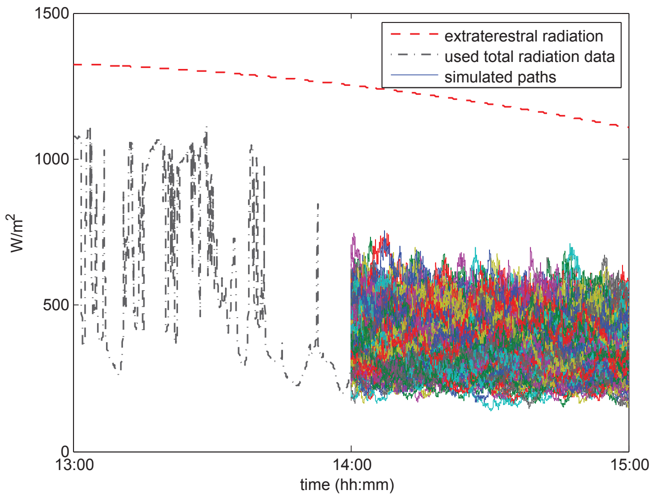


(b)

Figure 5.7: Paths simulated from 14h to 15h (a) using CTM-k with parameters estimated from k-DATA-I.2 (left), observed data (right), (b) using CTM-y with parameters estimated from y-DATA-I.2.

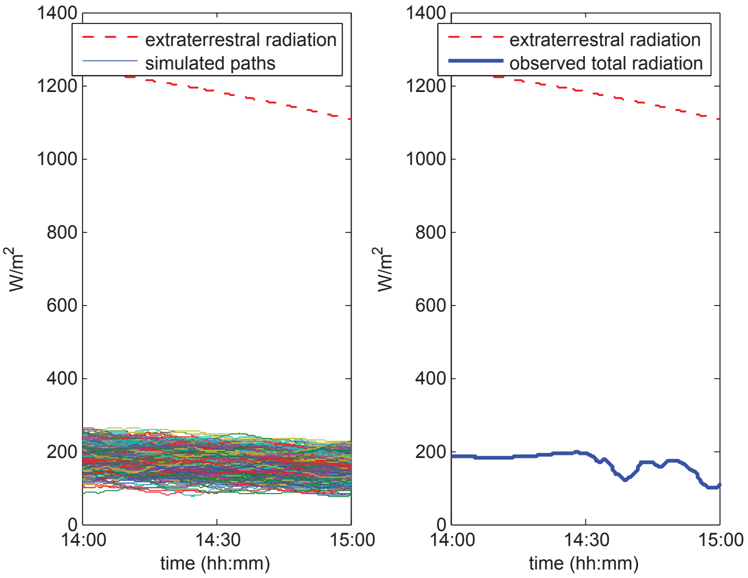


(a)

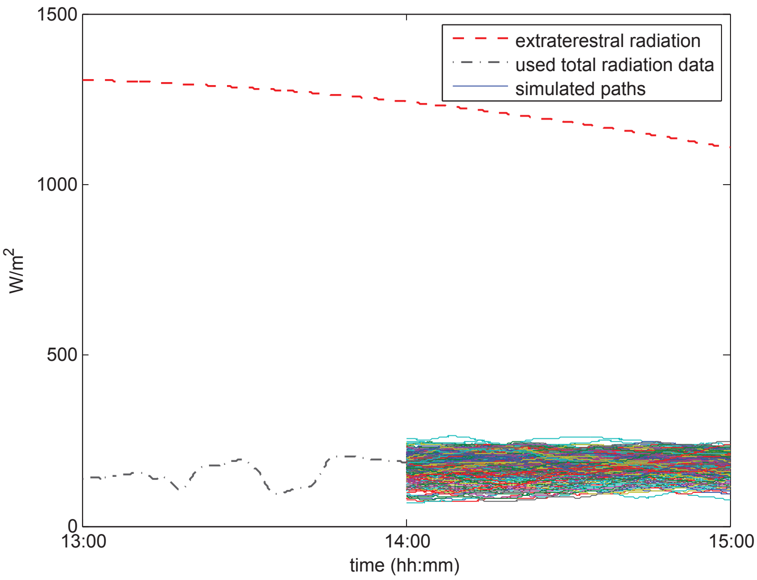


(b)

Figure 5.8: Paths simulated from 14h to 15h (a) using CTM-k with parameters estimated from k-DATA-II.2 (left), observed data (right), (b) using CTM-y with parameters estimated from y-DATA-II.2.



(a)



(b)

Figure 5.9: Paths simulated from 14h to 15h (a) using CTM-k with parameters estimated from k-DATA-III.3 (left), observed data (right), (b) using CTM-y with parameters estimated from y-DATA-III.3.

Data set	Mean obs. data (Wm^{-2})	95% C.I. from CTM-y (Wm^{-2})	err. (%)	95% C.I. from CTM-k (Wm^{-2})	err. (%)
DATA-I.1	937.0	(935.7,942.6)	0.29	(937.2,943.0)	0.33
DATA-I.2	838.6	(836.2,842.1)	0.07	(842.4,850.5)	0.94
DATA-I.3	802.0	(800.4,803.6)	0.52	(814.4,822.3)	1.51
DATA-II.1	797.2	(780.5,798.7)	0.75	(797.1,815.0)	1.10
DATA-II.2	370.5	(359.6,371.9)	1.28	(370.0,383.2)	1.64
DATA-II.3	488.5	(646.3,661.9)	33.9	(639.0,652.7)	32.2
DATA-III.1	210.2	(209.0,214.4)	0.72	(209.0,215.6)	1.01
DATA-III.2	249.5	(162.2,163.8)	34.7	(126.8,128.6)	48.8
DATA-III.3	169.6	(169.2,174.1)	1.18	(170.3,175.1)	1.81

Table 5.1: 95% confidence intervals for the mean total radiation during the hour next to T using the models CTM-k and CTM-y with parameters estimated from observations up to T.

5.2.2 Discussion on the prediction results

Our prediction method by simulating a large number of paths hinges on the strong law of large numbers.

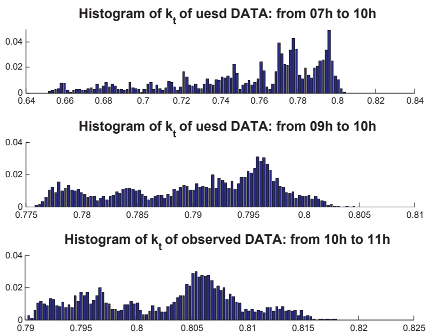
Predictions quality highly depends on the pdf, average and variance of the data observed during the learning time interval $[0, T]$ and used to estimate the parameters of the model. As it will be observed, prediction quality also depends on the number of states N of the model and on the distribution of the forecasting period data.

We illustrate this considering 4 examples of data recorded in Guadeloupe in 2006 and using CTM-y model. Predictions on 10h-11h are done from data observed in the interval 07h-10h (resp. 09h-10h).

Case 1: Measurements: day 262, type I (clear day).

Predictions and real value:

- Measurements: 07h-10h (DATA-07h10h), (k_t) average: $m_k = 0.76$, (k_t) variance $\sigma_k^2 = 0.0014$, 95% confidence interval for AVERAGE-10h11h: 95%C.I.07h10h $= (942.4, 948.6)$ W/m^2.
- Measurements: 09h-10h (DATA-09h10h), $m_k = 0.79$, $\sigma_k^2 = 0.0001$, 95%C.I.09h10h $= (935.7, 942.6)$ W/m^2.
- Forecasting period 10h-11h, Observed: $m_k = 0.80$, $\sigma_k^2 = 0.0001$, Average total radiation: 936.9 W/m^2.



Time	k_t m_k	k_t σ_k^2	95% C.I./ **Mean obs. data**	err. (%)
07h-10h	.76	.0014	(942.4,948.6)	0.91
09h-10h	.79	.0001	(935.7,942.6)	0.29
10h-11h	.80	.0001	**936.9** W/m^2	0.00

(simulated paths are plotted in Figure 5.10 and Figure 5.11)

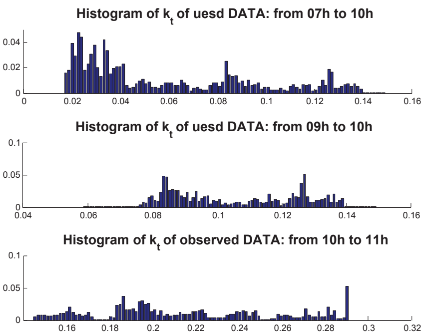
Comments:

- AVERAGE-10h11h does not belong to 95%C.I.07h10h, the relative error between the average of simulated data and the average of observed data is 0.91%), **probability distribution/histogram**, average, variance as well as the disposition of classes of DATA-07h10h and that of observed data in 10h-11h are different (see the figure and the table above).
- In contrast, 95%C.I.09h10h contains AVERAGE-10h11h (relative error is 0.29%) and similar statistical characteristics are observed.

Case 2: Measurements: day 285, type III (cloudy day).

Predictions and observed values: (see table below)

- Measurements: DATA-07h10h ($m_k = 0.06$, $\sigma_k^2 = 0.0015$). 95% confidence interval for AVERAGE-10h11h: $(118.3, 120.0)$ W/m^2.
- Measurements: DATA-09h10h ($m_k = 0.11$, $\sigma_k^2 = 0.0004$). 95% confidence for AVERAGE-10h11h: $(162.2, 163.8)$ W/m^2.
- Observed: AVERAGE-10h11h 249.5 W/m^2 ($m_k = 0.22$, $\sigma_k^2 = 0.0017$).

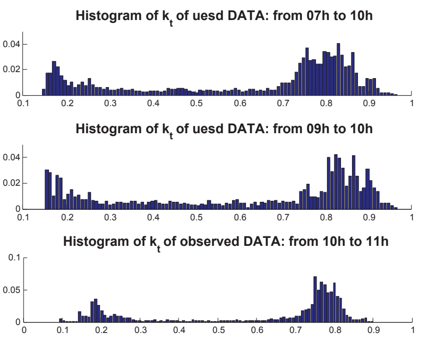


Time	k_t m_k	σ_k^2	95% C.I./ **Mean obs. data**	err. (%)
07h-10h	.06	.0015	(118.3,120.0)	52.3
09h-10h	.11	.0004	(162.2,163.8)	34.7
10h-11h	.22	.0017	**249.5** W/m^2	0.00

Simulated paths are plotted in Figure 5.12 and Figure 5.13

Comments: AVERAGE-10h11h does not belong to 95%C.I.07h10h (error: 52.3%) neither to 95%C.I.09h10h (error: 34.7%). Distribution characteristics of the data used for estimations are not similar to that of data in the prediction interval 10h-11h.

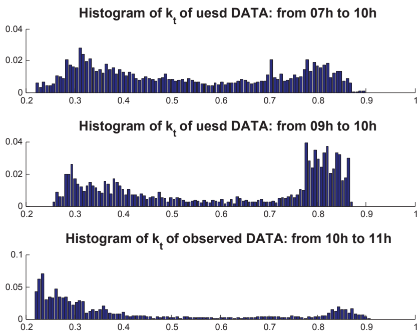
Case 3: Measurements: Day 118, type II (partially cloudy day).
Distribution characteristics of the data used for estimations (07h-10h, 09h10h) are similar to those of data observed in the prediction interval 10h-11h. The obtained predictions are quite nice (see the figure and the table below).



Time	k_t m_k	σ_k^2	95% C.I./ **Mean obs. data**	err. (%)
07h-10h	.63	.0643	(787.4,806.2)	0.06
09h-10h	.62	.0472	(780.5,798.7)	0.75
10h-11h	.61	.0627	**797.2** W/m^2	0.00

(Simulated paths are plotted in Figure 5.14 and Figure 5.15)

Case 4: Measurements Day 234, type II.
The distribution characteristics of data used for estimation (07h-10, 09h-10h) are not similar to that of data in the prediction interval 10h-11h. Predictions present large errors (see table and figure below).

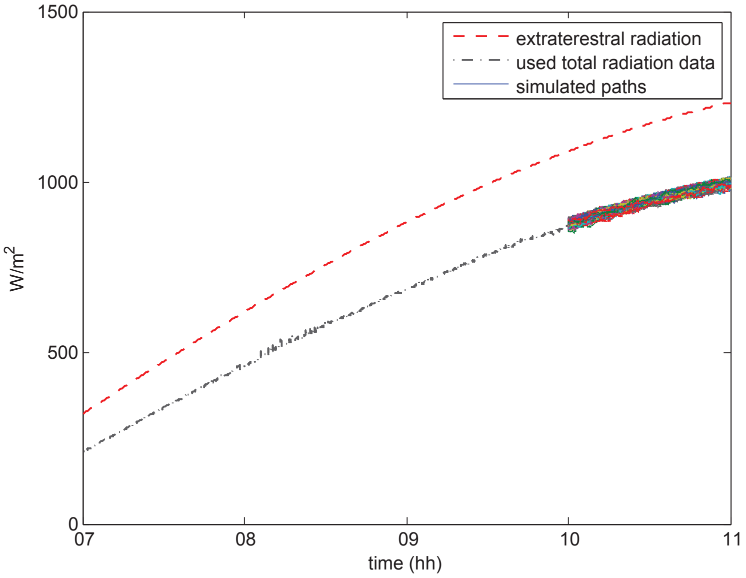


Time	k_t m_k	σ_k^2	95% C.I./ **Mean obs. data**	err. (%)
07h-10h	.54	.0397	(656.7,671.5)	36.0
09h-10h	.61	.0471	(646.3,661.9)	33.9
10h-11h	.42	.0499	**488.5** W/m^2	0.00

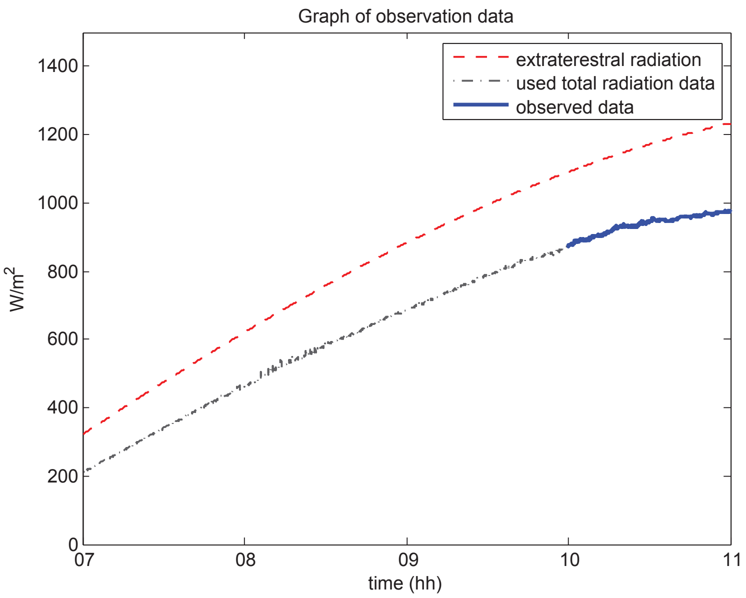
(Simulated paths are plotted in Figure 5.16 and Figure 5.17)

Our simulated data are distributed as the data used for estimating the models. So, if the meteorological conditions in the forecasting period change significantly (more clouds for instance), predictions (from CTM-y as well as CTM-k) do not work well.

Therfore we propose in our illustrations a short-horizon prediction interval $[T, T']$, $T' = T + 1$ (hour) and this horizon is also of interest for solar energy providers.

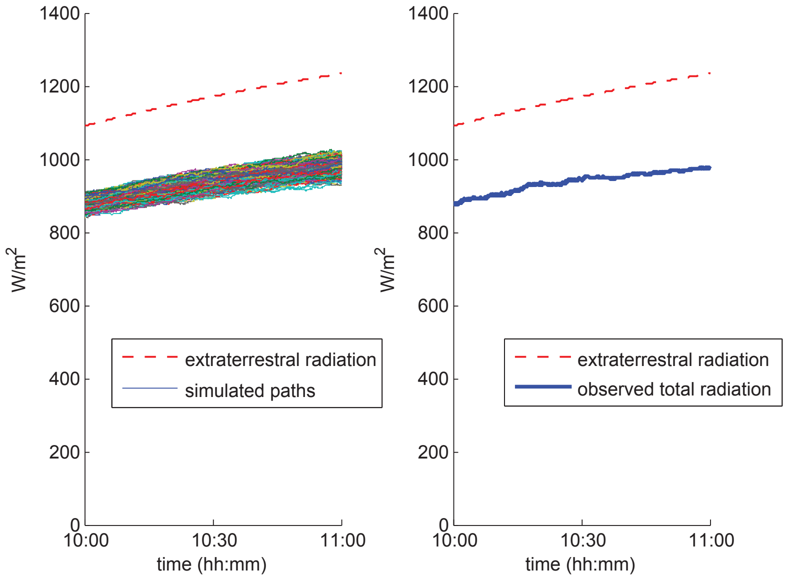


(a)

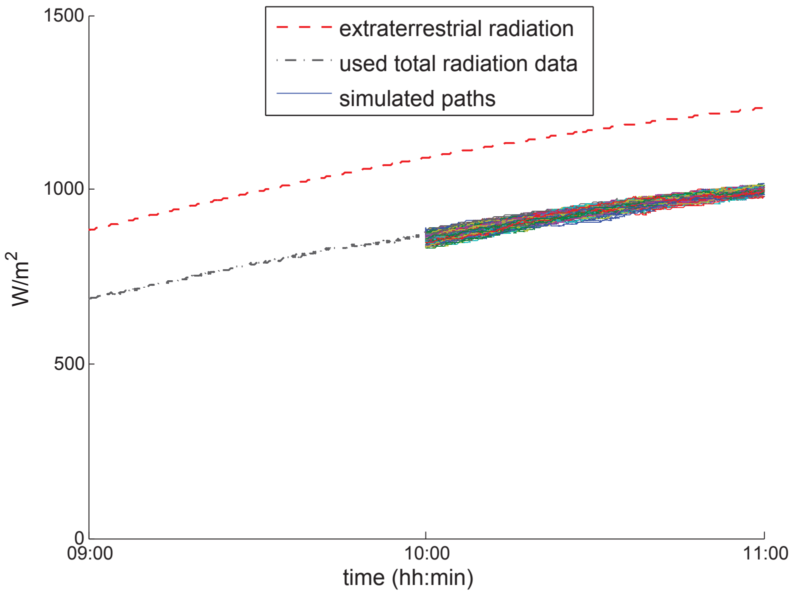


(b)

Figure 5.10: (a) Simulated paths in 10h-11h generated by CTM-y with parameters estimated from observations in 07h-10h, day 262 (type I, clear day), 2006, Guadeloupe. (b) Data observed during 10h-11h: blue solid line.

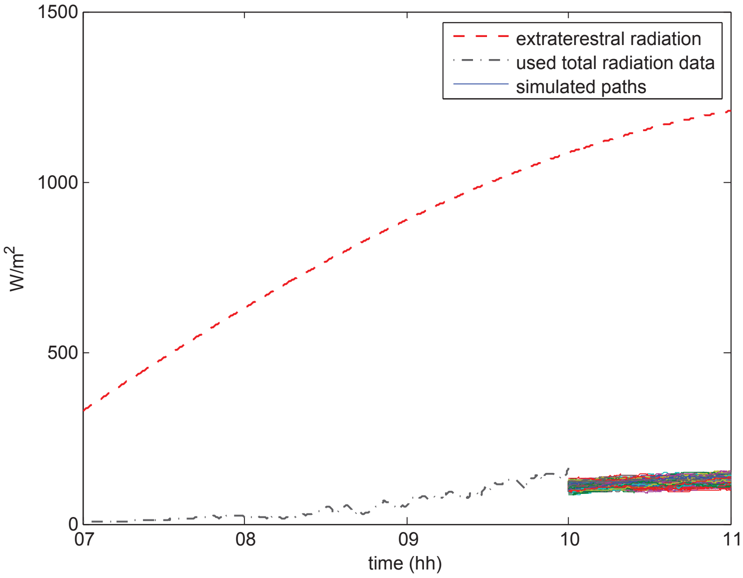


(a)

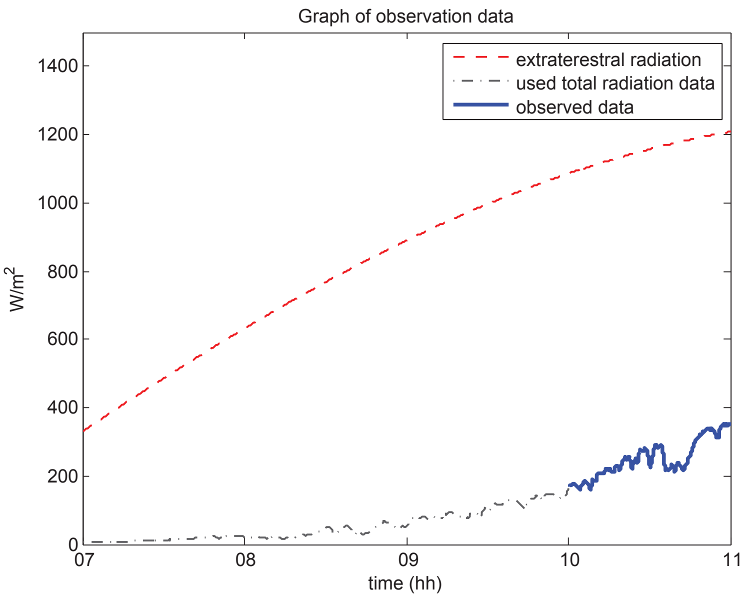


(b)

Figure 5.11: (a) left: Simulated paths in 10h-11h generated by CTM-k with parameters estimated from observations in 09h-10h, day 262 (type I, clear day), 2006, Guadeloupe (k-DATA-I.1), right: Observed data. (b) Simulated paths in 10h-11h generated by CTM-y with parameters estimated from y-DATA-I.1.

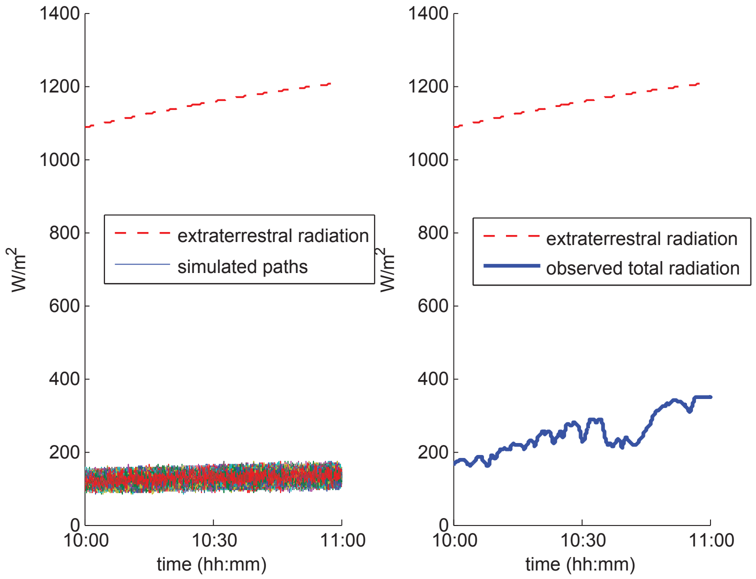


(a)

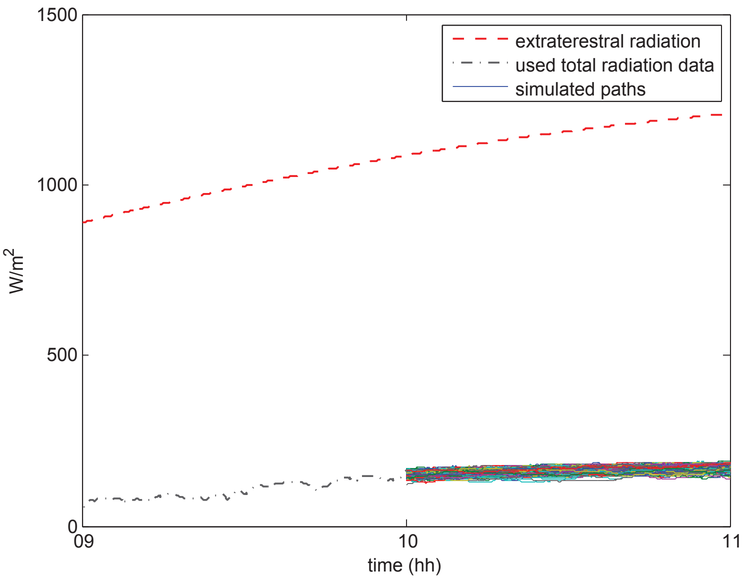


(b)

Figure 5.12: (a) Simulated paths in the next hour (10h-11h) generated by CTM-y with parameters estimated from observations in 07h-10h, day 28 (type III, cloudy day), 2006, Guadeloupe. (b) Observed data: blue solid line.

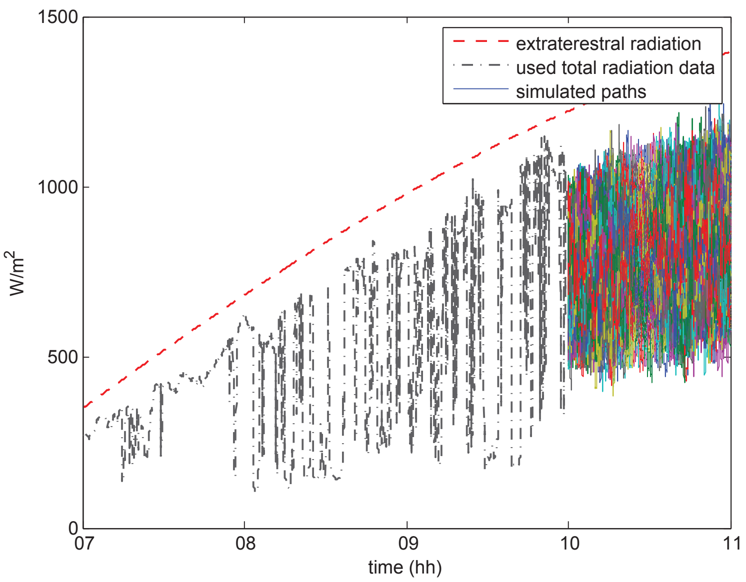


(a)

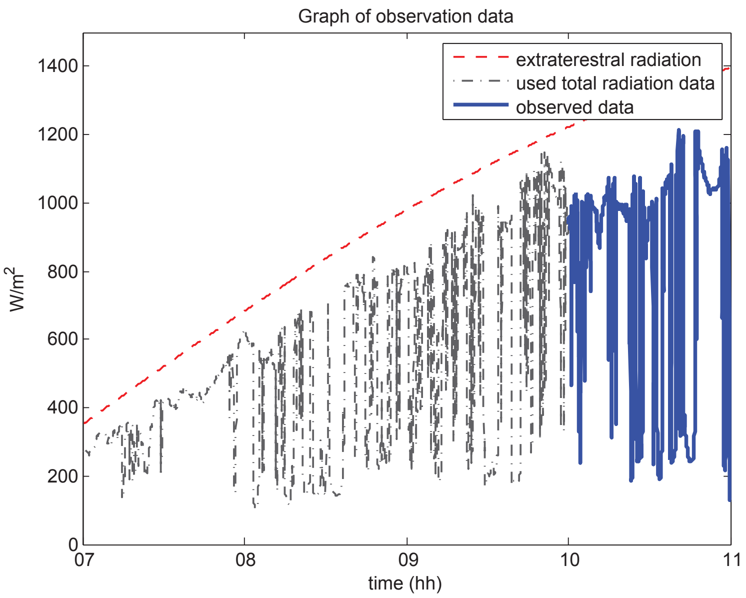


(b)

Figure 5.13: (a) left: Simulated paths in 10h-11h generated by CTM-k with parameters estimated from observations in 09h-10h, day 285 (type III, cloudy day), 2006, Guadeloupe (k-DATA-III.2), right: Observed data. (b) Simulated paths in 10h-11h generated by CTM-y with parameters estimated from y-DATA-III.2.

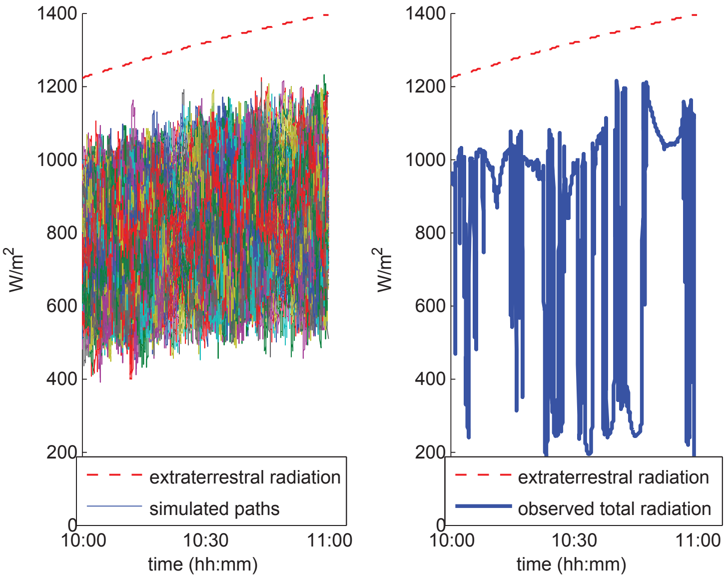


(a)

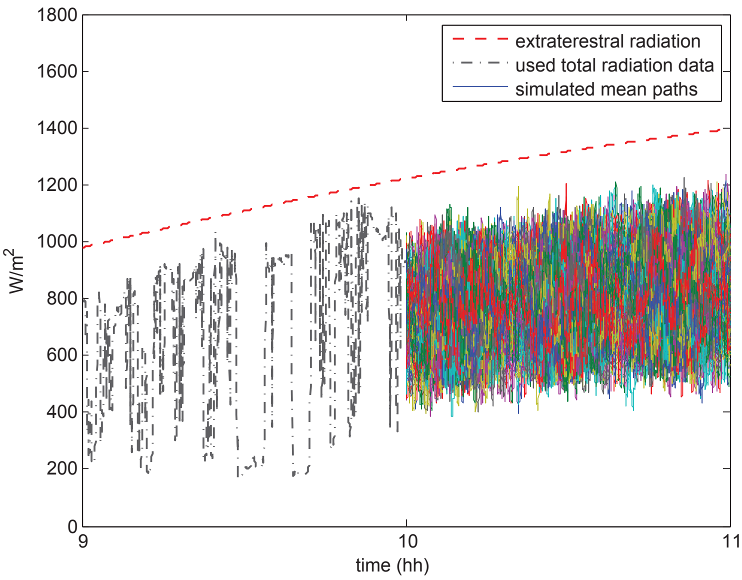


(b)

Figure 5.14: (a) Simulated paths in 10h-11h generated by CTM-y with parameters estimated from observations in 07h-10h, day 118 (type II, partially cloudy day), 2006, Guadeloupe. (b) Observed data: blue solid line.

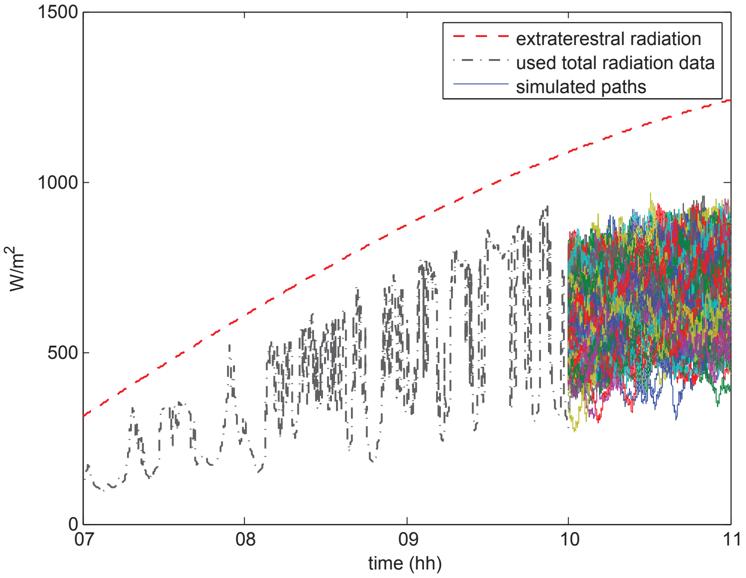


(a)

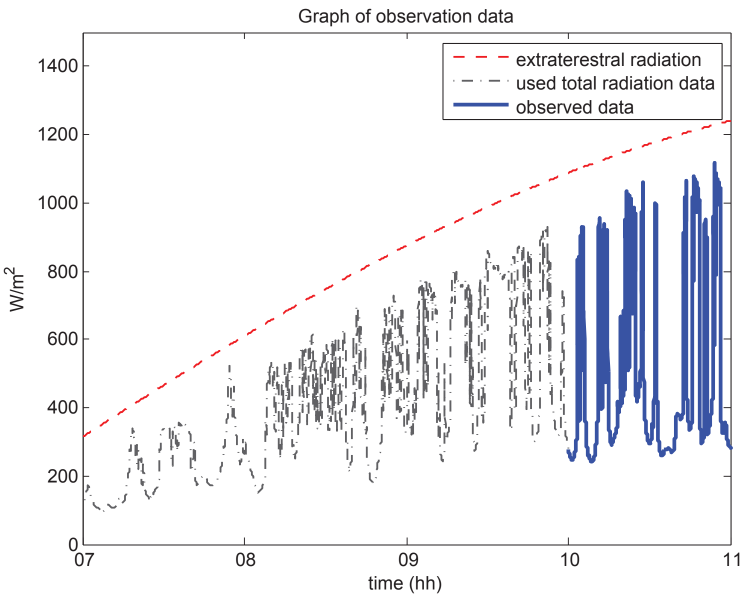


(b)

Figure 5.15: (a) left: Simulated paths in 10h-11h generated by CTM-k with parameters estimated from observations in 09h-10h, day 118 (type II, partially cloudy day), 2006, Guadeloupe (k-DATA-II.1), right: Observed data. (b) Simulated paths in 10h-11h generated by CTM-y with parameters estimated from y-DATA-II.1.

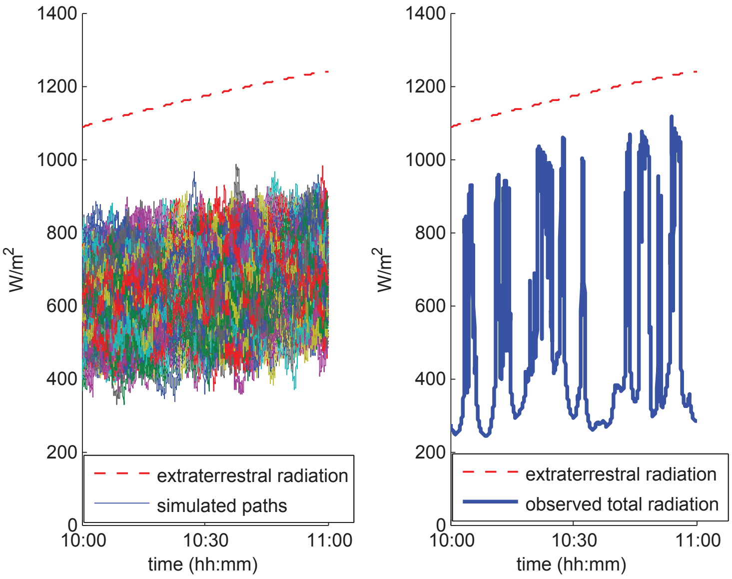


(a)

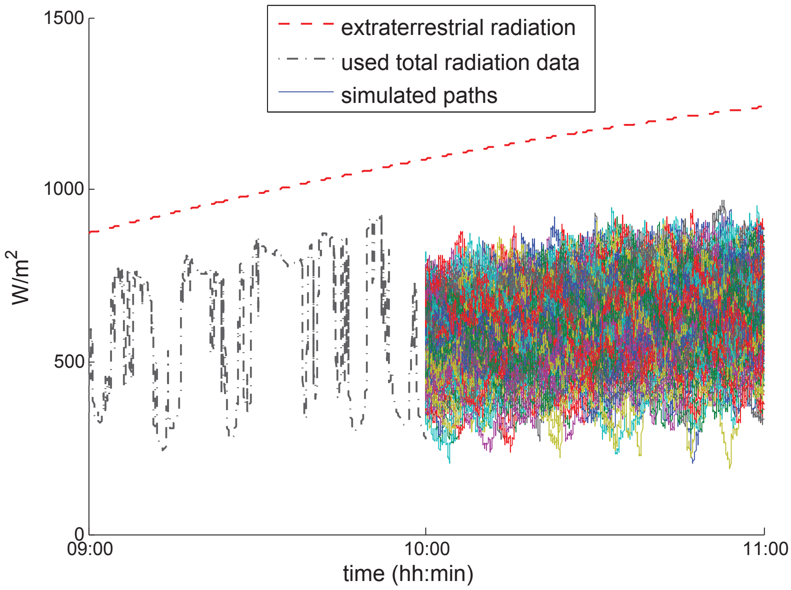


(b)

Figure 5.16: (a) Simulated paths in 10h-11h generated by CTM-y with parameters estimated from observations in 07h-10h, day 234 (type II, partially cloudy day), 2006, Guadeloupe. (b) Observed data observed: the blue solid line.



(a)



(b)

Figure 5.17: (a) left: Simulated paths in 10h-11h generated by CTM-k with parameters estimated from observations in 09h-10h, day 234-th (type II, partially cloudy day), 2006, Guadeloupe (k-DATA-II.3), right: Observed data. (b) Simulated paths in 10h-11h generated by CTM-y with parameters estimated from y-DATA-II.3.

Chapter 6

Conclusion

In order to understand and model the behaviour of solar radiation and clearness index, besides the two other approaches commonly used ("physical modelling" and "statistical solar climatology", as mentioned in the introduction), we have proposed a new approach, a HMM-type stochastic model for taking in account the influence of meteorological regimes.

The parameters of the model are estimated from real data using filtering equations, Girsanov change of measure theorem in stochastic calculus and the celebrated EM algorithm. The method has been tested and illustrated on several different types of observational data. The simulated data has been used to estimate the distribution of the daily clearness index and to predict the total solar radiation in a short horizon.

As a conclusion, we emphasize some elements that compare our model to physical models and other statistic models. We then take some notes in the problem of model construction, the problem of parameter estimation as well as the problem of simulated data application. From this, we determine the work which can be continued to study in the future.

Comparison with physical models

Our models are mainly data-driven as parameters of models are estimated from observed data and the number of regimes is chosen by observing the distribution of those data but they also include a physical model part through the extraterrestrial radiation.

In **the physical models**, direct beam (I_b) and diffuse (I_d) radiation components are obtained as a function of the specific atmospheric transmittances. They **require several physical parameters as inputs**: water vapor absorption (T_w), Rayleigh scattering (T_r), uniformly mixed gases absorption (T_g), ozone absorption (T_o), aerosol total extinction (T_a), ...

For instance, Psiloglou *et al.* [Psiloglou 2000] proposed a clear-sky radiation model with the following components:

- total radiation : $G_t = I_b + I_d$,
- direct-beam radiation : $I_b = I_0 \cos\theta_z T_w T_r T_g T_a$,
- diffuse radiation : $I_d = I_0 \cos\theta_z T_w T_g T_o T_{aa} \left(1 - \frac{T_a}{T_{aa}} T_r\right) /2 + I_{dm}$,

where T_{aa} is the absorption aerosol broadband transmittance function, T_{dm} is a multiple-scattering component, θ_z is the zenith angle and I_0 is the extraterrestrial normal solar radiation in the n_l-th day of the year:

$$I_0 = I_{SC}(1.00011 + .034221\cos\Gamma + .00128\sin\Gamma + .000719\cos 2\Gamma + .000077\sin 2\Gamma),$$

here $I_{SC} = 1373W/m^2$ is the solar constant and Γ (in radians) is the day angle which is represented by:

$$\Gamma = \frac{2\pi n_l - 1}{365}, \quad n_l = 1, 2, \ldots, 365. \tag{6.1}$$

However, our models also have a physical factor, namely the local extraterrestrial radiation I_t. The observation equation (4.4) of CTM-y depends on I_t. For DTM-K and CTM-k, the observations used in the parameter estimation are CISs computed using I_t and measurements of total solar radiation G_t:

$$(1.10): \; k_t = \frac{G_t}{I_t},$$
$$\text{or } (1.11): \; K^{\Delta t} = \frac{\int_{\Delta t} G_s ds}{\int_{\Delta t} I_s ds}.$$

As a consequen our models generate data having meteorological characteristics of areas where data used for estimation were collected.

Comparison with other statistical models

Statistical models and methods have played a wide range of applications in solar radiation and the climatology research.

As far as we know, most of statistical models used for solar radiation and clearness index hinge on regression models using the correlative relation between statistical variables (including main solar radiation components, clearness index and meteorological parameters such as sunshine duration, cloudiness, temperature, etc). The parameters of the statistical models are estimated from complete observation data of all these statistical variables.

For instance, Angstrom (1924) suggested the following linear expression for the relationship between daily clearness index K_h and sunshine duration ratio S_h [Tovar-Pescador 2008]:

$$K_h = a + bS_h,$$

where a, b are model parameters which are estimated by regression technique from observation data of statistical variables K_h and S_h (**complete data of model**). This linear expression was used in practical applications for many years to estimate the daily, monthly and annual total solar radiation from the comparatively simple measurements of sunshine duration. [Ogelmen 1984, Akinoglu 1990] proposed a quadratic expression by adding a non-linear term.

Our method is also based on several statistic techniques (ML parameter estimation, EM algorithm, filtered estimate, ...). However, in comparison with usual statistical models, our proposed models have three original features:

- Our models take in account the hidden (unobserved) dynamic of environmental regimes as well as observed variables (extraterrestrial, direct and diffuse radiations)
- Our model parameters can be estimated from incomplete data. This is important for solar radiation research as well as for climatological research because all climatological quantities and meteorological components are not always available.
- Our models having a random component, simulated data can be used to study the probability distribution of clearness index and total radiation.

Future works

Problems in the model construction

Our models depend on measurements of the total solar radiation and the extraterrestrial radiation I_t with astronomical calculations depending on the geographical location. A better modelling should consider more information on observed data and meteorological parameters, dealing with multidimensional vector as it is done in many physical models.

Problems in parameter estimation

The technique we used in the problem of parameter estimation is the filtering estimates, also called *forward estimates*, which are based on the observation history up to time h in discrete time, $\mathcal{Y}_h^K \triangleq \sigma\{K_1, K_2, \ldots, K_h\}$, $h = 1, 2, \ldots$ (or up to time t in continuous time, $\mathcal{Y}_t^Y \triangleq \sigma\{Y_s : 0 \leqslant s \leqslant t\}$, $t \in [0, T]$, respectively).

To achieve better estimates, we think of *backward estimates*, an estimating technique based on observations in the future, $\mathcal{Y}_{h:M}^K \triangleq \sigma\{K_h, K_{h+1}, \ldots, K_M\}$ (or $\mathcal{Y}_{t:T}^Y \triangleq \sigma\{Y_s : t \leqslant s \leqslant T\}$, resp.). Backward estimates are calculated as a backward recursion from the end of the batch of observations. The *forward-backward* estimate is termed *smoothing estimation*, based on the past, present and future of observation data, $\mathcal{Y}_h^K \vee \mathcal{Y}_{h:M}^K$ (or $\mathcal{Y}_t^Y \vee \mathcal{Y}_{t:T}^Y$, resp.) [Elliott 2010, James 1996].

Problems in prediction

First, our prediction results should be compared to recent prediction results obtained by EDF (French electricity company) using some data analysis techniques.
Next, our 1h prediction horizon was a short-term one and we need to refine our models to deal with very short-term (10mn) prediction and also long-term prediction (next day, next month, next year).
Moreover, as seen in the examples of chapter 5, prediction is connected to classification problems (classification of sequences, of days, of several months periods) and this classification aspect is not considered in our present models. Conversely it can be thought that estimated parameters can be used for classification purposes.

Bibliography

[Akinoglu 1990] B. G. Akinoglu and A. Ecevit. *Construction of a Quadratic Model Using Modified Angstrom Coefficients To Estimate Global Solar Radiation.* Solar Energy, vol. 45, no. 2, pages 85–92, 1990. (Cited on page 124)

[Benaglia 2009] T. Benaglia, D. Chauveau, D. R. Hunter and D. S. Young. *mixtools: An R Package for Analyzing Finite Mixture Models.* HAL : hal-00384896, 2009. (Cited on page 99)

[Campillo 1989] F. Campillo and F. Le Gland. MLE *for the patially observed diffusions: Direct maximization vs. the* EM *algorithm.* Stochastic Processes and their Applications, vol. 33, no. 2, pages 245–274, 1989. (Cited on page iii)

[Celeux 1989] G. Celeux and J. Diebolt. *Une version de type recuit simule de l'algorithme* EM. INRIA, Rappports de Recherche, N1123, 1989. (Cited on page iii)

[Charalambous 2000] C. D. Charalambous. *Maximum likelihood parameter estimation from incomplete data via the sensitivity equations: the continuous-time case.* Automatic Control, IEEE Transactions on, vol. 45, no. 5, pages 928–934, 2000. (Cited on page 26)

[Clark 1977] J. M. C. Clark. *The design of robust approximations to the stochastic differential equations of nonlinear filtering.* Communication Systems and Random Processes Theore, Darlington, pages 721–734, 1977. (Cited on pages iv, 48, 50 and 81)

[Clark 2005] J. M. C. Clark and D. Crisan. *On a robust version of the integral representation formula of nonlinear filtering.* Probab. Theory Relat. Fields, vol. 133, pages 43–56, 2005. (Cited on page iv)

[Delsaut 2013] M. Delsaut, P. Jeanty, L. Trovalet, H. Ralambondrainy, J.D. Lan-Sun-Luk, M. Bessafi, P. Charton and J.P. Chabriat. *Clustering daily solar radiation from Reunion Island using data analysis methods.* International Conference on Renewable Energies and Power Quality ICREPQ'13 - Bilbao, Spain, 2013. (Cited on page 53)

[Delta-T-Devices 2012] Delta-T-Devices. *Spn1 - sunshine pyranometer (May 2012).* URL http:www.delta-t.co.uk, 2012. (Cited on page 10)

[Dembo 1986] A. Dembo and O. Zeitouni. *Parameter estimation of partially observed continuous time stochastic processes via the EM algorithm.* Stochastic Processes and their Applications, vol. 23, pages 91–113, 1986. (Cited on pages iii, 24, 27 and 28)

[Dempster 1977] 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), vol. 39, no. 1, pages 1–38, 1977. (Cited on pages iii and 28)

[Duffie 1980] J. A. Duffie and W. A. Beckman. Solar engineering of thermal process. Wiley Interscience, New York, 1980. (Cited on page 3)

[Duffie 1991] J. A. Duffie and W. A. Beckman. Solar engineering of thermal process. Wiley Interscience, New York, 1991. (Cited on page 3)

[Duffie 2006] J. A. Duffie and W. A. Beckman. Solar engineering of thermal process. New York, 2006. (Cited on page 3)

[Elliott 1982] J. R. Elliott. Stochastis calculus and applications. Springer-Verlag, New York, 1982. (Cited on pages 13 and 22)

[Elliott 1995] J. R. Elliott, L. Aggoun and J. B. Moore. Hidden markov models : Estimation and control. Springer-Verlag New York, 1995. (Cited on pages iii, iv, 17, 24, 33, 40, 41 and 74)

[Elliott 2010] J. R. Elliott, L. Aggoun and J. B. Moore. Hidden markov models : Estimation and control. Springer, 2010. (Cited on pages iii, v, 12, 13, 17, 21, 24, 26, 44 and 125)

[Gross 2012] G. Gross. Markov models for reliability evaluation. University of Illinois at Urbana-Champaign, 2012. (Cited on page 23)

[Gueymard 1993] C. Gueymard. *Critical analysis and performance assessment of clear sky solar irrandiance models using theoretical and measured data*. Sol. Energy, vol. 51, pages 121–138, 1993. (Cited on page ii)

[Higham 2001] D. J. Higham. *An algorithmic introduction to Numerical Simulation of Stochatic Differential Equation*. Society for Introdustrial and Applied Mathematics (SIAM), no. 43, pages 525–546, 2001. (Cited on pages 49 and 82)

[Ianetz 2008] A. Ianetz and A. Kudish. Modeling solar radiation at the earth's surface, Chp.4: A Method for Determining the Solar Global and Defining the Diffuse and Beam Irradiation on a Clear Day, pages 92–113. Springer-Verlag Berlin Heidelberg, 2008. (Cited on page 53)

[Iqbal 1986] M. Iqbal. An introduction to solar radiation. Academic, Toronto, 1986. (Cited on page 5)

[James 1996] M. R. James, V. Krishnamurthy and F. Le Gland. *Time Discretization of Continuous-Time Filters and Smoothers for* HMM *Parameter Estimation*. IEEE TRANSACTIONS ON INFORMATION THEORY, vol. 42, no. 2, pages 593–604, 1996. (Cited on pages iv, 24, 26, 33, 41, 50, 74, 75, 81, 83 and 125)

[Kambezidis 1989] H. D. Kambezidis and N. S. Papanikolaou. *Total solar irradiance flux through inclined surfaces with arbitrary orientation in Greece: comparison between measurements and models.* In Proceedings of XIV Assembly of EGS, pages 13–17, 1989. (Cited on page ii)

[Krishnamurthy 2002] V. Krishnamurthy and R. Elliott. *Robust Continuous-Time Smoothers Without Two-Sided Stochastic Integrals.* IEEE TRANSACTIONS ON AUTOMATIC CONTROL, vol. 47, no. 11, pages 1824–1841, 2002. (Cited on pages iv, 22 and 81)

[Lanini 2010] F. Lanini. Division of global radiation into direct radiation and diffuse radiation. Master's thesis, University of Bern, 2010. (Cited on page 4)

[Lipster 2010] R. S. Lipster and A. N. Shiryayev. Statistics of random processes, parts i general theory. Springer-Verlag, Berlin-Heidelberg, 2010. (Cited on pages 21, 27 and 41)

[Liu 1960] B. Y. Liu and R. C. Jordan. *The interrelationship and characteristic distribution of direct, diffuse and total solar radiation.* Solar Energy, no. 4, pages 1–19, 1960. (Cited on page 2)

[Muneer 1997] T. Muneer, M. S. Gul and H. D. Kambezidis. *Solar radiation models based on meteorological data.* In Proc. ISES World Congress, Taegon, Korea, 1997. (Cited on page ii)

[Ogelmen 1984] H. Ogelmen, A. Ecevit and E. Tasmiroglu. *A new method for estimating solar radiation from bright sunshine data.* Solar Energy, vol. 33, no. 6, pages 619–625, 1984. (Cited on page 124)

[Psiloglou 2000] B. E. Psiloglou, M. Santamouris and D. N. Asimakopoulos. *Atmospheric broadband model for computation of solar radiation at the EarthâĂŹs surface. Application to Mediterranean climate.* Pure Appl. Geophys, vol. 157, pages 829–860, 2000. (Cited on pages ii, 2, 7 and 123)

[Psiloglou 2007] B. E. Psiloglou and D. Kambezidis. *Performance of the meteorological radiation model during the solar eclipse of 29 March 2006.* Atmospheric Chemistry and Physics, no. 7, pages 6047–6059, 2007. (Cited on pages ii and 7)

[Sen 2008] Z. Sen. Solar energy fundamentals and modeling techniques. Springer, 2008. (Cited on pages 2 and 6)

[Soubdhan 2009] T. Soubdhan, R. Emilion and R. Calif. *Classification of daily solar radiation distributions using a mixture of Dirichlet distributions.* Solar Energy, vol. 83, pages 156–163, 2009. (Cited on pages ii, iii and 53)

[Soubdhan 2011] T. Soubdhan and R. Emilion. *Prevision of typical solar days for PV production using Hidden Markov model.* Solar world congress SWC2011 - Kassel, Germany, 2011. (Cited on page iii)

[Spencer 1972] J. W. Spencer. *Fourier series reresentation of the position of the sun.* Search 2:172, 1972. (Cited on pages 3 and 5)

[Tingilinde 2006] Tingilinde. *Latest sunset, earliest sunrise.* http://tingilinde.typepad.com/starstuff/2006/06/latest_sunset_e.html, 2006. (Cited on page 4)

[Tovar-Pescador 2008] J. Tovar-Pescador and Z. Sein. Modeling solar radiation at the earth's surface, Chp.3: Modelling the Statistical Properties of Solar Radiation and Proposal of a Technique Based on Boltzmann Statistics, pages 55–91. Springer-Verlag Berlin Heidelberg, 2008. (Cited on pages 2 and 124)

[Tran a] V. L. Tran, R. Abraham and R. Emilion. *Modelling Clearness Index Sequences in Random Medium and Applications of Simulated Data. Preprint.* (Cited on page 30)

[Tran b] V. L. Tran, R. Abraham and R. Emilion. *A Stochastic Differential Equation in Random Medium for Modelling Total Solar Radiation Sequences. Preprint.* (Cited on page 70)

[Tran 2011] V. L. Tran, R. Abraham and R. Emilion. *Hidden Markov model and stochastic differential equation of solar radiation sequences.* Gravitation, Astrophysics and Cosmology. Proceeding of the Xth International Conference on Gravitation, Astrophysics and Cosmology, pages 66–67, 2011. (Cited on page 70)

[Tran 2013] V. L. Tran. *Stochastic Modelling for instantaneous clearness index sequences: Estimation and Prediction.* Proceeding of International Conference on Statistics and its Interractions with Other Disciplines, page 18, 2013. (Cited on page 30)

[Wong 1985] E. Wong and B. Hajek. Stochastic processes in engineering systems. Springer-Verlag, New York, 1985. (Cited on page 12)

Van Ly TRAN

Modèles Stochastiques des Processus de Rayonnement Solaire

Résumé :

Les caractéristiques des rayonnements solaire dépendent fortement de certains événements météorologiques **non observés** comme fréquence, taille et type des nuages et leurs propriétés optiques (aérosols atmosphériques, albédo du sol, vapeur d'eau, poussière et turbidité atmosphérique) tandis qu'une séquence du rayonnement solaire peut être **observée** et mesurée à une station donné. Ceci nous a suggéré de modéliser les processus de rayonnement solaire (ou d'indice de clarté) en utilisant un modèle Markovien caché (HMM), paire corrélée de processus stochastiques.

Notre modèle principal est un HMM à temps continu $(X_t, y_t)_{t\geq 0}$ est tel que (y_t), le processus observé de rayonnement, soit une solution de l'équation différentielle stochastique (EDS) : $dy_t = [g(X_t)I_t - y_t]dt + \sigma(X_t)y_t dW_t$, où I_t est le rayonnement extraterrestre à l'instant t, (W_t) est un mouvement Brownien standard et $g(X_t)$, $\sigma(X_t)$ sont des fonctions de la chaîne de Markov non observée (X_t) modélisant la dynamique des régimes environnementaux.

Pour ajuster nos modèles aux données réelles observées, les procédures d'estimation utilisent l'algorithme EM et la méthode du changement de mesures par le théorème de Girsanov. Des équations de filtrage sont établies et les équations à temps continu sont approchées par des versions *robustes*.

Les modèles ajustés sont appliqués à des fins de comparaison et classification de distributions et de prédiction.

Mots de clé : rayonnement solaire, indice de clarté, HMM, EDS, algorithme EM, Thérème de Girsanov, filtrage.

Stochastic Models of Solar Radiation Processes

Abstract :

Characteristics of solar radiation highly depend on some **unobserved** meteorological events such as frequency, height and type of the clouds and their optical properties (atmospheric aerosols, ground albedo, water vapor, dust and atmospheric turbidity) while a sequence of solar radiation can be **observed** and measured at a given station. This has suggested us to model solar radiation (or clearness index) processes using a hidden Markov model (HMM), a pair of correlated stochastic processes.

Our main model is a continuous-time HMM $(X_t, y_t)_{t\geq 0}$ is such that the solar radiation process $(y_t)_{t\geq 0}$ is a solution of the stochastic differential equation (SDE) : $dy_t = [g(X_t)I_t - y_t]dt + \sigma(X_t)y_t dW_t$, where I_t is the extraterrestrial radiation received at time t, (W_t) is a standard Brownian motion and $g(X_t)$, $\sigma(X_t)$ are functions of the unobserved Markov chain (X_t) modelling environmental regimes.

To fit our models to observed real data, the estimation procedures combine the Expectation Maximization (EM) algorithm and the measure change method due to Girsanov theorem. Filtering equations are derived and continuous-time equations are approximated by *robust* versions.

The models are applied to pdf comparison and classification and prediction purposes.

Keywords : solar radiation, clearness index, HMM, SDE, EM algorithm, Girsanov theorem, filtration.

UNIVERSITE D'ORLEANS

MAPMO UMR 7349, Fédération Denis Poisson
Université d'Orléans, UFR Science
Bâtiment de mathématiques - Rue de Chartres
B.P. 6759 - 45067 Orléans cedex 2
FRANCE

MAPMO Orléans