IMPACTS OF SPATIOTEMPORAL DEPENDENCY AND ASYMMETRIC INFORMATION ON THE ANALYSIS OF OPTIMAL CROP YIELD INSURANCE

OPTİMAL BİTKİSEL ÜRÜN VERİM SİGORTASI ANALİZİNDE MEKANSAL-ZAMANSAL BAĞIMLILIK VE ASİMETRİK BİLGİ ETKİLERİ

GÜVEN ŞİMŞEK

PROF. DR. ŞAHAP KASIRGA YILDIRAK
Supervisor

Submitted to Graduate School of Science and Engineering of Hacettepe University
as a Partial Fulfillment to the Requirements
for the Award of the Degree of Doctor of Philosophy
in Actuarial Sciences

2020

ABSTRACT

IMPACTS OF SPATIOTEMPORAL DEPENDENCY AND ASYMMETRIC INFORMATION ON THE ANALYSIS OF OPTIMAL CROP YIELD INSURANCE

Güven ŞİMŞEK

Doctor of Philosophy, Department of Actuarial Sciences

Supervisor: Prof. Dr. Şahap Kasırga YILDIRAK

July 2020, 113 pages

Having the most indispensable role in agricultural production, farmers need to protect themselves against the risks arising from agricultural production in order to continue producing. It is very essential to provide an insurance policy aimed at meeting the farmer's coverage need. Yield insurance is our specific interest since one of the most vulnerable segments of the society is farmers due to the fact that the farmers' wealth is highly exposed to environmental risks.

In order to discuss the equilibrium in an insurance market, we study the insurance demand in the frame of principal-agent problem which examines the optimal contract between the *principal* and the *agent*. The principal-agent problem that is widely investigated in economics and finance could be also used for insurance market to analyze the relationship between the insurer (*principal*) and the insured (*agent*). In this context, we mathematically deal with adverse selection which is one of the most essential concepts in insurance. Since individuals who need insurance have different expectations, preferences and justifications, it is very complicated to produce an insurance product which would satisfy all needs and cover several people. Therefore, it is very significant to analyze the demand for insurance accurately. In this thesis, we model the insurance demand in order to examine the equilibrium in insurance market. Having discussed the equilibrium for an insurance product generally, we study the asymmetric information concept in yield insurance specifically.

Yield insurance covers the yield loss occuring when the yield does not exceed a specified yield level. This insurance product has a special place among agricultural insurance types because the aim of keeping the agricultural production at a determined level provides sustainability in the ecosystem. In this thesis, we examine the effects of moral hazard in yield insurance in order to optimize the farmer's effort for both loss-prevention and loss-reduction. The impacts of moral hazard is also analyzed graphically and numerically by using certainty equivalent approach. Since the agricultural insurance is a government-backed system in Turkey, reducing risks related to asymmetric information leads the farmer to make an effort to search for more efficient ways of agricultural production.

Climate change, which is the biggest challenge for agricultural insurance, is also investigated through spatiotemporal modelling of the crop yield in the thesis. We consider changes in extreme weather events such as flood, drought and hail by taking into account of location and time related effects on agricultural production. In this thesis, we use the hierarchical Bayesian approach for the conditional distribution of the yield to reflect the spatiotemporal dependency.

In order to examine the district-based dependency, we consider spatial, temporal and spatiotemporal effects on the estimation of the crop yield by using hiearchical Bayesian structure. After we select the best model by using performance criteria, we compute the premium rates related to various coverage levels.

Keywords: Asymmetric information, certainty equivalent, loss-prevention, loss-reduction, moral

hazard, optimal effort, principal-agent, spatiotemporal dependency.

ÖZET

OPTİMAL BİTKİSEL ÜRÜN VERİM SİGORTASI ANALİZİNDE MEKANSAL-ZAMANSAL BAĞIMLILIK VE ASİMETRİK BİLGİ ETKİLERİ

Güven ŞİMŞEK

Doktora, Aktüerya Bilimleri Bölümü

Tez Danışmanı: Prof. Dr. Şahap Kasırga YILDIRAK

Temmuz 2020, 113 sayfa

Tarımsal üretimin omurgası olan çiftçinin üretimine devam etmesi için üretimden kaynaklanan risklere karşı kendini koruma altına alması gerekmektedir. Çiftçinin ortaya çıkan teminat ihtiyacına yönelik sigorta poliçelerinin sunulması önem kazanmaktadır. Çiftçinin varlığının büyük oranda çevresel risklere maruz olması, çiftçileri toplumun en savunmasız kesimlerinden biri haline getirmektedir. Bu nedenle de bu çalışmada verim sigortası ile ilgilenilmiştir.

Daha çok iktisat ve finans alanlarında çalışılan asil-vekil probleminde; asil ve vekil olarak ifade edilen iki taraf arasındaki optimal sözleşme incelenmektedir. Sigortacılık için de önemli olan ve asimetrik bilgiden kaynaklanan tersine seçim kavramı matematiksel olarak ele alınmaktadır. Sigorta şirketi (asil) ve sigortalı (vekil) arasındaki ilişki de, asil-vekil problemi bağlamında incelenebilmektedir. Sigortaya ihtiyaç duyan kişilerin beklentileri, tüketici tercihleri ve sigorta yaptırma gerekçeleri farklı olduğundan, birçok kişiyi kapsayacak ve memnun edecek bir

poliçe üretmek karmaşık bir hale gelmektedir. Bu sebeple, sigorta için talebin iyi bir şekilde analiz edilmesi önem kazanmaktadır. Bu tezde, tam bilgi ve tersine seçim durumları için asil-vekil problemi çerçevesinde, sigorta piyasasında talebin modellenmesi ve sigorta piyasası için denge durumunun incelenmesi amaçlanmaktadır. Genel bir sigorta ürünü için denge durumu tartışıldıktan sonra, spesifik olarak verim sigortasında asimetrik bilgi durumu ele alınmıştır.

Verim sigortası, verimin belirlenen bir seviyeyi aşmaması durumunda ortaya çıkan verim kaybını teminat altına almaktadır. Tarımsal üretimi belirlenmiş bir seviyede tutmayı amaçlayan verim sigortası ekosistemde sürdürülebilirliği sağladığı için bu sigorta ürününün tüm tarımsal sigortalar içerisinde özel bir yeri vardır. Bu çalışmada çiftçinin verim kaybını hem engellemek, hem de azaltmak için gösterdiği çabayı optimize etmek amacıyla asimetrik bilgi türlerinden biri olan ahlaki tehlikenin verim sigortasındaki etkisi incelenmektedir. Ahlaki tehlikenin etkileri kesinlik eşdeğeri yaklaşımı altında numerik ve grafiksel olarak analiz edilmiştir. Türkiye'de tarım sigortası devlet destekli olduğu için; asimetrik bilgiye ilişkin risklerin azaltılması, çiftçiyi tarımsal üretimin daha etkin yollarını araştırma çabasına teşvik etmektedir.

Tarım sigortası poliçeleri için uygun prim oranların belirlenmesi sistemin sürdürülebilmesi için hayati önem taşımaktadır. Örneğin, yüksek prim oranları sigortaya olan talebi düşüreceğinden, sigortayı sadece yüksek risk profiline sahip üreticiler tercih edecektir. Bu durumda, bu kişilerin hasar getirmesi olasılığı yüksek olduğundan ödenecek teminat miktarları da artacaktır ve sonuç olarak piyasa başarısızlığa uğrayacaktır.

Bu çalışmada aynı zamanda bitkisel ürün veriminin mekansal-zamansal modellemesi yoluyla tarımsal sigortalar için en büyük zorluk olan iklim değişikliği de incelenmiştir. Sel, kuraklık, dolu gibi şiddetli hava olaylarındaki değişiklikler, tarımsal üretim üzerindeki konuma ve zamana bağlı etkiler hesaba katılarak ele alınmıştır. Bu çalışmada, mekansal-zamansal bağımlılığı yansıtmak amacıyla verimin koşullu dağılımı için hiyerarşik Bayesyen yaklaşım kullanılmıştır.

Çalışmada buğday üretim hacmi bakımından üst sıralarda yer alan Ankara ve Konya illerine ait toplam 47 ilçe için 1991-2016 yılları arası buğday verim verileri kullanılmıştır. Buğday verimini tahmin etmek için ilk olarak açıklayıcı değişkenler olmaksızın sadece konum, zaman ve mekansal-zamansal etkilerin yer aldığı rastgele etkiler incelenmiştir. Model seçimi ve perfor-

mans kriterlerine göre tercih edilen model kullanılarak; bölge, il ve seçilen ilçeler için prim oranları elde edilmiş ve sonuçlar sunulmuştur.

Anahtar Kelimeler: Asimetrik bilgi, kesinlik eşdeğeri, hasar-engelleme, hasar-azaltma, ahlaki tehlike, optimal çaba, asil-vekil, mekansal-zamansal bağımlılık.

ACKNOWLEDGEMENT

First of all, I am grateful to my supervisor Prof. Dr. Şahap Kasırga Yıldırak for the continuous support in every part of my Ph.D study and research. I would like to express my sincere gratitude to my advisor for his patience, motivation, enthusiasm, and immense knowledge.

I would like to thank to my committee chair Prof. Dr. Fatih Tank, and committee members Assoc. Prof. Dr. Şule Şahin, Assoc. Prof. Dr. Özge Sezgin Alp and Assist. Prof. Dr. Uğur Karabey who have contributed to the development of the study with helpful suggestions. I am grateful to Assoc. Prof. Dr. Şule Şahin for her insightful comments, suggestions, questions and encouraging my research.

My sincere thanks also go to Dr. Ezgi Nevruz for encouraging my research, the stimulating discussions and supporting me for everything. I would also like to thank Assist. Prof. Dr. Başak Bulut Karageyik and Res. Assist. Hasan Güzel for their friendship. Thanks also deserve to go to my girlfriend Dr. Burcu Şabanoğlu for all her love and support.

Finally I would like to thank to my family for supporting me for everything and standing by me in all difficulties. Words cannot express how grateful I am.

CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

1. INTRODUCTION

Future threatening improvements such as growing population and decreasing agricultural areas have become very important for the sustainability of the production of the basic nutritional sources in a safe way. Having the most indispensable role in agricultural production, farmers need to protect themselves against the risks arising from agricultural production in order to continue producing. It is very essential to provide an insurance policy aimed at meeting this coverage need. Over the past few decades, crop insurance has played a growingly pervasive role in the agriculture market. Crop insurance helps farmers to make their incomes more stable gradually and also provides a considerable protection to farmer against the catastrophic risks such as drought, flood and hail. Crop failures could cause higher and more devastating negative impacts on agronomics.

Yield insurance is our specific interest since one of the most vulnerable segments of the society are farmers due to the fact that the farmers' wealth is highly exposed to environmental risks. Yield insurance covers the crop yield damage occuring when crop yield is below the predetermined threshold value. This insurance product has a special place among agricultural insurance types because the aim of keeping the agricultural production at a determined level provides sustainability in the ecosystem. In our study, we investigate the effects of "moral hazard" in crop yield insurance in order to optimize the farmer's effort for both "loss-prevention" and "loss-reduction".

In our thesis, we evaluate the factors having negative influences on yield insurance, one of which is asymmetric information. It is commonly accepted that "asymmetric information" is one of the most serious problems causing failure in the insurance market. "Asymmetric information" is known as "principal-agent problem" in contract theory. The Principal-Agent Problem is used to model the relationship between an agent and a principal. In this problem, one party, so-called "agent", agrees to work on behalf of another party, called "principle", in return for some incentives. Therefore, the ultimate aim of the principal-agent problem is to design an optimal contract which maximizes the principal's and the agent's utility.

Asymmetric information arises when the agent (insured) has better information about the decisions that he/she makes on behalf of the principal (insurer). In general, two types of asymmetric information is considered: "moral hazard (hidden action)" and "adverse selection (hidden type)". Moral hazard occurs when the insured's behavior cannot be fully monitored by the insurer. As an example of moral hazard, we consider a crop insurance contract. The insurer covers the insured against risk and the farmers pay the premiums for this coverage. The insured might take an action that is unobservable for the insurer. Due to the fact that the farmer expects to be paid by the insurer when the risk occurs, the farmer might decide to be lazy and work less to cultivate his/her fields. Therefore, the insured's action causes an unfavourable effect on the insurer's payoff. On the other hand, "adverse selection" turns out when the insured has private information which is not known to the insurer. For instance, the farmer might choose to buy a crop insurance contract on his/her cultivated agricultural areas which are distinctly possible to be faced with natural hazards such as flood, landslides and drought.

A contract design mechanism under the problem of moral hazard is shown in Figure 1.1.

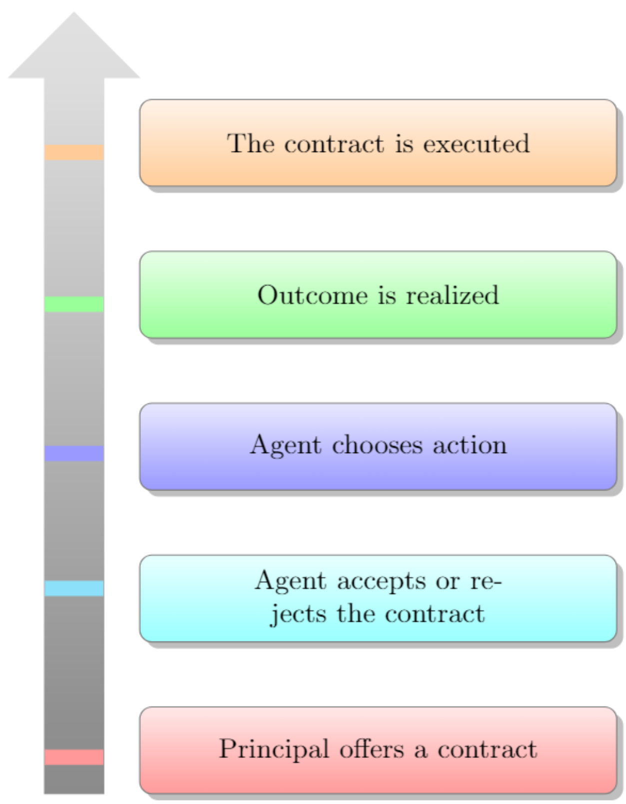

Figure 1.1: A contract design mechanism

Here, if the agent rejects the contract, then both the principal and the agent get their reservation utilities (i.e. the utility is what the agent gets if the agent resigns the contract). If the contract is accepted, it proceeds as how it is agreed on.

1.1. Problem Definition and Objectives of the Study

It is very essential for the farmer to buy an insurance contract in order to prevent loss income in connection with the agricultural activities and to keep production in an acceptable level. Agricultural insurance protect the farmer against the financial influences of the losses arising from the agricultural production. In this sense, traditional agricultural insurance products cover natural disaster risks and the risks which have negative effects on the farmer's income. However, the additional risks related to adverse selection and moral hazard causes higher indemnity payments. Asymmetric information leads insureds to earn unjustified benefit due to having more information about their own production than the insurer has. The insurance company thus increase premium amounts in order to minimize the negative impacts of "asymmetric information". Although the government pays 60 percent of the premiums, these higher premium rates lower the demand of agricultural insurance products. Because of this reason, classical agricultural insurance system could stay afloat by the help of the government support to the premium payment. Therefore, the solvency of the global agricultural insurance system could stay stable by reducing the behavioral risks of the insured side, i.e. asymmetric information.

In this study, we consider impacts of asymmetric information and spatiotemporal dependency on solvency of yield insurance. For this reason, we handle the farmer's optimal effort as an evidence of non-existence of the moral hazard and asymmetric information. On the other hand, we discuss taking into account of the spatiotemporal dependency of neighbourhoods in this study. By the help of the hierarchical Bayesian models applied to the district-based crop yield data; we investigate spatial, temporal and spatiotemporal effects on the crop yields.

First we propose an optimization approach used to maximize the expected utility of a farmer for a crop yield insurance. Here, it is considered that the farmer is able to control both the occurance and amount of the yield loss. We handle these two situations separately with the account of effort's being observable and non-observable. In this part, we examine the optimization problem from a farmer-based perspective.

The common tendency of the existing studies for modelling the conditional yield distribution is to use "Markov Chain Monte Carlo (MCMC)" methods for Bayesian estimation. In our study, we use "Integrated Nested Laplace Approximation (INLA)" to estimate parameters of the hierarchical Bayesian model. In comparison with the traditional Bayesian model, the hierarchical approach handles spatiotemporal dependencies among crop yields. The case study for 47 districts of two cities, where the wheat production is the highest, illustrates the results of hierarchical Bayesian modelling of the district-based crop yield sample data.

The main aim of this study is to analyze the crop yield insurance from the point of asymmetric information and spatiotemporal dependency. This thesis consists of three main parts. First of all, we consider a general insurance market in order to examine the demand for insurance in equilibrium situation. Having discussed the insurance demand for "low-risk" and "high-risk" insureds, the insured's effort turns out to be a critical issue for risk assessment. Optimal effort is examined for crop yield insurance in the second part. Asymmetric information is considered under loss-prevention and loss-reduction models. As seen in the summary chart in Figure 1.2, optimal effort is analyzed according to the cases of observable effort and non-observable effort for both loss-prevention and loss-reduction models. Here, we investigate all cases under expected utility theory (EUT) whereas we use certainty equivalent (CE) approach for the cases which consider that the effort is not observable. In addition to risk assessment, we also calculate premium rates by modeling the spatiotemporal dependency among subregions which are neighbours in the last part.

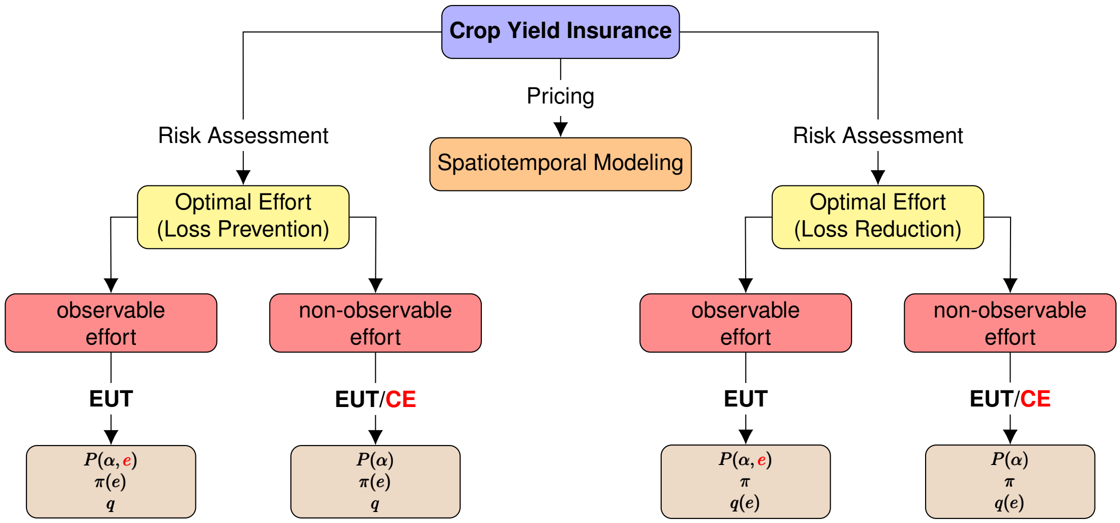

Figure 1.2: The summary diagram for the fundamental parts of this study

The thesis is organized as follows:

In the introduction part, we explain our motivation of the thesis. We also define our problem and give a summary of the literature review.

Chapter 2 discusses the analysis of insurance demand within the frame of "principal-agent problem". After we investigate the general equilibrium case for the insurance market, we economically analyze the equilibrium situation which maximizes the expected utilities of both the insured and the insurer. We also show the impact of the adverse selection on the equilibrium cases in this chapter.

In Chapter 3, we consider a crop insurance portfolio in order to provide the model setting. This chapter investigates the effects of asymmetric information on optimal yield insurance under various cases. The implementations of the maximization problem, which aims to find the farmer's optimal effort maximizing that farmer's expected utility, is based on the fact that the effort level of the farmer has an impact on both the probability and the amount of the yield loss. This study differs from the previous ones due to examining "asymmetric information" with respect to the "optimal effort" of the farmer in a crop yield insurance portfolio. The results obtained under EUT are also interpreted numerically and graphically by using CE approach.

Chapter 4 introduces the structure of the district-based yield and discusses the importance of the dependency between adjacent districts. The spatial or temporal dependencies among crop yields are reflected in the hierarchical Bayesian method which models the conditional distribution of the crop yield. We also introduce the approaches used for the analysis of model performance and model selection within the frame of Bayesian modelling.

In Chapter 5, we conduct a case study consisting of 47 districts in Ankara and Konya, the cities where the wheat production is the highest in Turkey, for the years 2004-2018. We examine the dependency between specified subregions by considering spatial and temporal effects. We model spatial, temporal and spatiotemporal effects for the estimation of the yield of wheat in Ankara and Konya. For this purpose, we use hiearchical Bayesian structure for the district-based crop yield data to handle space and time effects. After choosing the best model, we compute the premium rates related to various coverage levels by using this model.

Finally, we make our concluding remarks and present our ideas for further research in Chapter 6.

1.2. Literature Review

There have been numerous studies which investigate asymmetric information in crop insurance. During the 1970s, asymmetric information is discussed by Akerlof [1], Holmstrom et al. [2] and Raviv [3] for modelling the market failure. Rothschild and Stiglitz [4] and Stiglitz [5] examine adverse selection related to insurance market for competition and monopoly, respectively. Liu and Browne [6] present extentions of the paper of Rothschild and Stiglitz [4] and show the effect of tansaction costs in the insurance markets with "adverse selection". Boyer and Peter [7] represent a model for competitive insurance market, including the relationship between adverse selection and insurance fraud. Chambers [8] investigates the effect of moral hazard on indemnity payments. This study shows that the probability of loss will increase if the insurer can not observe the insured's actions. In this contex, it is inadequate indemnity payment in addition to the administrative costs since the premium is not defined accurately.

Coble et al. [9] investigate the insurance decisions of Kansas's farmer under moral hazard. Then, they claim that the moral hazard effect arise in poor production year for the farmers. Goodwin [10] examines factors which affect farmers' decisions for crop insurance. In this study, it is aimed to find the demand elasticities for the crop insurance. Due et al. [11] present the farmers make less effort with higher subsidies and lower premium rates. Chambers and Quiggin [12] study the effects of yield insurance by using a farmer's risk preferences in crop production. They conclude that there are not notable differences between other financial management tools and yield insurance. Gunnsteinsson [13] estimates the asymmetric information according to the different preferences for the same farmer. In this study, Gunnsteinsson represents that farmers use less inputs on insured areas. The optimal crop insurance contract and its implications are studied by Smith and Goodwin [14] Coble et al. [15], Mahul [16] and Ligon [17].

The considerations about moral hazard given by Eeckhoudt and Gollier [18] and Jaspersen and Richter [19] enable us to model moral hazard problem for the crop insurance contract. We extend these studies to find optimal effort for the farmers under loss prevention and loss reduction perspective.

Some studies consider the certainty equivalent as an alternative to EUT. Carter et al. [20] analyze the farmer's preferences based on "Constant Relative Risk Aversion". Berg [21] use stochastic optimization model to evaluate the yield insurance and the revenue insurance using mean and variance approach.

Ehrlich and Becker investigate [22] "self-insurance" and "self-protection" under EU hypothesis. In their paper, They represent some interesting conclusions the relationship between "market insurance", "self-insurance" and "self-protection". Subsequently, Dionne and Eeckhoudt [23] examine these insurance types according to risk behaviors. They show that a risk-averse insured demands more self-insurance than self-protection.

Donder et al. [24] propose a simple model with asymmetric information based on the term risk averse. They show that risk averse individuals employ more actions to mitigate their risk and are more willing to pay for insurance.

Roll [25] investigates a farmer's technical efficiency based on input use and yield using a stochastic frontier approach and shows that moral hazard exists in Norwegian salmon farming industry.

Over the past few decades, crop insurance has a significant position in the agriculture market. Therefore, the choice of a true statistical model which affects the distribution of the crop yield is a significant consideration to obtain reasonable premium rate. There have been many statistical approaches considered to determine the distribution of agricultural yields [26–29]. Nelson and Preckel [30] and Tirrapatur et al. [31] use “the beta distribution” to determine crop yield distribution. Jung and Ramezani [32] and Stokes [33] analyze the crop yield for the revenue insurance using the log-normal distribution.

The crop yields may vary by location, time and type of crop. Hence, it is not possible the use of classical parametric and nonparametric approachs for the modelling crop yield. In that case, using the models including spatial, temporal and spatiotemporal characteristics will be more appropriate for the crop yield. Vazaki et al. [34] investigate the spatio-temporal effects for Brazilian yield data for maize. They evaluate the premium rates for the state of Paraná. Saengseedam and Kantanantha [35] propose a linear mixed model with spatio-temporal process for the data of rice and cassava in Thailand. Park et al. [36] analyze the distribution for the crop yield of Iowa corn and Oklahoma wheat. A model is proposed using the Bayesian Kriging approach in order to estimate the parameters of the distribution. Moreover, the spatial effect across the related regions is examined.

“The Integrated Nested Laplace Approximation” approach has been recently proposed by Rue et al. [37]. It has a very broad and extensible class of models such as “linear mixed model” and spatiotemporal models [38–40]. Ruiz-Cárdenas and Krainski [41] propose a model to determine the premium rates of the crop yield insurance in Paraná state (Brazil). A dynamic spatio-temporal model is used under a Bayesian hierarchical framework.

2. ANALYSIS OF DEMAND FOR AN INSURANCE MARKET EQUILIBRIUM

2.1. Introduction

Since various expectations, preferences and justifications exist in the decision-making of buying insurance, building an insurance product which meets all needs of each individual is very difficult. The determination of the premium and the benefit that are two main factors of an insurance policy is very essential for the financial situation of the insurance company. In the principal-agent problem, the primary purpose for both the principal and the agent is maximizing their expected utilities. The insurer expects that individuals demand the offered insurance policy whereas the insured pays attention to the utility derived from the price and the indemnity of the policy. In this sense, we handle a relation between the demand and the supply of insurance. Since it is not sufficient to analyze this relationship under classical economic theory, we interpret this problem by taking the consideration of the dynamics in the insurance market.

For this purpose, we firstly aim to answer the question "How the insurance market works?" in Section 2.2. Within the frame of this question, we handle the equilibrium context in the insurance market. After we represent the required conditions which generate this equilibrium, we examine the drawbacks if an equilibrium failure occurs. After we analyze the equilibrium case for the insurance market in a general case, we discuss the insurance demand for the equilibrium situation which is a maximization problem taking consideration of both the insured's and the insurer's expected utilities [4, 42]. Section 2.3 discusses the influence of the models investigated in the previous section on the equilibrium cases for heterogenous risk groups. Lastly, this chapter is summarized in Section 2.4.

2.2. The Equilibrium in Insurance Market

Suppose that the random variables representing the wealth of an insured in the case where loss does not occur and loss occurs are S_1 and S_2, respectively. The consumption vector is positive, i.e. $(S_1, S_2) \in \mathbb{R}_+^2$ where $0 < S_2 < S_1$.

We firstly assume that the individual is not covered by an insurance. In the case that there is no loss, the wealth of this individual is as follows.

$$S_1 = w_0$$

where w_0 is the initial wealth. On the other hand, the wealth is given as follows if a loss occurs.

$$S_2 = w_0 - d.$$

Here, d repesents the loss amount. If π denotes the probability of loss occurance, the wealth of the individual whose initial wealth is w_0 is represented as

$$S_0 = \begin{cases} \pi & ;\ w_0 - d \\ 1-\pi & ;\ w_0 \end{cases}$$

Therefore, the expected wealth of the individual if no insurance exists is given as follows.

$$E_0\left(S_1, S_2\right) = (1-\pi)w_0 + \pi\left(w_0 - d\right) = w_0 - \pi d \tag{2.1}$$

where πd is the amount of the expected decrease in the individual's wealth. In addition, the expected utility of the individual is given as follows in this case.

$$U_0\left(S_1, S_2\right) = (1-\pi)u\left(w_0\right) + \pi u\left(w_0 - d\right) \tag{2.2}$$

On the other hand, we define an insurance contract as $C(P_I, I)$. Suppose that P_I is the insurance premium and I is the benefit which will be paid to the insured in case of loss. For the case that an insurance exists, we represent the change in the insured wealth as follows.

$$S = \begin{cases} \pi & ;\ w_0 - P_I - d + I \\ 1-\pi & ;\ w_0 - P_I \end{cases}$$

The expected utility of the insured is given as follows in this case.

$$U\left(S_1, S_2\right) = (1-\pi)u\left(w_0 - P_I\right) + \pi u\left(w_0 - P_I - d + I\right) \tag{2.3}$$

The insured's wealth according to the loss occurance is given as follows.

$$S_1 = w_0 - P_I$$
$$S_2 = w_0 - P_I - d + I$$

For all mentioned cases, the probability of loss π is assumed to be known by the insurer when asymmetric information does not exist. However, the insured has more information about π since the insured could observe his/her risk behaviours better in real life. On the other hand, the insurer could only predict the distribution of π by using the loss probabilities of the insureds in its own portfolio. When the insurer does not have a complete information about the probability of loss, some individuals having high loss probabilities (π_h) pretend to be low-risk individuals (π_l) in order to have an insurance with low premium rates. This situation could unbalance the asset-liability equilibrium of the insurance company. Covering the high-risk individuals with low premiums and paying high claims when loss occurs could increase the insurer's probability of ruin. This situation, analyzing or modeling the insureds' risk insufficiently, is called "adverse selection problem". Therefore, the insurer needs to offer policies which also considers the insured's own risk characteristics. In addition, the case which affects the insured's loss probability is called "moral hazard" is related to the insured's effort. The insured's effort for no loss occurrence will decrease the probability of loss π. We basically represent the moral hazard case by giving the insured's expected utility as follows.

$$U(S_1, S_2) = (1 - \pi(e))u(w_0 - P_I) + \pi(e)u(w_0 - P_I - d + I) - c(e) \tag{2.4}$$

Here, e denotes the effort that insured makes for no loss occurance whereas $c(e)$ indicates the cost of the effort e to the insurer. If the first derivative of the probability of loss $\pi'(e) < 0$, this means that $\pi(e)$ will decrease as the insured's effort e increase. On the other hand, if the effort increases, the cost of this effort to the insurer will also increase when $c'(e) > 0$. The impact of the insured's effort is analyzed for the crop yield insurance in Chapter 3 in detail.

In order to get an insurance policy, an individual requires to have more expected utility in the case where he/she has an insurance rather than having no insurance. This condition which is called “participation constraint” is represented as given.

$$(1-\pi)u\left(w_0-P_I\right)+\pi u\left(w_0-P_I-d+I\right)>(1-\pi)u\left(w_0\right)+\pi u\left(w_0-d\right)$$

Here, the insured’s utility function $u(.)$ has “*Von Neumann-Morgenstern utility function*” properties: $u(.)$ is an increasing and concave function. The positive first derivative of the utility function $u'(S)>0$ indicates that the insured prefers more wealth rather than less wealth. In addition to this, the negative second derivative of the utility function $u''(S)<0$ is the sufficient condition for the function’s being concave. In this study, we assume that the insured is risk-averse. The utility function’s being strictly concave, $u''(S)<0$, denotes that the insured is absolutely risk-averse. The condition for a concave utility function is given in the following definition.

Definition 2.2.1. *If the following inequality holds for all* w_1, w_2 *and* $\pi : 0 \leq \pi \leq 1$*, then* $u(.)$ *is a concave function.*

$$u\left(\pi w_1+(1-\pi)w_2\right)\geq \pi u\left(w_1\right)+(1-\pi)u\left(w_2\right)$$

A numerical example for specific values of w_1, w_2 and π is given as follows.

Example 2.2.1. *Define the utility function as* $u(w)=\sqrt{w}$*. Let the probability of loss be* $\pi=0.5$ *and* $w_1=64, w_2=16$*. Using the condition in Definition 2.2.1, we show that* $u(w)$ *is concave in Figure 2.1.*

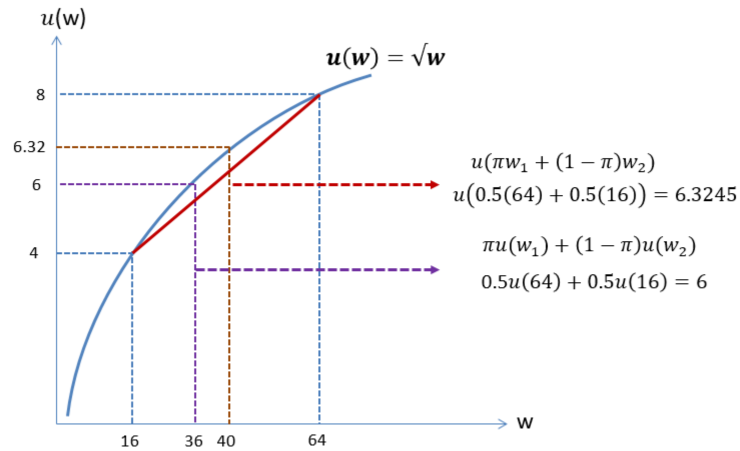

Figure 2.1: Representation of concave function

2.2.1. The analysis of the demand in insurance market

If the insured has an insurance, he/she pays the premium P_I for the the indemnity I which is paid according to the loss amount d. In this case, the insured's wealth will be $S_1 = w_0 - P_I$ if no loss occurs while it is $S_2 = w_0 - P_I - d + I$ if loss occurs. In this study, we consider the state contingent wealths S_1 and S_2 as "consumer goods" and the insured as "consumer" which are defined in classical microeconomics. Therefore, we can examine the demand for insurance in the frame of economics more easily.

We previously define an insurance contract as $C(P_I, I)$ supposing that P_I is the insurance premium and I is the indemnity paid to the insured when damage occurs. The reason for the use of subscript I in the notation P_I is that the premium amount is usually obtained by multiplying the indemnity with a determined premium rate (p). In other words, the equation $P_I = pI$ represents the premium which is paid by the insured.

The ultimate aim for an insurance company is not to lose money due to the policy it sells. In this section, we assume that the insurance company operates in a competitive market. The following condition is needed to be fulfilled for the insurer so that not losing money.

$$P_I - pI \geq 0$$

Since the insurance market is assumed to be competitive, the above condition turns out to be below equality.

$$P_I - pI = 0$$

When we replace the above equality in the definition of state contingent wealths S_1 and S_2, we rewrite these equations as follows.

$$S_1 = w_0 - pI, \tag{2.5}$$

and

$$S_2 = w_0 - pI - d + I \tag{2.6}$$

If Equation (2.5) is solved for I and replaced in Equation (2.6), the following solution is obtained.

$$\begin{aligned} I &= (w_0 - S_1)/p \\ S_2 &= w_0 - d + \frac{(1-p)(w_0 - S_1)}{p} \\ S_2 &= \frac{pw_0 - pd + w_0 - S_1 - pw_0 + pS_1}{p} \\ S_2 &= \frac{w_0 - pd}{p} - \frac{(1-p)S_1}{p} \end{aligned}$$

The above solution is summarized as follows.

$$(1-p)S_1 + pS_2 = w_0 - pd \tag{2.7}$$

Equation (2.7) represents the budget line for the insured. $(1-p)$ denotes the price of S_1 when no loss occurs whereas p is the price of S_2 when loss occurs. Here, the slope of the budget line is $-\frac{(1-p)}{p}$. Equation (2.7) is also stated as “fair odds line” in the literature. The graphical represantion of this equation is given below.

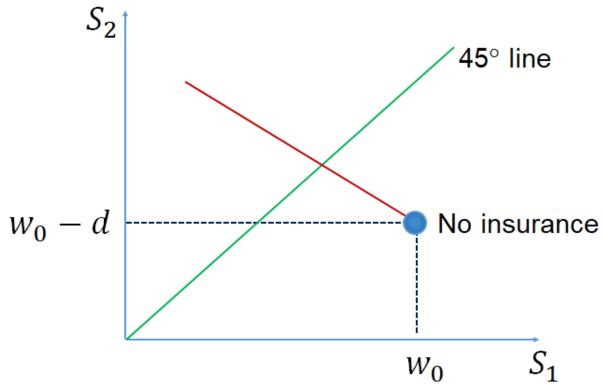

Figure 2.2: The budget line (fair odds line)

In Figure 2.2, the indifference curve obtained by using the utility function provides the same utility level for different commodity compositions which are denoted by S_1 and S_2 in our study. We will use the indifference curve to find the optimal insurance components. The insured's expected utility with insurance which is given in Equation (2.3) is edited for S_1 and S_2 as follows.

$$U(S_1, S_2) = (1-\pi)u(S_1) + \pi u(S_2) \tag{2.8}$$

Since the indifference curve provides the same utility level, we consider S_1 as a function of S_2 to find the equation for this curve. In this case, Equation (2.8) can be rewritten as follows.

$$(1-\pi)u(S_1(S_2)) + \pi u(S_2) = \bar{u}$$

Here, $\bar{u}$ indicates the utility level $U(S_1, S_2)$. The slope of the indifference curve at one point is obtained as

$$-\frac{\partial S_2}{\partial S_1} = \frac{(1-\pi)}{\pi}\frac{u'(S_1)}{u'(S_2)} \tag{2.9}$$

Equation (2.9) is also stated as "marginal rate of substitution" with the notation $\text{MRS}_{(S_1,S_2)}$. When we apply the "first order condition" to maximize the expected utility function under the insured's budget constraint condition to Equation (2.8), the solution of the problem in Equation (2.9) is given as below.

$$\frac{(1-\pi)}{\pi}\frac{u'(S_1)}{u'(S_2)} = \frac{(1-p)}{p} \tag{2.10}$$

The solution in Equation (2.10) is obtained using the Lagrange multipliers method given as follows.

$$\mathcal{L}(S_1, S_2, \lambda) = U(S_1, S_2) + \lambda (w_0 - pd - ((1-p)S_1 + pS_2)$$

Since $S_1 = S_2$ along the $45°$ line in Figure 2.2, the first derivations are equal, i.e. $u'(S_1) = u'(S_2)$, and therefore we obtain $\pi = p$ in Equation (2.10). This situation shows the case of "full insurance". The point where the budget constraint line and the indifference curve intersect on the $45°$ line is expressed as the *equilibrium point* for the insurance. Also, $d = I$ is obtained as follows for the full insurance case from the equation $S_1 = S_2$.

$$w_0 - pd = w_0 - pI - d + I$$
$$d(1-p) = (1-p)I$$
$$d = I$$

This equation indicates that the indemnity to be paid for the case of full insurance is equal to the occured loss amount. This situation is presented in the following figure.

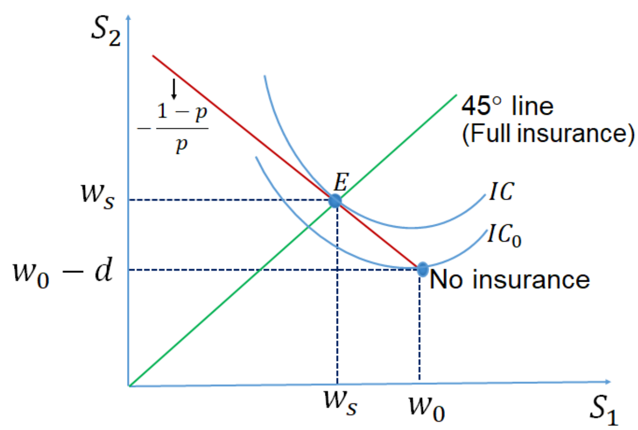

Figure 2.3: The equilibrium for the case of full insurance

In Figure 2.3, the point E where $S_1 = S_2 = w_s$ presents the equilibrium point for the full insurance case. Here, IC and IC_0 indicate indifference curves with insurance and no insurance, respectively. A numerical example for specific values of w, π and d is given as follows.

Example 2.2.2. *Define the utility function as* $u(s) = \sqrt{s}$*. Let the insured's initial wealth be*

$w = 40000$ and the probability of loss be $\pi = 0.25$. We assume that the loss amount is $d = 20000$ if loss occurs. We aim to determine the optimal premium to be asked from the insured and the amount of indemnity to be paid for the insurance company operating in a competitive market.

The point where the insured's utility is maximum is the point where the budget constraint line and the indifference curve intersect on the $45°$ line. First of all, we need to find the indifference curve and the budget line for the insured. The insured's utility function is obtained by Equation (2.8) as follows.

$$U(S_1, S_2) = (1 - 0.25)u(S_1) + 025u(S_2)$$

Since $S_1 = 40000 - P_I$ and $S_2 = 40000 - P_I - 20000 + I$, the budget line is found below using Equation (2.7).

$$S_2 = \frac{40000 - 0.25(20000)}{0.25} - \frac{(1 - 0.25)S_1}{0.25}$$
$$S_2 = 140000 - 3S_1(\textit{ budget constraint line})$$

In order to obtain the indifference curve for the insured, we use the following equation.

$$(1 - 0.25)\sqrt{S_1} + 0.25\sqrt{S_2} = \bar{u}$$

where $u(s) = \sqrt{s}$. Here, if S_1 is written as a function of S_2, we have

$$S_2 = \left(\frac{\bar{u} - 0.75\sqrt{S_1}}{0.25}\right)^2$$

The above equation is the indifference curve for the insured. By definition of the indifference curve, it gives the same utility level at all points on the curve. Therefore, $S_1 = S_2$ at the point where the indifference curve intersects the $45°$ line. If the equation $S_1 = S_2$ is replaced in the budget constraint line, the following solutions are obtained.

$$S_2 = 140000 - 3S_2 \Rightarrow S_1^*, S_2^* = S^* = 35000$$

where S^* indicates the equilibrium point. We could obtain the premium amount from $S_1 = 40000 - P_I$ as $P_I = 5000$ and the indemnity amount from $S_2 = 40000 - P_I - 20000 + I$ as

$I = 20000$. *The expected utility at this equilibrium point is*

$$(1-0.25)\sqrt{35000} + 0.25\sqrt{35000} = \bar{u}$$
$$\bar{u} = 187.0829.$$

As a result, $S_2 = \left(\frac{187.0829-0.75\sqrt{S_1}}{0.25}\right)^2$ *represents the indifference curve for the insured. The graph of the budget constraint line and the indifference curve are shown in Figure 2.4.*

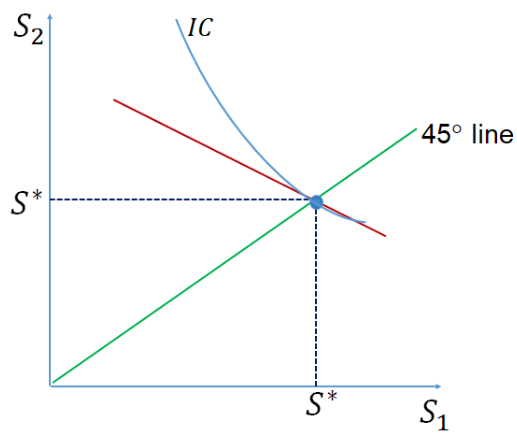

Figure 2.4: The equilibrium for the case of full insurance: A numerical example

After discussing the equilibrium for the full insurance case, we also examine equilibrium situation for the cases such as over-insurance and under-insurance where full insurance does not exist. For this aim, Equation (2.10) is reorganized as follows.

$$u'(S_1) = u'(S_2)\,\frac{\pi(1-p)}{(1-\pi)p} \tag{2.11}$$

Bu using Equation 2.11, we obtain the following results.

i. Full insurance: $\pi = p \Leftrightarrow u'(S_1) = u'(S_2) \Leftrightarrow S_1 = S_2$

ii. Over-insurance: $\pi > p \Leftrightarrow u'(S_1) > u'(S_2) \Leftrightarrow S_1 < S_2$

iii. Under-insurance: $\pi < p \Leftrightarrow u'(S_1) < u'(S_2) \Leftrightarrow S_1 > S_2$

The first result shows that insured's wealths are equal whether there is loss or not, that is, full insurance. Here, there is a full coverage for the insured against loss that may occur. Since the insured is risk-averse, $u'(.)$ decreases as the wealth increases. Because of this property, the second result states that the wealth for the case of loss is greater and this relationship provides more coverage than the full insurance does which turns out to be over-insurance. The final result shows that the insured's wealth for the case of no loss is greater, that is, under-insurance. The graphical representation related to these results is given below.

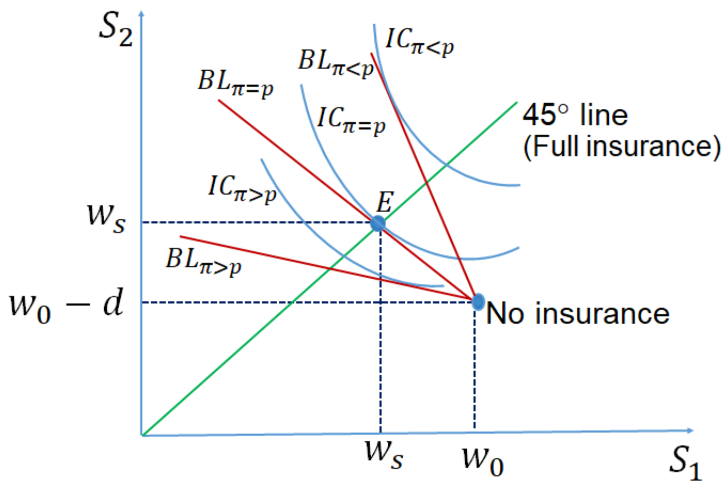

Figure 2.5: Review of equilibrium for the cases of full insurance, under-insurance and over-insurance

In Figure 2.5, the point E shows the full insurance case. At this point, the profit of the insurance company is 0 (*zero profit condition*). The insurance company makes a loss on the budget line $BL_{\pi=p}$ which provides the equilibrium whereas the company makes profit in the areas below this line. The insured has more expected utility at every point on the line $BL_{\pi=p}$ and the related indifference curve for each case (provided that the budget constraint is also provided).

2.2.2. Comparative statistics for the components of an insurance policy

Considering $P_I = pI$, the insured's expected utility function is rewritten below.

$$U(S_1, S_2) = (1-\pi)u(w_0 - pI) + \pi u(w_0 - pI - d + I)$$

By the help of this equation, economic interpretations of how the indemnity I, which will be paid to the insured in case of loss, changes according to the components w_0, P_I and d. The discussion of "what kind of economic goods the insurance policy is" is explained using the comparative statistics given below [42].

If the derivative of the insured's expected utility function according to I is equal to 0, the demand model for the indemnity to be paid to the insured is obtained in the following equation.

$$U'_I\left(S_1, S_2\right) = -p(1-\pi)u'\left(w_0 - pI\right) + (1-p)\pi u'\left(w_0 - d + (1-p)I\right) \tag{2.12}$$

Using Equation (2.12), it can be analyzed how indemnity changes according to other variables. The impact of a change in initial wealth w_0 on indemnity is given below.

$$\frac{\partial I}{\partial w_o} = -\frac{U'_{Iw_0}\left(S_1, S_2\right)}{U'_{II}\left(S_1, S_2\right)} \tag{2.13}$$

Here, the numerator and the denominator of the right-hand side of Equation (2.13), which are given below, represents the derivative of Equation (2.12) according to w_0 and I, respectively.

$$U'_{Iw_0}\left(S_1, S_2\right) = -p(1-\pi)u''\left(w_0 - pI\right) + (1-p)\pi u''\left(w_0 - d + (1-p)I\right) \tag{2.14}$$

$$U'_{II}\left(S_1, S_2\right) = p^2(1-\pi)u''\left(w_0 - pI\right) + (1-p)^2\pi u''\left(w_0 - d + (1-p)I\right) \tag{2.15}$$

If we replace these two equations in Equation (2.13), we obtain the following result.

$$\frac{\partial I}{\partial w_o} = -\frac{U'_{lw_0}\left(S_1, S_2\right)}{U'_{II}\left(S_1, S_2\right)} = -\frac{-p(1-\pi)u''\left(w_0 - pI\right) + (1-p)\pi u''\left(w_0 - d + (1-p)I\right)}{p^2(1-\pi)u''\left(w_0 - pI\right) + (1-p)^2\pi u''\left(w_0 - d + (1-p)I\right)} \tag{2.16}$$

In the case of $\pi = p$, $d = I$ since the insured has full insurance. Therefore, the value of the Equation (2.16) is 0. Hence, any change in w_0 will not have an impact on the insurance demand. On the other hand, it is previously mentioned that the case of $p > \pi$ is the under-insurance case. For this case, the sign of the term $U'_{lw_0}\left(S_1, S_2\right)$ in the numerator of the above equation is not certain. The term in the denominator is always negative because $u''(S) < 0$. Since the sign of the numerator is not known, it is difficult to find a relationship between I and w_0. Rewriting

Equation (2.10),we have

$$p(1-\pi)=\frac{(1-p)\pi u'\left(S_2\right)}{u'\left(S_1\right)}$$

When we replace this expression in Equation (2.14), we obtain the result given below where $S_1=w_0-P_I$ and $S_2=w_0-P_I-d+I$.

$$\begin{aligned} U'_{Iw_0}\left(S_1,S_2\right) &= -u''\left(S_1\right)\tfrac{(1-p)\pi u'(S_2)}{u'(S_1)}+u''\left(S_2\right)\tfrac{(1-p)\pi u'(S_1)}{u'(S_2)} \\ &= (1-p)\pi u'\left(S_2\right)\left[\tfrac{u''(S_2)}{u'(S_2)}-\tfrac{u''(S_1)}{u'(S_1)}\right] \end{aligned} \tag{2.17}$$

We can express the above equation in terms of risk aversion coefficient. Risk aversion coefficient is defined as follows.

$$r_S=-\frac{u''(S)}{u'(S)}$$

Hence, Equation (2.17) is rewritten for the risk aversion coefficient as below.

$$U'_{Iw_0}\left(S_1,S_2\right)=(1-p)\pi u'\left(S_2\right)\left[r_{S_1}-r_{S_2}\right] \tag{2.18}$$

According to the relationship between r_{S_1} and r_{S_2}, the sign of the term $U'_{Iw_0}\left(S_1,S_2\right)$ is determined. If the insured's risk aversion coefficient r_S is increasing, the relationship $r_{S_1}>r_{S_2}$ is obtained since $S_1>S_2$ for $p>\pi$. Hence, the sign of the term $U'_{Iw_0}\left(S_1,S_2\right)$ is zero or positive. By the help of the results, it can be concluded that the change of indemnity according to inital wealth, i.e. $\frac{\partial I}{\partial w_o}$ in Equation (2.16) will be in the same direction. This result shows that the insurance demand will increase as the initial wealth increases. We can make an economic inference from this result as well. In economics, if the demand for goods increases as the individual's income increases, goods are defined as "normal goods". Therefore, insurance policies can be economically defined as "normal goods" if r_S is increasing.

On the other hand, the direction of the relation will be reversed if r_S is decreasing. In this case, insurance policies can be economically defined as "inferior goods" since the insurance demand will decrease as the initial wealth increases, i.e. $\frac{\partial I}{\partial w_o}<0$. We interpret these results under the assumption that other variables apart from I and w_0 are constant (*ceteris paribus*).

We examine the change of the indemnity amount I according to the loss amount d by using Equation (2.12). The impact of change in d on indemnity is given below.

$$\frac{\partial I}{\partial d} = -\frac{U'_{Id}\left(S_1, S_2\right)}{U'_{II}\left(S_1, S_2\right)} \tag{2.19}$$

Here, $U'_{Id}\left(S_1, S_2\right) = -(1-p)\pi u''\left(w_0 - d + (1-p)I\right)$ is the derivative of Equation (2.12) with respect to d. Since $u''(S) < 0$, the numerator is positive, i.e. $U'_{Id}\left(S_1, S_2\right) > 0$. Therefore, the amount of indemnity to be paid for the insurance will increase as the amount d increases. Again, there is an assumption that other variables remain the same.

On the other hand, the relationship between I and the premium rate p is examined by the help of the following equation.

$$\frac{\partial I}{\partial p} = -\frac{U'_{Ip}\left(S_1, S_2\right)}{U'_{II}\left(S_1, S_2\right)} \tag{2.20}$$

Here, $U'_{Ip}\left(S_1, S_2\right)$ is the derivative of Equation (2.12) with respect to p and it is obtained as follows.

$$\begin{aligned} U'_{Ip}\left(S_1, S_2\right) &= -(1-\pi)u'\left(S_1\right) + p(1-\pi)Iu''\left(S_1\right) - \pi u'\left(S_2\right) - (1-p)\pi I u''\left(S_2\right) \\ &= -\left[(1-\pi)u'\left(S_1\right) + \pi u'\left(S_2\right)\right] + I\left[p(1-\pi)u''\left(S_1\right) - (1-p)\pi u''\left(S_2\right)\right] \end{aligned} \tag{2.21}$$

According to this resutlt,we rewrite Equation (2.20) below.

$$\frac{\partial I}{\partial p} = \frac{\left[(1-\pi)u'\left(S_1\right) + \pi u'\left(S_2\right)\right]}{U'_{II}\left(S_1, S_2\right)} + I\frac{\left[-p(1-\pi)u''\left(S_1\right) + (1-p)\pi u''\left(S_2\right)\right]}{U'_{II}\left(S_1, S_2\right)}$$

The term $\left[-p(1-\pi)u''\left(S_1\right) + (1-p)\pi u''\left(S_2\right)\right]$ on the right hand side of the above equation equals the expression $U'_{Iw_0}\left(S_1, S_2\right)$ given in Equation (2.14). Therefore, the above equation can be obtained as follows.

$$\frac{\partial I}{\partial p} = \frac{\left[(1-\pi)u'\left(S_1\right) + \pi u'\left(S_2\right)\right]}{U'_{II}\left(S_1, S_2\right)} + I\frac{U'_{Iw_0}\left(S_1, S_2\right)}{U'_{II}\left(S_1, S_2\right)} \tag{2.22}$$

The first term on the right hand side of the above equation indicates the substitution effect between S_1 and S_2 whereas the second term denotes the effect of the initial wealth w_0. Since $U'_{II}\left(S_1, S_2\right)$ is always negative, the substitution effect is negative. In addition, it is shown above that the term $U'_{Iw_0}(S_1, S_2)$ is positive or negative according to the cases for the risk aversion coefficient. If $U'_{Iw_0}(S_1, S_2) > 0$, then the effect of wealth in Equation (2.22) is negative. Therefore, the direction of change in the term $\frac{\partial I}{\partial p}$ in Equation (2.22) is opposite. In other words, the insurance demand for indemnity decreases as the value of premium ratio p increases.

2.3. The Equilibrium in Insurance Market for Heterogeneous Risk Groups

In this section, we deal with equilibrium cases for different risk groups. Here, heterogeneity implies that the probability of loss π varies according to the riskiness of the insured. We assume in this section that there are two types of insureds with different risk structures. In this context, the loss probabilities of high-risk and low-risk insureds are denoted as π_h and π_l respectively ($\pi_h > \pi_l$).

In addition, it is assumed that the insureds' initial wealth and loss amounts are equal and they have the same utility function $u(.)$. In order to build a model for individuals with different risks, we deal with the asymmetric information where the insureds' loss probabilities are not known by the insurer but known only by insureds. On the other hand, heterogeneity will not matter in the case of complete information according to the assumption mentioned above since the insureds have the same risk characteristics.

In case of heterogeneity, *"Rothschild-Stiglitz equilibrium"* assumptions are taken into account in the analyze of equilibrium conditions. These assumptions are given below [4]:

i. Each insured prefers the insurance policy that maximizes their expected utilities. In other words, the insured's expected utility with insurance should be higher than the expected utility when no policy exists.

ii. The insurer's profit should be 0 because of the competitive market assumption.

iii. Apart from the insurance policy that the insurer offer, there should not be any policy that another insurance company could offer and make profit.

In the case of full insurance, the probability of loss and the premium rate are equal, $\pi = p$.

For a competitive market, this equation also represents the actuarial fair price. Budget lines and indifference curves for high-risk and low-risk insureds are given below.

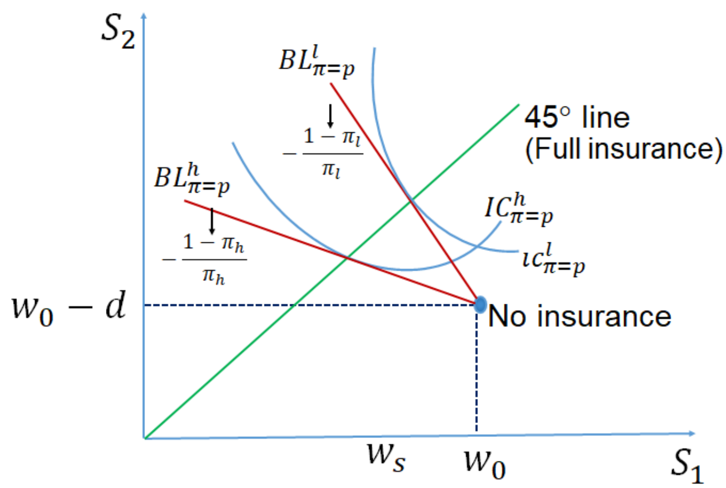

Figure 2.6: Budget lines and indifference curves for high-risk and low-risk insureds

As seen in Figure 2.6, the budget line and indifference curve are steeper for low-risk insureds. Let us explain the reason for this by using marginal substitution rates (MRS). The *MRS*s for high-risk and low-risk insureds are obtained respectively using Equation (2.9) as follows.

$$\text{MRS}^h_{S_1,S_2} = -\frac{\partial S_2}{\partial S_1} = \frac{(1-\pi_h)}{\pi_h}\frac{u'(S_1)}{u'(S_2)}$$

$$\text{MRS}^l_{S_1,S_2} = -\frac{\partial S_2}{\partial S_1} = \frac{(1-\pi_l)}{\pi_l}\frac{u'(S_1)}{u'(S_2)}$$

Since the insureds have the same characteristics except the probability of loss, we can compare the budget lines and indifference curves of high-risk and low-risk insureds when the first equation above is divided into the second equation.

$$\frac{\text{MRS}^h_{S_1,S_2}}{\text{MRS}^l_{S_1,S_2}} = \frac{(1-\pi_h)}{\pi_h}\frac{\pi_l}{(1-\pi_l)}$$

In the above equation, $\frac{\text{MRS}^h_{S_1,S_2}}{\text{MRS}^l_{S_1,S_2}} < 1 \Rightarrow \text{MRS}^l_{S_1,S_2} > \text{MRS}^h_{S_1,S_2}$ because of the relation $\pi_h > \pi_l$. This relationship provides steeper budget line and indifference curve for low-risk insureds.

For an insurance policy with asymmetric information, we consider an insurance contract $C(P_s, I_s)$ where the same policy is sold to all insureds. In Equation (2.23), we define the joint probability of loss for the policy that the insurer offer.

$$\pi_s = \delta\pi_h + (1-\delta)\pi_l \tag{2.23}$$

Here, δ parameter denotes the ratio of high-risk insureds (π_h) in the insurance portfolio. According to this insurance policy, the insured's budget line is obtained as follows using Equation (2.7).

$$(1-\pi_s)\, S_1 + \pi_s S_2 = w_0 - \pi_s d \tag{2.24}$$

Under the assumption that w_0 and d are the same for low-risk and high-risk insureds, the probability of π_s is used for both risk groups. The situation where the same policy is sold to all insureds is referred as "pooling equilibrium". All policies to be proposed in this equilibrium case must be on the budget line (fair odds line) obtained in Equation (2.24). The graphical representation of this situation is given below.

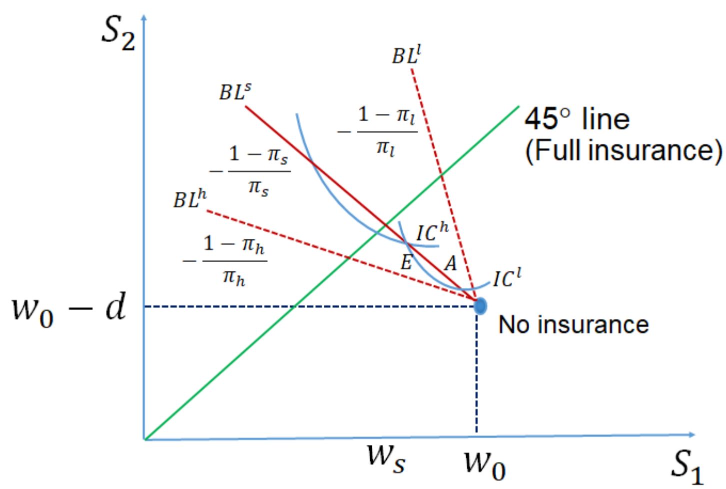

Figure 2.7: Pooling equilibrium

BL^s represents the budget line for a "pooling" policy in full insurance case. The "Policy E" in Figure 2.7 is a "pooling" policy since it is located on the line BL^s. This policy provides the same utility level for the high-risk and low-risk groups. The expected profit of the insurance company operating in a competitive market is 0 for the Policy E where the "pooling equilibrium" exists as well. Low-risk and high-risk insureds pay the same premium for this policy. Since the relation $\pi_h > \pi_l$ holds, high-risk insureds cause more losses. Therefore, the low-risk insureds' expected cost increase as they have to pay higher premiums for this policy. If another insurance company offer a policy above IC^l or below IC^h, low-risk individuals will choose this policy. "Policy A" is given as an example in Figure 2.7. Low-risk individuals will leave Policy E and prefer Policy A since Policy A is above the indifference curve IC^l and above the budget line for the "pooling" policy, and low-risk individuals will thus move away from equilibrium.

Instead of offering the same policy, the situation where different policies are offered to the insureds is called "separating policy". In order to analyze the separating policy case for insurance market, we examine the equilibrium situation for both low-risk and high-risk insureds together. In this case, the question is whether there exists a point which includes both risk groups and where equilibrium is managed. For this question, we examine the Policy D represented in the following figure.

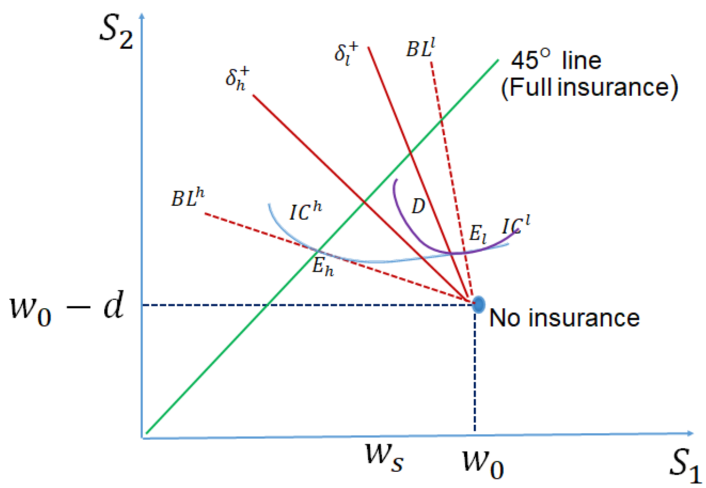

Figure 2.8: Separating equilibrium

In Figure 2.8, δ_h^+ and δ_l^+ indicates that the high-risk and low-risk insureds in the insurance portfolio are proportionally higher, respectively. Policy D is preferred because it is located above the indifference curve for both risk groups. If the majority of the insureds who purchase Policy D are low-risk insureds (δ_l^+), the insurance company will make profit (since Policy D is located under the line representing δ_l^+). Therefore, the positive profit of the insurer will disrupt the equilibrium. If the majority of high-risk insureds is higher, the insurer will not be able to make a positive profit as the Policy D is located above the budget line representing δ_h^+). In this case, the Policy D is the equilibrium point. As a result, if another insurance company which has the same equilibrium situation but has less number of high-risk insureds, this insurance company will have a positive profit opportunity by offering the Policy D. This situation shows that there is no "separating" equilibrium.

2.4. Interim Conclusion: Demand and Equilibrium in Insurance Market

In this chapter, we discuss the economic conditions for the insurance market. In this sense, we seek for answers of the following questions:

i. What does the demand mean for a general insurance policy?

ii. How does the insurance market work?

iii. How does the indemnity to be paid in case of loss according to the wealth, the premium amount and the loss amount?

iv. What are the equilibrium conditions for insurance under "Rothschild-Stiglitz equilibrium" assumptions?

These questions lead us to discuss asymmetric information in crop yield insurance in the next chapter as well .

3. IMPACTS OF MORAL HAZARD ON OPTIMAL CROP YIELD INSURANCE

3.1. Introduction

In this chapter, we concentrate on how to maximize the expected utility of a farmer at the end of an insurance period under a yield insurance. We consider that the insurer make a payment when the yield does not exceed a specified yield threshold.

The aim of this chapter is to examine the moral hazard effect in crop insurance for yield losses under EUT approach. We handle two cases for crop yield losses as Rees et al. [42] suggest. Firstly, the farmer is assumed to control the probability of occurence of a natural hazard (e.g. flood, drought) (*loss-prevention*) while he/she cannot control the distribution of losses. In the second case, we consider that the farmer has an influence on the size of the crop yield losses (*loss-reduction*) where the farmer might also have an influence on the distribution of crop yield losses. We also obtain numerical results and interpret them graphically under CE approach for both cases considering non-observable effort.

In this chapter, we discuss the idea of the farmer's optimal effort which is taken into account for the evaluation of the moral hazard of that farmer.

In Section 3.2 we introduce our main model setting of the yield insurance in order to examine the expected utility of the farmer under various cases. Section 3.3 provides an approach for the optimal effort under loss-prevention case within the frame of the optimal contract with observable and non-observable effort. In Section 3.4, we examine the optimal effort under loss-reduction case for observable/non-observable effort. Section 3.5 discusses how the farmer's moral hazard, the coverage rate and the risk aversion coefficient could be considered when the effort is not observable. Finally, Section 3.6 concludes the chapter.

3.2. Main Model Setting of the Yield Insurance

The expected utility model is a commonly used theory to study an individual's behaviors under uncertainty. For a risk averse farmer, it is assumed that von-Neumann-Morgenster utility's function U_I is strictly increasing, concave and continously twice differentiable, i.e, $U_I'(.) > 0, U_I''(.) < 0$. $U_I'(.)$ denotes the first derivative (marginal utility). The expected utility of the farmer is aimed be to maximized for the wealth at the end of the period.

In the crop yield insurance, the farmer's wealth at the end of the period depends on the cases where the indemnity payments are made or not. We define the wealth as follows:

$$W = W_0 + A[yp_c + I(\alpha) - [1 + \beta(\alpha)][1 - s(\alpha)]P(\alpha)], \tag{3.1}$$

where W_0 denotes initial wealth. y represents the farmer's observable crop yield per hectar $y \in [0, y_{max}]$ where y_{max} is the maximum yield. A and p_c denote harvested area and crop price per hectar, respectively. In this study, A and p_c are assumed to be constant and equal to 1. The crop insurance premium, $P(\alpha)$, represents the cost of purchasing insurance to the coverage level α. Also, $\beta(\alpha)$ denotes the premium loading factor which is taken for the transaction costs whereas $s(\alpha)$ indicates the premium subsidy rate which varies with respect to the chosen coverage level α.

In order to determine the value of the indemnity payment $I(\alpha)$ for the yield insurance, we consider two cases. If the yield y falls below a strike level yield αy^*, then the indemnity payment $I(\alpha)$ is made by the insurer, otherwise no payment is made. Here, y^* is the long-term average of the yield. Therefore, we model the indemnity payment $I(\alpha)$ stochastically as if the yield insurance is a put option, i.e. $I(\alpha) = \max(\alpha y^* - y, 0)$. For the simplicity, we use the following notations to express the end-of-period wealth of the farmer.

$$W_l = W_0 + y + (\alpha y^* - y) - [1 + \beta(\alpha)][1 - s(\alpha)]P(\alpha), \ y < \alpha y^* \tag{3.2}$$

$$W_h = W_0 + y - [1 + \beta(\alpha)][1 - s(\alpha)]P(\alpha), \ y \geq \alpha y^* \tag{3.3}$$

where W_l and W_h denote end-of-period wealth below and above the strike level yield, respectively. W_h does not consist of an indemnity payment since the yield y is higher than the strike level yield αy^*. Equations (3.2) and (3.3) simply represent the wealth at the end of the period where the indemnity payments is added to and the premium payments substracted from the initial wealth. Hence, we could describe the insured's profit (Π_I) from the crop yield insurance contract by the following equation:

$$\Pi_I = \begin{cases} (\alpha y^* - y) - [1 + \beta(\alpha)][1 - s(\alpha)]P(\alpha), & \text{if } y < \alpha y^* \\ -[1 + \beta(\alpha)][1 - s(\alpha)]P(\alpha), & \text{if } y \geq \alpha y^* \end{cases} \tag{3.4}$$

3.3. Optimal Effort Level under Loss-Prevention

Having introduced the end-of-period wealth, we extend the model according to the consideration of moral hazard. We assume that, after accepting the insurance contract, the farmer chooses an effort level $e \in [0, \infty]$, to avoid the effects of natural disasters. This effort level e lowers the probability of risk occuring $\pi(e) \in (0, 1)$. Here, we assume that $\pi(e)$ is strictly decreasing and convex, i.e. $\pi'(e) < 0, \pi''(e) > 0$ since more effort e causes lower probability of loss and the same change in the effort amount where e gets greater leads to smaller changes in the probability of risk event.

For a general insurance contract, in order to define the expected utility of the insured in case the insurance contract is agreed, we first assume that the insured is risk averse with the initial wealth W_0. In this situation , the insured pays the premium P to be covered against a random loss d. Hence, the insured will take an indemnity payment I if d occurs. In this sense, the expected utility of the insured is defined as:

$$E(U) = [1 - \pi(e)]U(W_0 - P) + \pi(e)U(W_0 - P - d + I) - c(e) \tag{3.5}$$

where $U(.)$ is the utility function of the insured. Here, $c(e)$ represents the cost of effort function e with the assumption that it is twice differentiable, strictly increasing and strictly convex, i.e. $c'(e) > 0$ and $c''(e) > 0$. It does not depend on whether the risk occurs or not. In this model we assume that $\pi(e)$ is controlled by the insured, which is the idea behind the loss-prevention case. As seen in Equation (3.5), the utility of the insured depends on two cases: (a) the risk occurs with the probability $\pi(e)$, and (b) the risk does not occur with the probability $1 - \pi(e)$. Then, if the case (a) arises, the insured will receive the indemnity payment. However, the states happen to be different for a crop yield insurance. In order to be paid as an insured, the case that the risk occurs is not enough, the yield y must also fall below the strike level yield αy^* to receive the indemnity payment. If the risk occurs, it could affect the yield, or not. As a result, we define three states: (*i*) "Risk with low yield" case ($y < \alpha y^*$) with possibility q, (*ii*) "Risk with high yield" case and no-risk ($y \geq \alpha y^*$), and (*iii*) "no-risk " case. The payoffs to the insured for these states under crop yield insurance are given separately in Table 3.1:

Table 3.1: The probability of the states and the premiums, indemnities and the insured's wealths in different states under the crop yield insurance

State	**Probability**	**Premium**	**Indemnity**	**Wealth**
Risk with low yield	$\pi(e)q$	$P(\alpha)$	$I(\alpha) = (\alpha y^* - y_l)$	$W_l = W_0 + y_l + (\alpha y^* - y_l) - P(\alpha)$
Risk with high yield	$\pi(e)(1-q)$	$P(\alpha)$	0	$W_h = W_0 + y_h - P(\alpha)$
No risk	$1 - \pi(e)$	$P(\alpha)$	0	$W_h = W_0 + y_h - P(\alpha)$

Table 3.1 represents the payoffs of the farmer under the crop yield insurance defined as follows: Let we consider that the risk occurs with the probability $\pi(e)$. If the realized yield y is lower than the strike level yield αy^* (so-called low yield), the insured will take the payment from the insurer. On the other hand, if y is higher than αy^* under "Risk with high yield" state, the insured will not take any payment. Lastly, if there is no loss with the probability $(1 - \pi(e))$, the insured will not be paid, either. The insured's yield becomes y_l for the low yield case with probability q, and y_h for the high yield with probability $1 - q$. We do not take the loading factor $\beta(\alpha)$ and the premium subsidy rate $s(\alpha)$ as functions since we do not consider the gross premium case.

3.3.1. Optimal Contract with observable effort

As seen in Figure 1.1, the farmer either rejects or accepts the contract under specific conditions. If the expected utility of the farmer is higher than his/her reservation utiliy U_0 (Here, $U_0 = [1 - \pi(e)]U(W_0 + y_h) + \pi(e)U(W_0 + y_l)$ is the utility when there is no crop yield insurance), he/she will accept the contract, i.e. $E(U_I) > U_0$ (also known the participation constraint where U_1 is the utility under crop yield insurance) in the Principal-Agent Problem. After accepting the contract, then farmer chooses what effort level to make. Farmer might exert either low or high effort, which are associated with different costs ($0 \leq c(e_{low}) < c(e_{high})$).

Here, we assume that the effort level is observable by the insurer, that is the insurer has full information about the farmer's activities, then the premium $P(\alpha, e)$ paid by the insured can be described as a function of the farmer's effort level e and the coverage level α. Thus, according to the three states given in Table 3.1, we describe the farmer's the expected utility by the following equation:

$$E(U_I) = [1 - \pi(e)]U(W_h) + \pi(e)(1 - q)U(W_h) + \pi(e)qU(W_l) - c(e) \tag{3.6}$$

where U_I is the utility function of the insured, $W_l = W_0 + y_l + (\alpha y^* - y_l) - P(\alpha, e)$ and $W_h = W_0 + y_h - P(\alpha, e)$ are the wealths of the insured for low and high yield, respectively. The first term on the right hand side of Equation (3.6) represents the farmer's expected utility for "No risk" state, the second term is the farmer's expected utility for "Risk with high yield" case and the third term represents the insured's expected utility under risk and low yield case.

Equation (3.6) can be rewritten as:

$$E(U_I) = [1 - \pi(e)q]U(W_h) + \pi(e)qU(W_l) - c(e) \tag{3.7}$$

we could define the expected utility of the farmer under two cases as shown in Equation (3.7) which are (i) the indemnity payment is made with probability $\pi(e)q$ and is not made with probability $1 - \pi(e)q$.

Based on the expected utility of the insured's expressed by Equation (3.7), the insured chooses optimal effort e by maximizing the expected utility of his/her final wealth:

$$\max E(U_I) = [1 - \pi(e)q]U(W_h) + \pi(e)qU(W_l) - c(e) \tag{3.8}$$

which is subject to the zero profit condition of the insurer given as:

$$P(\alpha, e) - \pi(e)qI(\alpha) = 0 \tag{3.9}$$

where $I(\alpha) = (\alpha y^* - y_l)$. Here, we assume that the insurer has zero profit and the insurance market is competitive. Thus, the expected profit of the insurer is zero in equilibrium.

The derivative of Equation (3.8) under zero profit condition with respect to e gives the optimal effort level e:

$$\frac{\partial E(U_I)}{\partial e} = \pi'(e)q[U(W_l) - U(W_h)] - P'_e(\alpha, e)([1 - \pi(e)q]U'(W_h) + \pi(e)qU'(W_l)) - c'(e) \tag{3.10}$$

and it is computed as, $P'_e(\alpha, e) = \pi'(e)qI(\alpha)$, $P'_e(\alpha, e)$ denotes the partial derivatives of the premium with respect to e. The optimal level of the effort, indicated as e, must verify the first order condition, i.e. $\frac{\partial E(U_I)}{\partial e} = 0$. Therefore, Equation (3.10) can be rewritten as :

$$\pi'(e)q[U(W_l) - U(W_h)] - P'_e(\alpha, e)([1 - \pi(e)q]U'(W_h) + \pi(e)qU'(W_l)) = c'(e) \tag{3.11}$$

which implies that the expected marginal benefit of the effort (the left hand side of Equation (3.11) equals to the marginal cost of the effort (the right side of Equation (3.11)). The first term on the left hand side captures the effect of a reduction of the probability of risk occuring $\pi(e)$. The marginal benefit consists of the second term shows the marginal benefit in premium reduction because the insurance premium decreases in effort e. Two sides of Equation (3.11) are positive because $\pi'(e) < 0, P'_e(\alpha, e) = \pi'(e)qI(\alpha) < 0$ (since $q > 0$ and $I(\alpha) > 0$) and $c'(e) > 0$.

Taking the derivative of Equation (3.10) with respect to e, the second order optimality condition is satisfied as given:

$$\begin{aligned}\frac{\partial^2 E(U_I)}{\partial e^2} = \; & \pi''(e)q[U(W_l) - U(W_h)] \\ & -\pi'(e)qP_e'(\alpha,e)[U'(W_l) - U'(W_h)] \\ & -P_e''(\alpha,e)([1-\pi(e)q]U'(W_h) + \pi(e)qU'(W_l)) \\ & -P_e'(\alpha,e)[\pi'(e)q(U'(W_l) - U'(W_h)] \\ & +(P_e'(\alpha,e))^2([1-\pi(e)q]U''(W_h) + \pi(e)qU''(W_l)) \\ & -c''(e)\end{aligned} \tag{3.12}$$

with $W_l = W_0 + y_l + (\alpha y^* - y_l) - P(\alpha, e)$ and $W_h = W_0 + y_h - P(\alpha, e)$. In order to determine the sign of Equation (3.12), we define the following expressions for simplicity:

$a = \pi''(e)q[U(W_l) - U(W_h)]$

$b = -\pi'(e)qP_e'(\alpha,e)[U'(W_l) - U'(W_h)]$

$c = -P_e''(\alpha,e)([1-\pi(e)q]U'(W_h) + \pi(e)qU'(W_l))$

$d = -P_e'(\alpha,e)[\pi'(e)q(U'(W_l) - U'(W_h)]$

$f = (P_e'(\alpha,e))^2([1-\pi(e)q]U''(W_h) + \pi(e)qU''(W_l))$

We are able to decide the sign of a, b, c, d, f under the following assumptions:

1. $U^{'}(W) > 0, U^{''}(W) < 0$.
2. $U(W_l) - U(W_h) < 0$, because of the concavity of U, ($W_h > W_l$).
3. $U'(W_l) - U'(W_h) > 0$, this assumption can also be seen in Figure 3.1.

4. $\pi'(e) < 0$ and $\pi''(e) > 0$.

5. $P'_e(\alpha, e) < 0$ and $P''_e(\alpha, e) > 0$.

6. $c''(e) > 0$

Based on the assumptions mentioned above, we can rewrite Equation (3.12) as:

$$\frac{\partial^2 E(U_I)}{\partial e^2} = \underbrace{a}_{<0} + \underbrace{b}_{<0} + \underbrace{c}_{<0} + \underbrace{d}_{<0} + \underbrace{f}_{<0} - \underbrace{c''(e)}_{>0} < 0 \tag{3.13}$$

As seen in Equation (3.13), the second-order condition is negative at the optimal level e, i.e. $\frac{\partial^2 E(U_I)}{\partial e^2} < 0$. Hence, the effort level e is the optimal solution for the crop yield insurance under perfect information.

3.3.2. Optimal Contract with non-observable effort

In the previous section, we discuss that the farmer must obtain at least that expected utility U_0 to make agreement with insurer. We also define the premium as a fuction of the effort level e. However, in case of moral hazard, the farmer has personal information on her/his choice of crop insurance. Thus, the insurer cannot observe the farmer's effort e, which is the reason that moral hazard problem exists. In that case the insurance premium cannot be described as a function of the effort. It only depends on the coverage level α, i.e. $P(\alpha)$, not $P(\alpha, e)$. Under the non-observable effort case, the maximization problem turns out to be as follows:

$$\max E(U_I) = [1 - \pi(e)q]U(W_0 + y_h - P(\alpha)) + \pi(e)qU(W_0 + y_l + (\alpha y^* - y_l) - P(\alpha)) - c(e) \tag{3.14}$$

which is subject to the zero profit condition of the insurer:

$$P(\alpha) - \pi(e)qI(\alpha) = 0 \tag{3.15}$$

Taking the derivative of Equation (3.14) with respect to e and setting it equal to zero, i.e. $\frac{\partial E(U_I)}{\partial e} = 0$, we obtain the first order condition as follows.

$$\pi'(e)q[U(W_l) - U(W_h)] = c'(e) \tag{3.16}$$

with $W_l = W_0 + y_l + (\alpha y^* - y_l) - P(\alpha)$ and $W_h = W_0 + y_h - P(\alpha)$. The second order optimality condition is satisfied as given:

$$\frac{\partial^2 E(U_I)}{\partial e^2} = \pi''(e)q[U(W_l) - U(W_h)] - c''(e) < 0 \tag{3.17}$$

Clearly, the term $U(W_l) - U(W_h)$ is negative because of the concavity of U, and $\pi''(e) > 0, c''(e) > 0$. Thus, $sgn[\frac{\partial^2 E(U_I)}{\partial e^2}]$ is negative.

In order to compare the optimal contract results with observable and non-observable efforts, we examine Equations (3.11) and (3.16). Marginal benefit in premium reduction in Equation (3.11) is removed for non-observable effort case. This is a conclusion of moral hazard and leads to an inefficient price in crop insurance.

In our study, we assume that the insurer cannot observe the insured's effort e. In order to observe the impact of the coverage level α on effort level e, we need to take the total differential of Equation (3.16) with respect to e and α by using *Implicit Function Theorem* :

$$\frac{\partial e}{\partial \alpha} = -\frac{\frac{\partial^2 E(U_I)}{\partial e \partial \alpha}}{\frac{\partial^2 E(U_I)}{\partial e^2}} \tag{3.18}$$

Taking the derivative of Equation (3.16) with respect to α with zero profit condition $P(\alpha) = \pi(e)qI(\alpha)$ then we substitue $P(\alpha)$ into Equation (3.16) for W_l and W_h, we have:

$$\frac{\partial^2 E(U_I)}{\partial e \partial \alpha} = \pi'(e)q[y^*[(1 - \pi(e)q)U'(W_l) + \pi(e)qU'(W_h)]] \tag{3.19}$$

where $P'(\alpha) = \pi(e)qy^* > 0$, and because of the concavity of U and $\pi'(e) < 0$ implies that the sign of Equation (3.19) is negative, i.e. $[\frac{\partial^2 E(U_I)}{\partial e^2 \partial \alpha}] < 0$. Hence, the sign of Equation (3.18) is negative because the signs of the terms $[\frac{\partial^2 E(U_I)}{\partial e \partial \alpha}]$ and $[\frac{\partial^2 E(U_I)}{\partial e^2}]$ are negative in Equations (3.17) and (3.19) . We have:

$$\frac{\partial e}{\partial \alpha} = -\frac{\frac{\partial^2 E(U_I)}{\partial e \partial \alpha}}{\frac{\partial^2 E(U_I)}{\partial e^2}} < 0 \tag{3.20}$$

According to Equation (3.20), the optimal effort e exercised by the farmer, is a strictly decreasing function of the coverage level α. For a farmer in low wealth state, the higher coverage level α provides more marginal utility than the one under high wealth state, i.e., $U'(W_l) > U'(W_h)$ as a result of *"The Law of Diminishing Marginal Utility"*. The higher coverage level α causes that the farmer will make less effort to be in a high wealth state, and the farmers will thus not try to make more effort to increase their wealths. The explanation of this inference can also be seen in Figure 3.1.

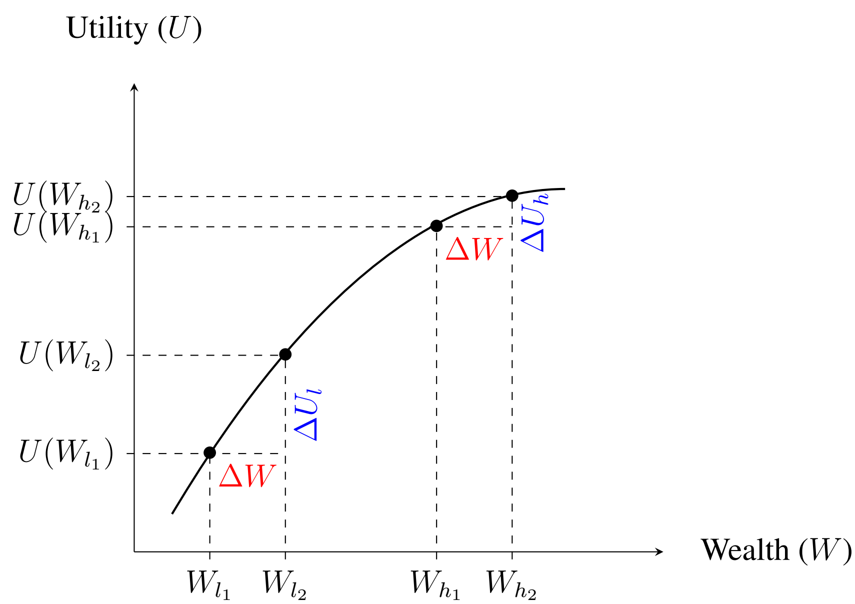

Figure 3.1: Comparisons of the wealth states using utility function U (strictly increasing, concave)

As seen in Figure 3.1, an increase in wealth W gives less additional utility for higher wealth amounts. If the same increase amount ΔW is made for two wealth states (W_l and W_h), the change in the marginal utility in the low wealth state is higher than the one under the high wealth state, i.e. $\Delta U_h < \Delta U_l$.

3.4. Optimal Effort Level under Loss-Reduction

In the previous section, we examined the moral hazard effect under loss prevention perspective. In that case, the farmer is assumed to reduce the probability of risk occuring to put more effort. In this section, we now suppose that the farmer can affect the size of the crop yield losses. Therefore, the crop yield $y(e)$ can be described as a function of the farmer's effort level. Here, we assume that $y(e)$ is strictly increasing, concave and twice differentiable on e, i.e, $y'(e) > 0$ and $y''(e) < 0$. $y'(e)$ shows the marginal yield arising from the effort level e. At this point, we create the variable $\theta(e)$ to see the effect of e on the crop yield. $\theta(e)$ shows the rate of increase in the crop yield that lies between 0 to 1 ($0 < \theta(e) < 1$). $y(e)$ and $\theta(e)$ have the same mathematical properties i.e. $\theta'(e) > 0$ and $\theta''(e) < 0$. That is, the crop yield level increases as effort e increases, i.e. the crop yield level is the increasing function of e, however the increase at the yield level is decreasing. Hence, we will define $\theta(e)$ separately for y_l and y_h. $\theta_l(e)$ and $\theta_h(e)$ denote the rate of increase in low and high yield, respectively. Also, we assume that $\theta'_l(e) > \theta'_h(e)$ because of the concavity of $\theta(e)$.

We model the indemnity payment as in the previous section. If the yield falls below a strike level yield αy^*, then the indemnity payment $I(\alpha)$ is made by the insurer, otherwise no payment is made. In order to define the payoffs and the states, we use the same methodology as in Table 3.1. So, the payoffs to the insured under crop yield insurance are given separately in Table 3.2:

Table 3.2: The payoffs to the insured under crop yield insurance in case of loss-reduction

State	**Probability**	**Premium**	**Indemnity**	**Wealth**
Risk with low yield	$\pi q(e)$	$P(\alpha)$	$I(\alpha) = (\alpha y^* - y_l)$	$W_l = W_0 + y_l(1 + \theta_l(e)) + (\alpha y^* - y_l) - P(\alpha)$
Risk with high yield	$\pi(1 - q(e))$	$P(\alpha)$	0	$W_h = W_0 + y_h(1 + \theta_h(e)) - P(\alpha)$
No risk	$1 - \pi$	$P(\alpha)$	0	$W_h = W_0 + y_h(1 + \theta_h(e)) - P(\alpha)$

$q(e)$ denotes the probability of low yield case since the farmer is assumed to affect the distribution of crop yield in case of loss-reduction. Here, we assume that $q(e)$ is strictly decreasing and convex, i.e. $q'(e) < 0, q''(e) > 0$ since more effort e provides higher crop yield and decreases the probability of the low yield case. On the other hand, π represents the probability of risk event occuring such as drought, hail and does not depend on the effort. W_l and W_h denote end-of-period wealth below and above the strike level yield, respectively. $y_l(1+\theta_l(e))$ and $y_h(1+\theta_h(e))$ represent the crop yield for low and high yield, respectively. In this section, we suggest a new insurance policy. If the yield falls below a strike level yield αy^*, the farmer only receive the indemnity payment $I(\alpha) = (\alpha y^* - y_l)$ in classical model. Here, we also provide an additional benefit $y_l\theta_l(e)$ to the farmer from his/her effort as well as the indemnity payment in case of low yield.

3.4.1. Optimal contract with observable effort

In this section, we assume that the insurer is able to observe the farmer's effort e. So, the crop insurance premium $P(\alpha, e)$ depends on the effort level e as well as the coverage level α. The zero profit condition of the insurer is defined as $P(\alpha, e) = \pi q(e) I(\alpha)$. Using these results, according to the three states mentioned in Table 3.2, we describe the farmer's expected utility by the following equation:

$$E(U_I) = (1-\pi)U(W_h) + \pi[1-q(e)]U(W_h) + \pi q(e)U(W_l) - c(e) \tag{3.21}$$

with $W_l = W_0 + y_l(1+\theta_l(e)) + (\alpha y^* - y_l) - P(\alpha, e)$ and $W_h = W_0 + y_h(1+\theta_h(e)) - P(\alpha, e)$. We can rewrite Equation (3.21) after some arrangement:

$$E(U_I) = [1-\pi q(e)]U(W_h) + \pi q(e)U(W_l) - c(e) \tag{3.22}$$

The expected utility of the farmer is defined under two cases as shown in Equation (3.22). The indemnity payment $I(\alpha, e)$ is made by the insurer with probability $\pi q(e)$ and otherwise no payment is made with probability $1-\pi q(e)$. Then, the insured chooses optimal effort e in order to maximize their expected utility.

$$\max E(U_I) = [1-\pi q(e)]U(W_h) + \pi q(e)U(W_l) - c(e) \tag{3.23}$$

The derivative of Equation (3.23) under zero profit condition with respect to e gives the optimal effort level e:

$$\frac{\partial E(U_I)}{\partial e} = \pi q'(e)[U(W_l) - U(W_h)] + [1 - \pi q(e)]y_h\theta_h'(e)U'(W_h) + [\pi q(e)]y_l\theta_l'(e)U'(W_l)$$
$$-P_e'(\alpha, e)([1 - \pi q(e)]U'(W_h) + \pi q(e)U'(W_l)) - c'(e) \tag{3.24}$$

The optimal level of the effort must verify the first order condition, i.e. $\frac{\partial E(U_I)}{\partial e} = 0$. Therefore, Equation (3.24) can be rewritten as :

$$\underbrace{q'(e)\pi[U(W_l) - U(W_h)]}_{(i)} + \underbrace{[1 - \pi q(e)]y_h\theta_h'(e)U'(W_h) + [\pi q(e)]y_l\theta_l'(e)U'(W_l)}_{(ii)}$$
$$\underbrace{-P_e'(\alpha, e)([1 - \pi q(e)]U'(W_h) + \pi q(e)U'(W_l))}_{(iii)} = c'(e) \tag{3.25}$$

The straightforward interpretation of Equation (3.25) is that marginal benefit (the left hand side of this equation) equals the marginal cost (the right hand side of this equation). The term *(i)* shows that the effort level e decreases the probability of the low yield case and thus it provides the marginal utility benefit from increasing the probability of the high yield state i.e. $q'(e)\pi[U(W_l) - U(W_h)] > 0$. The second term *(ii)* is also positive and represents an increase in the effort level e increases the wealth for low and high level yield cases. In the remaining term *(iii)* in Equation (3.25), the effort e increases the wealth because $P(\alpha, e)$ decreases with an increase in effort. Also, the second-order condition is negative at the optimal level e, i.e. $\frac{\partial^2 E(U_I)}{\partial e^2} < 0$. The details are given in Equation (3.26).

Taking the derivative of Equation (3.24) with respect to e, the second order optimality condition is satisfied as given:

$$
\begin{aligned}
\frac{\partial^2 E(U_I)}{\partial e^2} = \; & \pi q''(e)[U(W_l) - U(W_h)] \\
& +\pi q'(e)[y_l\theta_l'(e)U'(W_l) - y_h\theta_h'(e)U'(W_h)] \\
& -\pi q'(e)P_e'(\alpha,e)[y_l\theta_l'(e)U'(W_l) - y_h\theta_h'(e)U'(W_h)] \\
& -\pi q'(e)[y_h\theta_h'(e)U'(W_h) - y_l\theta_l'(e)U'(W_l)] \\
& +(1-\pi q(e))y_h\theta_h''(e)U'(W_h) + \pi q(e)y_l\theta_l''(e)U'(W_l) \\
& +(1-\pi q(e))y_h^2\theta_h'(e)^2U''(W_h) - (1-\pi q(e))y_h\theta_h'(e)P_e'(\alpha,e)U''(W_h) \\
& +(1-\pi q(e))y_l^2\theta_l'(e)^2U''(W_l) - (1-\pi q(e))y_l\theta_l'(e)P_e'(\alpha,e)U''(W_l) \\
& -P_e''(\alpha,e)[(1-\pi q(e))U'(W_h) + \pi q(e)U'(W_l)] \\
& +\pi q'(e)P_e'(\alpha,e)[U'(W_h) - U'(W_l)] \\
& -P_e'(\alpha,e)(1-\pi q(e))y_h\theta_h'(e)U''(W_h) \\
& +P_e'(\alpha,e)^2(1-\pi q(e))U''(W_h) \\
& -P_e'(\alpha,e)\pi q(e)y_l\theta_l'(e)U''(W_l) \\
& +P_e'(\alpha,e)^2\pi q(e)U''(W_l) - c''(e)
\end{aligned} \tag{3.26}
$$

Here, we assume that the term $[y_l\theta_l'(e)U'(W_l) - y_h\theta_h'(e)U'(W_h)]$ is positive since for a farmer in low wealth state, the same effort level e provides more marginal utility than the one in high wealth state. Thus, $sgn[\frac{\partial^2 E(U_I)}{\partial e^2}]$ is negative. Hence, the effort level e is the optimal solution for the crop yield insurance under perfect information.

3.4.2. Optimal contract with non-observable effort

We mentioned in the Section 3.3.2 that the insurer cannot observe the farmer's effort e in case of moral hazard. Hence, the premium cannot be described as a function of the effort and only depends on the coverage level α. The farmer then chooses e to maximize his/her expected utility

$$\begin{aligned} \max E(U_I) = \ & [1 - \pi q(e)]U(W_0 + y_h(1 + \theta_h(e)) - P(\alpha)) \\ & + \pi q(e)U(W_0 + y_l(1 + \theta_l(e)) + (\alpha y^* - y_l) - P(\alpha)) - c(e) \end{aligned} \tag{3.27}$$

which is subject to the zero profit condition of the insurer. The first order condition is given as:

$$\frac{\partial E(U_I)}{\partial e} = \pi q'(e)[U(W_l) - U(W_h)] + [1 - \pi q(e)]y_h\theta_h'(e)U'(W_h) + [\pi q(e)]y_l\theta_l'(e)U'(W_l) = c'(e) \tag{3.28}$$

The marginal benefit from the premium is removed from above equation as in Equation (3.16). As we mentioned before, this result indicates the existence of moral hazard problem.

According to the maximization problem, the second order optimality condition is fulfilled as given:

$$\begin{aligned} \frac{\partial^2 E(U_I)}{\partial e^2} = \ & \pi(e)q''(e)[U(W_l) - U(W_h)] + \pi q'(e)[y_l\theta_l'(e)U'(W_l) - y_h\theta_h'(e)U'(W_h)] \\ & - \pi q'(e)[y_h\theta_h'(e)U'(W_h) - y_l\theta_l'(e)U'(W_l)] \\ & + [1 - \pi q(e)]y_h\theta_h''(e)U'(W_h) + [\pi q(e)]y_l\theta_l''(e)U'(W_l) \\ & + [1 - \pi q(e)]y_h^2\theta_h'(e)^2U''(W_h) + [\pi q(e)]y_l^2\theta_l'(e)^2U''(W_l) - c''(e) \end{aligned} \tag{3.29}$$

Here, we assume that the term $[y_l\theta_l'(e)U'(W_l) - y_h\theta_h'(e)U'(W_h)]$ is positive since for a farmer in low wealth state, the same effort level e provides more marginal utility than the one in high wealth state. Thus, $sgn[\frac{\partial^2 E(U_I)}{\partial e^2}]$ is negative.

In order to examine the impact of the coverage level α on effort level e, we need to take the total differential of Equation (3.28) with respect to e and α:

$$\frac{\partial e}{\partial \alpha} = -\frac{\frac{\partial^2 E(U_I)}{\partial e \partial \alpha}}{\frac{\partial^2 E(U_I)}{\partial e^2}} \tag{3.30}$$

Given $\frac{\partial^2 E(U_I)}{\partial e^2} < 0$, the sign of $\dfrac{\partial e}{\partial \alpha}$ depends on $\frac{\partial^2 E(U_I)}{\partial e \partial \alpha}$.

$$\begin{aligned}\frac{\partial e}{\partial \alpha} = \; & \underbrace{\pi q'(e)[y^*[(1-\pi q(e))U'(W_l) + \pi q(e)U'(W_h)]]}_{a} \\ & \underbrace{- y^*(1-\pi q(e))\pi q(e)[\underbrace{y_h\theta_h'(e)U''(W_h)}_{i} - \underbrace{y_l\theta_l'(e)U''(W_l)}_{ii}]}_{b}\end{aligned} \tag{3.31}$$

Here, the sign of Equation (3.31) is ambiguous even if the term a is negative. Hence, we cannot determine whether a change in the coverage level α has a direct impact on the optimal effort level e, or not. Each possible case which changes the sign of $\dfrac{\partial e}{\partial \alpha}$ should be investigated separately.

3.5. Analyzing Loss-Prevention and Loss-Reduction Models under The Certainty Equivalent Approach

In this section, we will give some numerical examples to evaluate the effects of loss prevention and loss reduction model. The most used method to assess farmer preferences is the EUT. The possible losses that may occur under this theorem is needed to be measured within the frame of the utility. Here, we use the CE approach to analyze the behaviour of the insured against the risk. Unlike EUT, the CE turns possible outcomes into a monetary amount. The CE is the guaranteed amount of money that produces the same utility to the uncertain outcome of a possibly higher amount of wealth. The CE is obtained by finding the inverse of the utility function, i.e. $CE = U^{-1}(E(U))$. The expected utility and CE for a risk-averse insured is presented below.

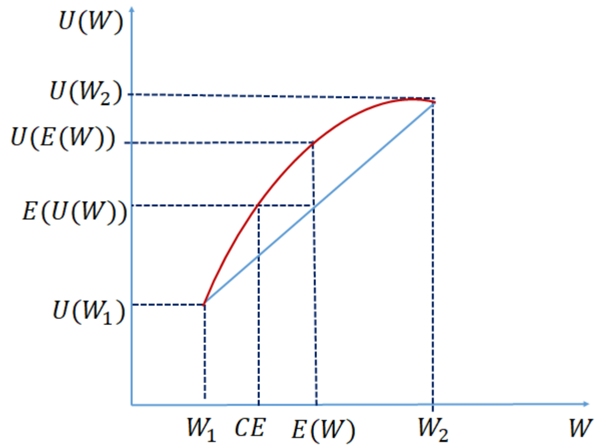

Figure 3.2: Graphical illustration of certainty equivalent and expected utility

As seen in Figure 3.2, a farmer with concave utility function (risk-averse) prefers the expected wealth to the random wealth, i.e. $U(E(W)) \geq E(U(W))$. Figure 3.2 above presents, if the insured is risk-averse, the value of CE is smaller than the average value of wealth. The difference between $E(W)$ and CE is also known as "the risk premium". This premium shows that the amount of money that an insured will want to pay to avoid the interested risk.

We use the mean-variance approach to derive the CE to show the numerical illustrations. Several studies investigate the CE to model the crop insurance by using mean and variance of wealth [13, 21, 43].

We use the following equation to calculate the CE of the insured.

$$CE(W) = E(W) - \frac{r}{2}Var(W) - c(e) \tag{3.32}$$

where $r \geq 0$ denotes the risk aversion coefficient of the insured. $E(W)$ and $Var(W)$ denote expected wealth and the variance of wealth, respectively [44, 45].

3.5.1. Analyzing loss-prevention model under the certainty equivalent approach

The details of loss-prevention model under moral hazard are given in the Section 3.3. To derive the CE for this model, we firstly obtain the expected value and the variance of the wealth. We describe the expected value of wealth $E(W)$ by the following equation.

$$E(W) = [1 - \pi(e)q]W_h + \pi(e)qW_l \tag{3.33}$$

where $W_l = W_0 + y_l + (\alpha y^* - y_l) - P(\alpha, e)$, $W_h = W_0 + y_h - P(\alpha, e)$ and $P(\alpha, e) = \pi(e)q\,(\alpha y^* - y_l)$. We rewrite Equation (3.33) as follows.

$$E(W) = [1 - \pi(e)q]y_h + \pi(e)qy_l \tag{3.34}$$

The variance of wealth is given below.

$$\begin{aligned} Var(W) = &\ [1 - \pi(e)q]\left(W_h - E(W)\right)^2 + \pi(e)q\left(W_l - E(W)\right)^2 \\ = &\ \pi(e)q[1 - \pi(e)q]\left(y_h - \alpha y^*\right)^2 \end{aligned} \tag{3.35}$$

In the case that the farmer, who is expected to produce low yield, has an insurance policy, αy^* is the upper level of the indemnity that the farmer could take from the insurer. If the farmer produce high yield, the difference $(y_h - \alpha y^*)$ denotes the indemnity amount that the farmer could not take from the insurance policy. Therefore, the variance of this term represents the loss. Thus, the CE can be written as follows.

$$CE(W) = W_0 + [1 - \pi(e)q]y_h + \pi(e)qy_l - \frac{r}{2}\pi(e)q[1 - \pi(e)q]\left(y_h - \alpha y^*\right)^2 - c(e) \tag{3.36}$$

A numerical example for the model under the CE approach is given below. We first analyze the relationship between the coverage rate α and the CE under different crop yield levels. The parameters of the CE are given in Table 3.3.

Table 3.3: The parameter values in the certainty equivalent

W_0	y_l	y^*	e	$\pi(e)$	q	r
10,000	200	270	0.37	0.4	0.25	2

We assume that $\pi(e) = (1-e)^2 \in (0,1)$ for $e \in (0,1)$ and $c(e) = e^2$.

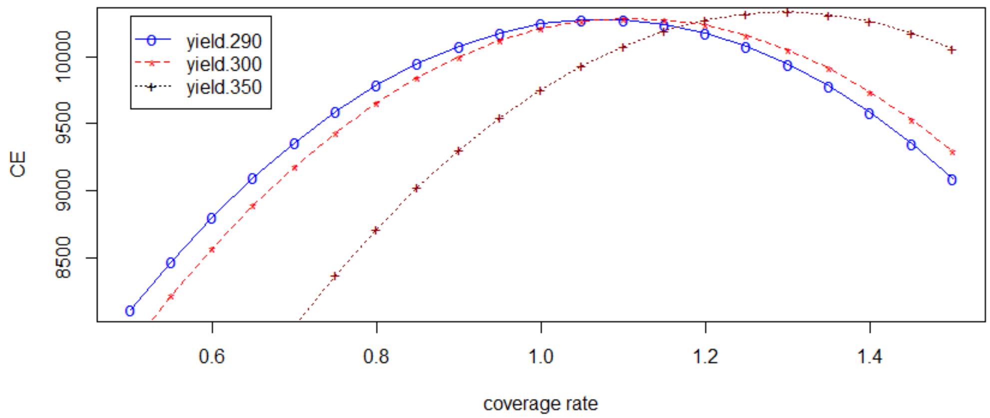

Figure 3.3: The coverage effects under different yield levels in the certainty equivalent

Figure 3.3 presents the CE values for each coverage level α for three crop yield levels. In this example, three yield amounts, 290, 300 and 350, are chosen to determine the relationship between the yield level and the coverage rate. Here, it should be noted that the chosen values are above the strike yield level αy^*. In this example, we assume that the insured produce high yield y_h ($y \geq \alpha y^*$). The optimal coverage rates (α^*) and the related CE values for the chosen yield levels are given in the following table.

Table 3.4: The optimal coverage levels for the selected yield levels

	$y_h = 290$	$y_h = 300$	$y_h = 350$
$\boldsymbol{\alpha^*}$	1.05	1.10	1.30
CE	10,275.10	10,287.09	10,332.81

As seen in Figure 3.3, this optimal coverage rates maximize the CE value. The CE values for each yield level decreases above the optimal coverage rate points ($1.05, 1.10$ and 1.30 respectively). The yield $y_h = 290$ has the highest CE value with the smallest coverage rate. An important result that can be inferred here is that the needed coverage rate is higher as the difference $(y_h - \alpha y^*)$ increases. Since the optimal coverage rate is lower, the premium amount is also lower at the point $\alpha^* = 1.05$. In addition, since the indemnity amount that cannot be taken from the insurer $(y_h - \alpha y^*)$ is lower for $y_h = 290$, the variance of wealth is also lower for that point (see Equation (3.35)). Therefore, the policy under loss-prevention model is more attractive to the farmer having $y_h = 290$ which also have the highest CE value calculated from Equation (3.36). We extend this example by using different risk aversion coefficients, $r = 1, r = 2$ and $r = 4$. The chosen parameters of the CE are given in Table 3.5.

Table 3.5: The parameter values in the certainty equivalent

W_0	y_l	y_h	y^*	e	$\pi(e)$	q
10,000	200	300	270	0.37	0.4	0.25

The relationship between the risk aversion coefficients and the CE is represented in Figure 3.4.

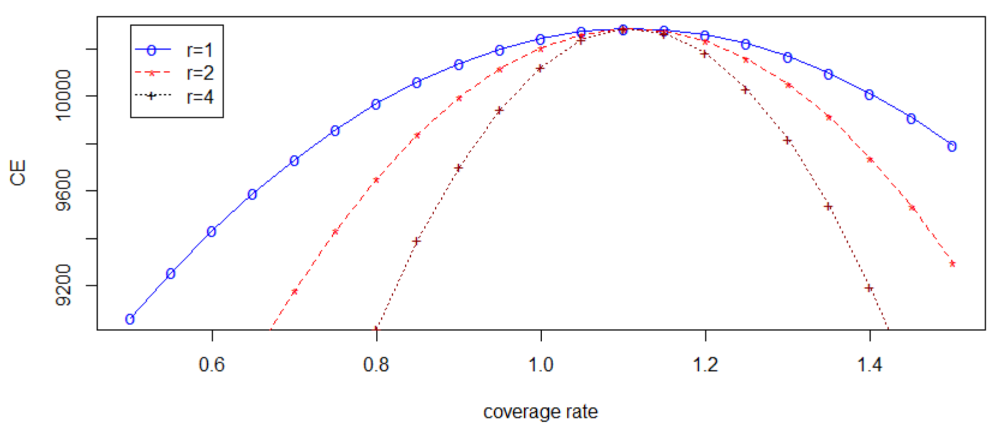

Figure 3.4: The CE under different the risk aversion coefficients

All risk aversion coefficients reach the highest CE value at the coverage rate $\alpha = 1.10$. The smallest the risk aversion coefficient $r = 1$ has the highest CE for all coverage rates. Hence, individuals having low risk aversion coefficient might prefer the policy provided under loss-prevention model.

3.5.2. Analyzing loss-reduction model under the certainty equivalent approach

The expected value of wealth $E(W)$ for loss-reduction model is given as follows:

$$E(W) = [1 - \pi q(e)]W_h + \pi q(e)W_l \tag{3.37}$$

where $W_l = W_0 + y_l(1 + \theta_l(e)) + (\alpha y^* - y_l) - P(\alpha, e)$, $W_h = W_0 + y_h(1 + \theta_h(e)) - P(\alpha, e)$ and $P(\alpha, e) = \pi(e)q\,(\alpha y^* - y_l)$. We can rewrite Equation (3.37) as follows.

$$E(W) = W_0 + [1 - \pi q(e)]y_h(1 + \theta_h(e)) + \pi q(e)y_l(1 + \theta_l(e)) \tag{3.38}$$

The variance of wealth is given below.

$$\begin{aligned} Var(W) = \;& [1 - \pi q(e)]\left(W_h - E(W)\right)^2 + \pi q(e)\left(W_l - E(W)\right)^2 \\ = \;& \pi q(e)[1 - \pi q(e)]\left(y_h(1 + \theta_h(e)) - [\alpha y^* + y_l\theta_l(e)]\right)^2 \end{aligned} \tag{3.39}$$

Thus, the CE can be written as follows.

$$\begin{aligned} CE(W) = \;& W_0 + [1 - \pi q(e)]y_h(1 + \theta_h(e)) + \pi q(e)y_l(1 + \theta_l(e)) \\ & - \frac{r}{2}\pi q(e)[1 - \pi q(e)]\left(y_h(1 + \theta_h(e)) - [\alpha y^* + y_l\theta_l(e)]\right)^2 - c(e) \end{aligned} \tag{3.40}$$

In this section, we will only examine how the CE values change according to the increasing effort levels for numerical illustration. The CE values are obtained for two models mentioned in Equation (3.36) and Equation (3.40). Another advantage of the CE approach is that it evaluates the effectiveness of the models [46–48].

Table 3.6: The parameter values used in CE calculation

	W_0	y_l	y_h	α	y^*	π	r
Loss-Prevention	10,000	200	300	0.85	270	0.25	2
Loss-Reduction	10,000	200	300	0.85	270	0.25	2

We assume that $q(e) = (1 - e)^2 \in (0, 1)$ for $e \in (0, 1)$, $\theta_l(e) = 0.5(2e - e^2)$; $\theta_h(e) =$

$0.25(2e - e^2)$ and $c(e) = e^2$.

According to the results presented in Table 3.7, the CE values increase as effort level inreases. The CE value in the loss-reduction model is higher than the one in the loss-prevention model for each case. Hence, the loss-reduction model might be preferred rather than the loss-prevention model.

Table 3.7: The CE values for different effort levels

	Loss-Reduction	**Loss-Prevention**
e	**CE**	**CE**
0.1	9,595.59	9,475.86
0.2	9,802.61	9,614.51
0.3	9,967.24	9,751.66
0.4	10,094.95	9,881.71
0.5	10,191.54	9,999.81
0.6	10,262.43	10,101.82
0.7	10,312.31	10,184.38
0.8	10,344.88	10,244.84
0.9	10,362.79	10,281.31

The results in Table 3.7 are represented in Figure 3.5.

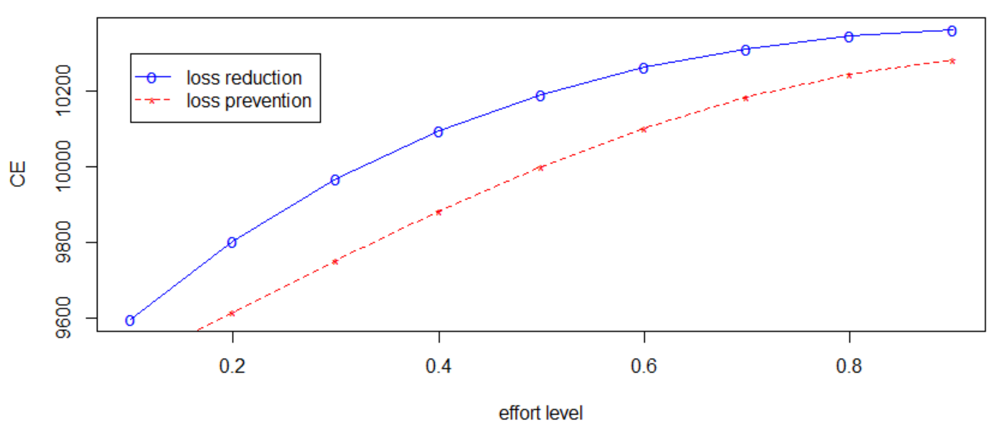

Figure 3.5: The comparison of the CE under loss-reduction model and the loss-prevention model for different effort levels

The difference between the CE values of the loss-reduction model and the loss-prevention model decreases after effort level 0.3. That is, the difference between the CE values of the models decreases as the level of effort increases. The farmer might be indifferent to these two models at high levels of effort.

3.6. Interim Conclusion: Moral Hazard in Yield Insurance

Our study is mainly based on the effects of moral hazard on optimal yield insurance. In this chapter, we consider various cases to provide an understanding about our motivation of taking into account of moral hazard in this thesis.

We firstly introduce a main expected utility formulation from a farmer's perspective. The profit of the farmer is investigated for all cases, which the yield insurance suggests, i.e. the yield's being below/above a determined threshold. In this sense, the expected utility of that farmer could be maximized.

A high level of the farmer's effort results in a lower probability of occurance of the events affecting the crop yield. Therefore, the optimal effort is obtained to determine the impact of the moral hazard for the case where the farmer could prevent the loss, i.e. loss-prevention. In this case, the optimal effort is evaluated under the scenarios where the effort is observable or not. Indeed, the insurer's ability of observing the farmer's effort is the concept of asymmetric information which is another highly-debated topic in the literature.

Moreover, it is also considered in this chapter that the farmer's effort could have an influence on the size of the yield. Hence, the solution of the optimal effort is investigated in order to measure the moral hazard where the farmer could reduce the size of the loss, i.e. loss-reduction. The impact of the asymmetric information is considered for this case as well.

The solutions of the optimal effort under both loss-prevention and loss-reduction cases are the fundamental achievements of our study in terms of yield insurance. The contribution of this chapter is to find the farmer's optimal effort which is a very significant indication of the desired situation, i.e. non-existence of the moral hazard. In addition to this, we propose an assessment of asymmetric information through the optimal contracts considering the difference between observable and non-observable effort.

We also provide the reasons why we could only propose a theoretical solutions in order to evaluate the moral hazard and asymmetric information in this chapter. Since the farmer-based yield data does not exist, the time and cost of collecting insured-based data lead us to use distirct-based yield data. More detailed information for the use of district-based data is presented in Chapter 4.

Lastly, we analyze the efficiency of loss-prevention and loss-reduction models according to different coverage rates, risk aversion coefficients, crop yield levels and effort levels. One of the important results is that the policy under loss-prevention model is more attractive to the farmer who produces high yield close to strike yield level. In addition, we find that the insureds having low risk aversion coefficients might prefer the policy provided under loss-prevention model. Another important result is that the loss-reduction model has higher CE values as the effort level increases. However, the insured might be indifferent to these two models for high effort levels according to the design of the insurance policy (due to the costs such as premium calculation, input costs etc.).

4. BAYESIAN SPATIOTEMPORAL MODELLING OF CROP YIELD DATA

4.1. Introduction

The spatial and temporal effects for the crop yield is investigated in this chapter. We aim to model the distribution of the yield by " the hierarchical Bayesian method" which reflects the spatial and temporal dependencies among geographical areas. For this aim, we deal with district-based yield data and neighbourhoods between specified subregions.

In geographical studies, it is not always possible to have information for each individual. In our study, we could reach district-based crop yields instead of farmer-based crop yields. This drawback leads us to define yields for subregions. This approach provides an approximation based on the "Law of Large Numbers" which could also be handled as an advantage that might decrease the individual impact of moral hazard.

In this chapter, we introduce the hierarchical Bayesian approach that is used to take into account of the spatiotemporal dependency impacts on the conditional distribution of the yield. In order to do this, the farmer-based yield is assumed to approximate to the district-based yield due to the lack of point-based data.

Section 4.2 explains the use of district-based crop yield data by means of neighbourhoods. We introduce traditional Bayesian model in Section 4.3. In Section 4.4, we provide the hierarchical structure that allows us to take into consideration of spatial or temporal dependencies among crop yields. Section 4.5 describes "Integrated Nested Laplace Approximation (INLA)" which performs better than "Markov Chain Monte Carlo (MCMC)" approach in computing time especially for geographical studies. In this section, we also present various measures for the model's prediction performance and the model selection. Finally, we summarize this chapter in Section 4.6.

4.2. District-Based Crop Yield

In this chapter we focus on spatial, temporal and spatiotemporal effects of yield in crop insurance. We assume that the agricultural output has a spatiotemporal process, that is crop yield conditional to the temporal and spatial processes. Hence, the conditional density of crop yield is denoted as $f(y_{s,t}|\Omega_{s,t})$ where $\Omega_{s,t}$ represents any information belong to the location s and the time t related to the realized crop yield [49].

In geographical studies, it is not always possible to have information for each individual. Since the district-based crop yields are available instead of farmer-based crop yields, we define yields for subregions.

In order to examine spatial effect of the crop yield, we define farmer-based crop yield and district-based crop yield as y_s and Y_s, respectively where s is the vector of latitude and longitude of the data point. We assume that farmer-based yield approximates district-based yield which means that $y_s \approx Y_s; s \in \boldsymbol{D} \subset \mathbb{R}^2$ where s belongs to a domain $\boldsymbol{D}$. For spatiotemporal case, the approaximation of the crop yield data turns out to be $y_{s,t} \approx Y_{s,t}; (s,t) \in (\boldsymbol{D}, \boldsymbol{T}) \subset \mathbb{R}^2 \times \mathbb{R}$.

Considering that a spatial region is divided into subregions $(1, \ldots, S)$, spatial data of crop yield Y_s denotes the areal-based data. Y_s is measured over the predetermined region according to this data type. An illustration of this approximation is seen in Figure 4.1.

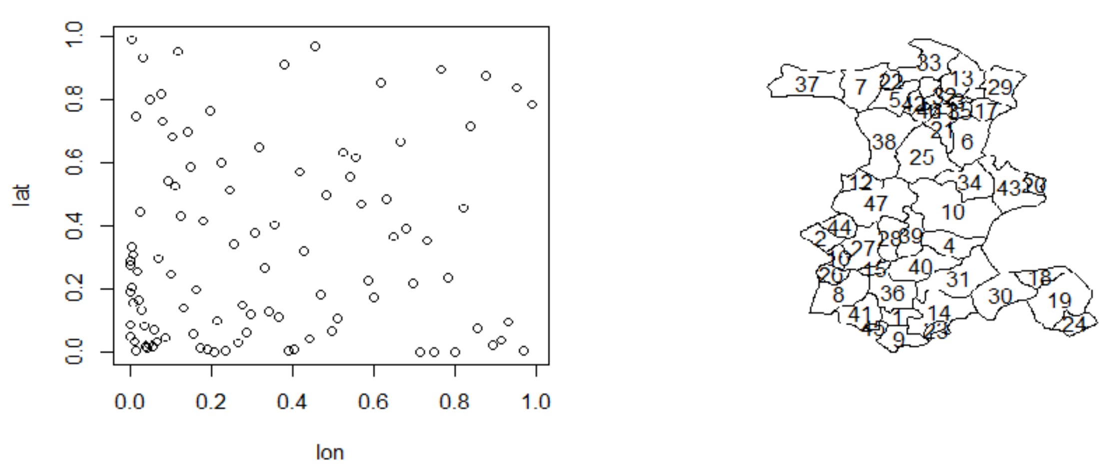

Figure 4.1: Point-based data (left) and areal-based data (right)

In this figure, the latitude-longitude couples as shown points on the left figure could be considered as locations of the farmer-based yields whereas the numbers relating to areas on the right figure represent the subregions of the district-based yields.

The conditional density of crop yield is examined under Bayesian models in our study. In the basic Bayesian approach, the observations are supposed to be independent. However this is not the case for the hazards in the same region since they are spatially dependent. Therefore, we use hierarchical Bayesian models in order to take into account dependence among realized crop yields in a district. In this study, we use district-based data for the application part. In the district-based data, the spatial neighbourhood is used to examine dependence among the specified subregions. The neighbourhoods, which have a common border, between the subregions are shown in below Figure 4.2.

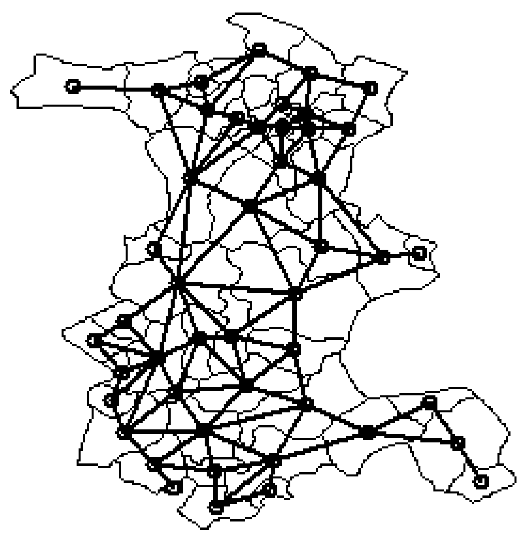

Figure 4.2: Neighbourhood between the predetermined subregions

In Figure 4.2, the lines connecting the points stand for the neighbourhood between subregions.

4.3. Basic Bayesian Modelling

Let $\pi(y|\theta)$ is “the likelihood” for the observed data, which shows the connection between data and parameters of the model. Here, y and θ represent observations from data and random parameters $\theta = (\theta_1, \ldots, \theta_n)$, respectively. The main issue is to obtain “the posterior distribution” of θ, i.e. $\pi(\theta|y)$ when the observed data is available. Given the likelihood, the posterior distribution is defined as

$$\pi(\theta|y) = \frac{\pi(y|\theta)\pi(\theta)}{\pi(y)} \propto \underbrace{\pi(y|\theta)}_{\text{likelihood}}\underbrace{\pi(\theta)}_{\text{prior}}, \tag{4.1}$$

where $\pi(\theta)$ is “the prior distribution” that provides prior information of θ. When θ is a continous random variable then the denominator of Equation (4.1) is defined as

$$\pi(y) = \int_{R_\theta} \pi(y|\theta)\pi(\theta)d\theta$$

where R_θ is the doman of θ. In order to obtain Bayesian estimation of θ, we use the following formula which is known as minumum mean squared error (MMSE)

$$\arg\min_{\hat{\theta}} \int (\theta - \hat{\theta})^2 \pi(\theta|y)d\theta. \tag{4.2}$$

In order to obtain point estimate of θ, we take the derivative of Equation (4.2) with respect to $\hat{\theta}$ as follows

$$\int \frac{\partial}{\partial\hat{\theta}}(\theta - \hat{\theta})^2 \pi(\theta|y)d\theta = \int 2(\theta - \hat{\theta})(-1)\pi(\theta|y)d\theta = 0. \tag{4.3}$$

Having taken the derivation, we set Equation (4.3) equal to 0. Since $\pi(\theta|y)$ is a probability density function, $\int_{R_\theta} \pi(\theta|y)d\theta = 1$ is known. As a result, the point estimation is obtained from Equation (4.3) as follows:

$$\hat{\theta} = \int \theta\pi(\theta|y)d\theta, \tag{4.4}$$

where $\hat{\theta}$ is “the posterior mean” of θ. Also, we aim to make prediction for a non-observed variable. In that case, we use “the posterior predictive distribution” to obtain an estimation of non-observed variable and it is given as

$$\pi(y_{n+1}|y) = \int_\theta \pi(y_{n+1}|\theta)\pi(\theta|y)d\theta. \tag{4.5}$$

In some cases, it is not possible to handle posterior distribution analytically. Hence, MCMC is a simulation-based method used to overcome this problem. MCMC solves this problem by drawing samples from the posterior distribution and computing the basic statistics of the parameter of interest. Gibbs Sampling and Metropolis-Hastings algorithms are the most known MCMC algorithms.

Since the dynamics of the data vary according to time in our data set, the parameters reflecting this variation can be more complex. In such cases, the interested model might have a large number of parameters. Therefore, hierarchical models are used to model parameters according to our data including yearly-recorded crop yields.

4.4. Hierarchical Bayesian Modelling

We could define a model in hierarchical structure if the distribution of one parameter is conditional to another parameter. To explain this clearly, the prior distribution is assigned to another prior parameter which is called hyperparameter. Hence, we can investigate the spatial or temporal dependence between observations by using hierarchical structure. Relations between the parameters can be expressed through “the joint probability distribution” [34]. The Hierarchical Bayesian model can be divided into three levels [50] :

i. Data (Likelihood) level

$$Y|\nu_1, \nu_2 \sim \pi_1(Y|\nu_1, \nu_2) \tag{4.6}$$

ii. Process (Parameter) level

$$\nu_1|\nu_2 \sim \pi_2(\nu_1|\nu_2) \tag{4.7}$$

iii. Prior (Hyperparameter) level

$$\nu_2 \sim \pi_3(\nu_2) \tag{4.8}$$

The first level represents the likelihood function $\pi_1(Y|\nu_1, \nu_2)$ and the data Y which is conditionally independent from given ν_1, ν_2. ν_1 and ν_2 represent parameter and hyperparameter, respectively. By the help of the second level given in Equation (4.7), we could analyze spatial or temporal dependence among the observations in Y. The last equation indicates the prior level with the distribution of hyperparameter π_3. Using these levels, the posterior distribution can be defined through Bayes' theorem as follows:

$$\begin{aligned}\pi(\nu_1, \nu_2|y) &= \frac{\pi(Y, \nu_1, \nu_2)}{\pi(Y)} \\ &= \frac{\pi_1(Y|\nu_1, \nu_2)\pi(\nu_1, \nu_2)}{\int_{R_{\nu_1}}\int_{R_{\nu_2}} \pi_1(Y|\nu_1, \nu_2)\pi(\nu_1, \nu_2)d\nu_1 d\nu_2} \propto \pi_1(Y|\nu_1, \nu_2)\pi(\nu_1, \nu_2)\end{aligned} \tag{4.9}$$

where R_{ν_i}s are the domains of ν_is. Here, the term $\pi(\nu_1, \nu_2)$ is the prior distribution for the data (likelihood) level. Also, $\pi(\nu_1, \nu_2)$ can be written by dividing into two parts as follows:

$$\pi(\nu_1, \nu_2) = \pi_2(\nu_1|\nu_2)\pi_3(\nu_2). \tag{4.10}$$

Therefore, if we rearrange Equation (4.10), then it becomes as

$$\pi(\nu_1, \nu_2|Y) \propto \pi_1(Y|\nu_1, \nu_2)\pi_2(\nu_1|\nu_2)\pi_3(\nu_2). \tag{4.11}$$

As seen in Equation (4.11), the joint posterior which belongs to the model, is proportional to *likelihood, parameter* and *hyperparameter*.

As an example, we consider that the number of car accidents y has a negative binomial distribution where the number of no accidents is k and the probability of no accident is p. Let us define this problem in a hierarchical structure. Firstly, we write *likelihood* level mentioned in Equation (4.6) as follows:

$$\pi_1(y_1, \ldots, y_n|p_1, \ldots, p_n, k) = \prod_{i=1}^{n} \binom{k + y_i - 1}{y_i} p_i^k (1 - p_i)^{y_i}, \tag{4.12}$$

where $p_1, \ldots, p_n$ are the no accident probabilities. Now, we assume that p_is have conditional beta distribution with α and β parameters. Then the *parameter* level for this model is written as:

$$\pi_2(p_1, \ldots, p_n|\alpha, \beta) = \prod_{i=1}^{n} \frac{p_i^{\alpha-1}(1-p_i)^{\beta-1}}{B(\alpha, \beta)}, \tag{4.13}$$

where $B(\alpha, \beta)$ represents "the beta function". The beta function can also be written with "the gamma functions", i.e. $B(\alpha, \beta) = \dfrac{\text{Gamma}(\alpha)\text{Gamma}(\beta)}{\text{Gamma}(\alpha+\beta)}$. Lastly, in the *prior* level, we assign gamma distribution as a "prior distribution" for α and β, that is $\alpha \sim \text{Gamma}(\alpha|\alpha_1, \alpha_2)$ and $\beta \sim \text{Gamma}(\beta|\beta_1, \beta_2)$. Hence, the *prior* level becomes as given:

$$\begin{aligned} \pi_3(\alpha|\alpha_1, \alpha_2) &= \frac{\alpha_2^{\alpha_1}\alpha^{\alpha_1-1}e^{-\alpha\alpha_2}}{\text{Gamma}(\alpha_1)} \text{ , and} \\ \\ \pi_3(\beta|\beta_1, \beta_2) &= \frac{\beta_2^{\beta_1}\beta^{\beta_1-1}e^{-\beta\beta_2}}{\text{Gamma}(\beta_1)} \end{aligned} \tag{4.14}$$

where (α_1, β_1) and (α_2, β_2) are scale and rate parameters, respectively. As a result, we derive the joint posterior density for negative binomial distribution through the equation by Equation (4.11) as follows:

$$\pi(p_1, \ldots, p_n, \alpha, \beta|y) \propto \underbrace{\pi_1(y_1, \ldots, y_n|p_1, \ldots, p_n, k)}_{\text{likelihood level}}\underbrace{\pi_2(p_1, \ldots, p_n|\alpha, \beta)}_{\text{parameter level}}\underbrace{\pi_3(\alpha|\alpha_1, \alpha_2)\pi_3(\beta|\beta_1, \beta_2)}_{\text{prior level}}. \tag{4.15}$$

If we rewrite Equation (4.15) by using Equation (4.12) and Equation (4.14), then it turns out to be as follows:

$$\begin{aligned} \pi(p_1, \ldots, p_n, \alpha, \beta|y) \propto &\prod_{i=1}^{n} \binom{k+y_i-1}{y_i} p_i^k (1-p_i)^{y_i} \prod_{i=1}^{n} \frac{p_i^{\alpha-1}(1-p_i)^{\beta-1}}{B(\alpha, \beta)} \\ &\times \frac{\alpha_2^{\alpha_1}\alpha^{\alpha_1-1}e^{-\alpha\alpha_2}}{\text{Gamma}(\alpha_1)} \frac{\beta_2^{\beta_1}\beta^{\beta_1-1}e^{-\beta\beta_2}}{\text{Gamma}(\beta_1)}. \end{aligned} \tag{4.16}$$

After giving this example, we adapt a similar approach to our data set in order to see how we could analyze a spatial data using hierarchical Bayesian model. We simply suppose that the response variable Y_s is the crop yield in a subregion as follows:

$$Y_s = \beta X_s^T + \gamma_s + \epsilon_s;\ s = 1, \ldots, n \tag{4.17}$$

where X_s denotes "the vector of predictor variables" and β indicates the slope of X_s. Here, γ_s and ϵ_s is the random effect and random error of Y_s, respectively. Using γ_c, we could investigate the areal effect for the observations in data. In this sense, we consider the following properties which are proposed by Awondo et al. [51]:

$$Y_s|\beta, \gamma_s, \sigma_\epsilon^2 \sim N(\beta X_s^T + \gamma_s, \sigma_\epsilon^2)$$

$$\gamma_s|\sigma_\gamma^2 \sim N(0, \sigma_\gamma^2)$$

$$\tau_\gamma \sim \text{IGamma}(a_1, b_1)$$

$$\tau_\epsilon \sim \text{IGamma}(a_2, b_2)$$

$$\beta \sim N(0, C)$$

where IGamma denotes the Inverse Gamma distribution with a and b parameters. τ and C indicates the precision parameter ($\tau = 1/\text{variance}$) and the variance-covariance matrix, respectively. Using the properties mentioned above, we define the joint posterior density as follows:

$$\pi(\beta, \sigma_\gamma^2, \sigma_\epsilon^2|Y_s) = \pi_1(Y_s|\beta, \gamma_s, \sigma_\epsilon^2)\pi_2(\gamma_s|\sigma_\gamma^2)\pi_3(\sigma_\gamma^2)\pi_3(\sigma_\epsilon^2)\pi_3(\beta). \tag{4.18}$$

Here, we use the random effect for Y_s that is the randomness arising directly from the observations $Y_s^{(i)}$ where i denotes the i-th district with $i = 1, \ldots, n$. Here, we have not taken into account spatial properties yet. For this aim, we use a common approach suggested by Besag et al. [52] which is an "intrinsic conditional autoregressive structure (CAR)" to examine the spatial random effect. Under this approach, the parameter γ mentioned in Equation (4.18) is defined as follows:

$$\gamma_i|\gamma_{i \neq j} \sim N\left(\bar{\gamma}_i, \frac{1}{\tau_\gamma N_i}\right), \tag{4.19}$$

where $\bar{\gamma}_i = N_i^{-1} \sum_{j \in N_i} \gamma_j$ and N_i shows the areas which are neighbours of the district i. The expectectation of γ_i under the condition γ_j is the mean of "spatial random effects" in the set of neighbours whereas the conditional precision parameter τ_γ controls the spatial dependence between the observations. As an example for the term $\bar{\gamma}_i$, we assume a region consisting of eight subregions.

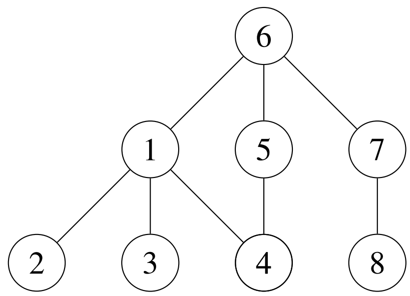

Figure 4.3: The neighbourhood structure of subregions

For example, area 1 share boundries with $2, 3, 4$ and 6. Hence, the conditional expectation of γ_1 is given as follows:

$$\gamma_1|\gamma_{1\neq j} \sim N\left(\frac{\gamma_2+\gamma_3+\gamma_4+\gamma_6}{4}, \frac{1}{4\tau_\gamma}\right). \tag{4.20}$$

The MCMC methods for Bayesian modelling may last a long time because of the complexity of the models or high dimension of the data. INLA approach has been recently introduced by Rue et al. [37] as an alternative to MCMC. INLA is an analytically well-organized approach which can be performed in a large range of models such as spatial, temporal, spatiotemporal, generalized linear mixed models and stochastic volatility models. The most significant advantage of INLA is that it reduces the computational time comparison with the MCMC methods. Moreover, a useful property of INLA is that it allows us to approximate "the posterior distribution" which belongs to the parameter. More detailed information about INLA is given in the following subsection.

4.5. Integrated Nested Laplace Approximation (INLA)

The INLA approach was primarily formulated for latent Gaussian model (LGM). LGM is defined as a hierarchical structure mentioned in Section 4.4.

$$\begin{aligned} &Y|\nu, \psi \sim \pi(y|\nu, \psi) = \prod_{i=1}^{n} \pi(y_i|\nu_i, \psi) \ , \\ &\nu|\psi \sim [\pi(\nu|\psi) = N(0, Q^{-1}(\psi))] \ , \\ &\psi \sim \pi(\psi) \end{aligned} \tag{4.21}$$

where the second part of Equation (4.21) represents latent process. Q^{-1} denotes the precision matrix and equals to the inverse of covariance matrix Σ. Here, ν is the parameter vector and it is described by a "Gaussian Markov Random Field (GMRF)" as follows [53]:

$$\pi(\nu|\psi) = \frac{|Q(\psi)|^{1/2}}{(2\pi)^{n/2}} e^{\left(-\frac{1}{2}\nu^T Q(\psi)\nu\right)}. \tag{4.22}$$

The LGM can be identified through additive regression models. These models are defined similarly as generalized linear models. Nevertheless, additive predictor for the LGM includes nonlinear effects such as seasonal and spatially structured random effects in contrast to the linear predictor in GLM. In our study, we assume that the response or dependent variable $y_i; i = 1, \ldots, n$ belongs to exponential family [54]. Let μ_i represent the mean of i-th observation of y. Then it is characterized by the the additive predictor η_i with a function $g(.)$. Here $g(.)$ denotes the link function, i.e. $g(\mu_i) = \eta_i$. The common form of the predictor η_i is

$$\eta_i = \alpha_0 + \sum_{k=1}^{N_\beta} \beta_k x_{ki} + \sum_{j=1}^{N_f} f_j(z_{ji}) + \epsilon_i \ . \tag{4.23}$$

The terms used in Equation (4.23) are defined as follows:

- α_0 is a scalar which represents the intercept.
- $\beta = (\beta_1, \ldots, \beta_{N_\beta})$ represents the linear term and measures the effect of the vector of covariates x on the response variable.

- $f_j(.)$ is the function of the vector of covariates z and can be used to investigate nonlinear effects of z, i.e. spatial or temporal random effects on the dependent variable.

- N_β and N_f represent the number of corresponding covariates, respectively.

We can define the additive predictor η_i as a hiearchical structure using Equation (4.21). The term $\pi(y|\nu, \psi) = \prod_{i=1}^{n} \pi_1(y_i|\nu_i, \psi)$ means that each observation in data y is linked to i-th element of ν_i in latent field ν [55]. The random vector (latent field) $\nu = \left(\alpha_0, \beta_{k=1}^{N_\beta}, f_{j=1}^{N_f}(.), \eta\right)$ includes all parameters that are not seen directly in the data. Lastly, $\psi = (H_1, \ldots, H_{n_p})$ indicates the n_p-dimensional vector of hyperparameters H_is. Using these definitions, the joint posterior of ν and ψ can be defined as follows.

$$\begin{aligned} \pi(\nu, \psi) &\propto \pi(y|\nu, \psi)\pi(\nu|\psi)\pi(\psi) \\ &\propto \left(\prod_{i=1}^{n} \pi(y_i|\nu_i, \psi)\right)\pi(\nu|\psi)\pi(\psi) \end{aligned} \tag{4.24}$$

where the density function $\pi(\nu|\psi)$ is defined in Equation (4.22). Thus, if we replace Equation (4.22) into Equation (4.24), the following form is obtained.

$$\pi(\nu, \psi) \propto \pi(\psi)|Q(\psi)|^{1/2} exp\left(-\frac{1}{2}\nu^T Q(\psi)\nu + \sum_{i=1}^{n} log(\pi(y|\nu, \psi))\right) \tag{4.25}$$

The purpose of INLA method is to provide an approximation for the marginal posterior distribution of each parameter vector ν and ψ. These marginals for ν and ψ are given separately as follows:

$$\pi(\nu_i|y) = \int \pi(\nu_i, \psi|y)d\psi = \int \pi(\nu_i|\psi, y)\pi(\psi|y)d\psi, \text{ and} \tag{4.26}$$

$$\pi(\psi_h|y) = \int \pi(\psi|y)d\psi_{-h} \tag{4.27}$$

where ψ_{-h} denotes the vector of the remaining hyperparameters when h-th hyperparameter is omitted. As seen in Equation (4.26) and (4.27), both equations have the term $\pi(\psi|y)$ in common. Thus, in order to find an approximation for the marginal posterior distributions for all hyperparameters, we can define $\pi(\psi|y)$ as:

$$\begin{aligned}\pi(\psi|y) = \frac{\pi(\nu_i, \psi|y)}{\pi(\nu|\psi, y)} &\propto \frac{\pi(y|\nu, \psi)\pi(\nu|\psi)\pi(\psi)}{\pi(\nu|\psi, y)} \\ &\approx \left.\frac{\pi(y|\nu, \psi)\pi(\nu|\psi)\pi(\psi)}{\tilde{\pi}(\nu|\psi, y)}\right|_{\nu=\nu^*(\psi)}\end{aligned} \tag{4.28}$$

where $\tilde{\pi}(\nu|\psi, y)$ is the Gaussian approximation of $\pi(\nu|\psi, y)$ and $\nu^*(\psi)$ represents the mode of ν for a given ψ.

The approximation for the posterior conditional distributions $\pi(\nu_i|\psi, y)$ given ψ and y can be more complex due to the fact that the parameter vector ν has more elements than ψ in general. Three approaches can be used to approximate $\pi(\nu_i|\psi, y)$ [56]. These approaches are given below.

i. The marginals from $\tilde{\pi}(\nu|\psi, y)$ are used to approximate $\pi(\nu_i|\psi, y)$ by the help of Normal distribution. Also, precision matrix is obtained by using the Cholesky decomposition. According to the other two approaches, this approach is relatively fast to approximate $\pi(\nu_i|\psi, y)$. However, this approach is not usually very good at the approximation [57].

ii. An alternative approach to the approximation for $\pi(\nu_i|\psi, y)$ is the "Laplace Gaussian approximation". We can rewrite the vector of parameters as $\nu = (\nu_i, \nu_{-i})$ where ν_{-i} denotes the vector of the remaining parameters when i-th parameter is omitted. Then the joint posteriors belong to the parameter ν can be approximated by using "Laplace approximation" as follows:

$$\begin{aligned}\pi(\nu_i|\psi, y) &= \frac{\pi(\nu_i, \nu_{-i}|y)}{\pi(\nu_{-i}|\nu_i, \psi, y)} \\ &\approx \left.\frac{\pi(y|\nu, \psi)\pi(\nu|\psi)\pi(\psi)}{\tilde{\pi}(\nu_{-i}|\nu_i, \psi, y)}\right|_{\nu_{-i}=\nu^*_{-i}(\nu_i,\psi)}\end{aligned} \tag{4.29}$$

where $\tilde{\pi}(\nu_{-i}|\nu_i, \psi, y)$ represents Laplace Gaussian approximation of $\pi(\nu_{-i}|\nu_i, \psi, y)$ and $\nu^*_{-i}(\nu_i, \psi)$ is the mod of ν_{-i}. Laplace Gaussian approximation works effectively, however it takes a long computational time.

iii. The last approach, "simplified Laplace approximation" is based on Taylor's series of

Equation (4.29). It is more rational and computationally efficient in comparison with the other two methods mentioned above.

We can write the approximated marginal posteriors of $\pi(\nu_i|y)$ and $\pi(\psi_h|y)$ in Equations (4.26) and (4.27) as follows:

$$\tilde{\pi}(\nu_i|y) = \int \tilde{\pi}(\nu_i|\psi, y)\tilde{\pi}(\psi_l|y)d\psi, \tag{4.30}$$

$$\tilde{\pi}(\psi_h|y) = \int \tilde{\pi}(\psi|y)d\psi_{-h}. \tag{4.31}$$

The solution of the given integral in the Equation (4.30) can be obtained by numerical integration as

$$\tilde{\pi}(\nu_i|y) = \sum_{l=1}^{L} \tilde{\pi}(\nu_i|\psi^{(l)}, y)\tilde{\pi}(\psi^{(l)}|y)\Delta_l \tag{4.32}$$

where Δ_l denotes the set of weights and $\psi^{(l)}$ represents some integration points. More detailed information about these approximations and INLA can be found in Rue et al. [37] and Blangiardo and Cameletti [58].

4.5.1. Model adequacy based on predictive distribution

Assessing the adequacy of the model is very crucial to make sensible decisions after modelling process. The most popular approaches related to Bayesian modelling are based on predictive distribution. The data is splitted into two groups as training and validation. The training data is used to fit the model whereas the validation data is used to examine the accuracy of the prediction. Posterior predictive check method and/or "leave-one-out-cross-validation" can be used for testing the model adequacy.

- **Posterior predictive check**

  "The posterior predictive distribution" and "the posterior predictive p-value" are two important quantities for the model posterior predictive checks [59]. The mathematical representations of these quantities are given by

$$\pi(y_i^{rep}|y) = \int_\theta \pi(y_i^{rep}|\theta_i)\pi(\theta_i|y)d\theta_i, \tag{4.33}$$

$$P(Y_i^{rep} \leq y_i|y). \tag{4.34}$$

Here, the first equation represents the posterior predictive distribution where $\pi(y_i^{rep}|y)$ denotes the density of a replicated observation Y_i^{rep}. $P(Y_i^{rep} \leq y_i|y)$ represents the posterior predictive p-value [60]. Using this values, it is possible to make an inference about whether the model is appropriate for data or not. If $\pi(y_i^{rep}|y)$ has many small values, it means that the related observation is an outlier. Thus, that model is not suitable for the data set. Morever, if the values belonging to $P(Y_i^{rep} \leq y_i|y)$ are close to zero or one, the model appears to be not valid for the data.

- **Leave-one-out-cross-validation**

  "The conditional predictive ordinate (CPO)" [61] and "the probability integral transform (PIT)" [62] are measures of goodness of fit, i.e. the performance of prediction. The main idea behind these measures is to assign a numerical score to the models according to their predictive distribution. The first measure is defined as

$$\text{CPO}_i = \pi(y_i|y_{-i}) \tag{4.35}$$

  where y_{-i} represents the vector of the remaining observations in the data when i-th observation is omitted. High values for CPO represent good predictive performance of the model for y_i whereas very small CPO values mean that the i-th observation might be an outlier observation. Using the values of CPO, we can also obtain the logarithmic score to choose the appropriate model. This score is computed as follows:

$$\text{Lscore} = -\frac{\sum_i^n \log(CPO_i)}{n}. \tag{4.36}$$

  Here, small log-score value denotes that the interested model is reasonable. The logarithmic score and the Akaike information criterion (AIC) is asymptotically equivalent in the case where each observation in the data set is independent from each other [63].

  The second measure is calculated as follows:

$$\mathrm{PIT}_i = P\left(Y_i^{rep} \leq y_i | y_{-i}\right) \tag{4.37}$$

where PIT_i value represents the calibration of i-th observation to the rest of the data, i.e. i-th observation is removed. If PIT values are extremes, i.e. they are very small or very high, the observations belonging to these values may be outlier. Moreover, the PIT values must have standard uniform distribution, otherwise the model fit is not reasonable. Also, we can use the histogram of PIT to evaluate the convenience of the model for the data.

The PIT values and the histogram of PIT values [53] are presented following figure.

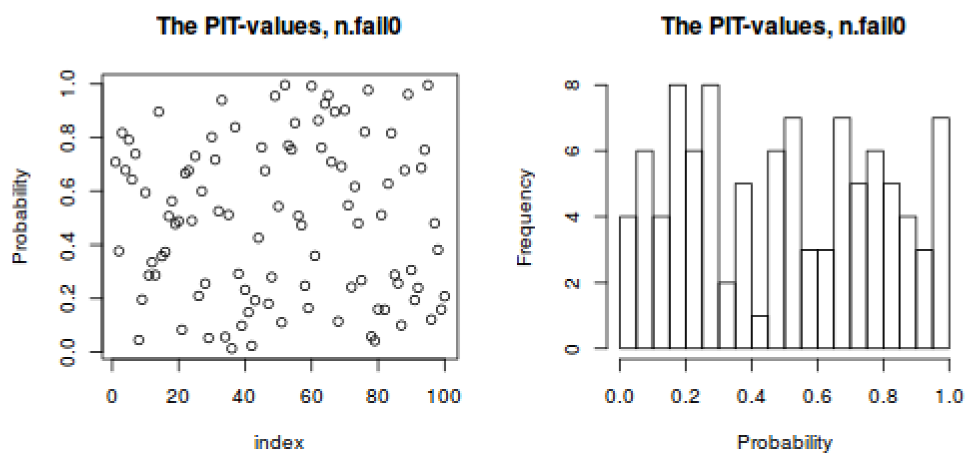

Figure 4.4: Graph of the PIT values (left) and the histogram of PIT (right)

As seen in Figure 4.4, the histogram of these values established on the model represents that the distribution of PIT values is close to uniform.

4.5.2. Model selection

"The Deviance Information Criterion (DIC)" introduced by [64] is used to make a comparison among the fit of different models. The DIC model consists of two terms as "goodness of fit" and "penalty" and is given by

$$DIC = \bar{D} + p_D \tag{4.38}$$

where $\bar{D}$ is posterior expectation of the Bayesian deviance $D(\nu)$ and it is used to evaluate as measure of the model fit.

$$D(\nu) = -2log(\pi(y|\nu)) + 2log(\pi(y)) \tag{4.39}$$

where $\pi(y|\nu)$ denotes the likelihood function, and the term $2log(\pi(y))$ can be removed when comparing different models for the same sample. p_D given in Equation (4.38) is called ”the number of effective parameters” and is used to measure the complexity of the model. It equals to “the expectation of deviance” minus “the deviance of the expectations”.

$$p_D = \bar{D} - D(\bar{\nu}). \tag{4.40}$$

Here, $\bar{\nu}$ denotes the expected value of ν and can be represented by $E_{\nu|y}(\nu)$. Lastly, the model which has the smaller DIC value should be preferred.

4.6. Interim Conclusion: Spatiotemporal Bayesian Modelling of the Conditional Yield Distribution

The conditional yield distribution, which is modelled by hierarchical Bayesian method, takes into consideration of spatial and temporal dependencies among geographical subregions, i.e. districts.

We present hierarchical Bayesian model results and premium calculations in Section 5.

5. CALCULATION OF CROP YIELD INSURANCE PREMIUM USING SPATIOTEMPORAL MODELS

5.1. Introduction

Obtaining the premium rates based on statistical models becomes more of an issue for the modelling of the crop yield due to some factors such as moral hazard, dependency between the catastrophic losses in terms of spatial and temporal effects. Therefore, we take spatial, temporal and spatiotemporal effects into account to model the crop yield. For this purpose, we use hiearchical Bayesian structure for the district-based crop yield data to capture spatial and temporal effects. A general introduction of the hiearchical structure is given in Chapter 4.

We obtain the estimations of the considered models and calculate the premium rates using the R-INLA package provided by R software [56, 58]. For displaying some of the graphs, we use the tools of ArcGIS software.

Firstly, we introduce the data which is used as a representative sample for the case study in Section 5.2. Having explain the structure and properties of the data, Section 5.3 presents the model results. In this section, the models vary according to the consideration of spatial, temporal and spatiotemporal effects on crop yield estimations obtained by the hierarchical Bayesian approach. The performance of each model is also examined by the help of INLA in this part. In Section 5.4, we use two criteria, DIC and Lscore, to determine the best fit among the tested models. We represent the parameter estimations and graphical interpretations for each model. The PIT values are also used to test the model selection in this section. In Section 5.5, we give a short introduction for the theory of the premium calculation. According to this calculation procedure, we obtain the premium rates for the chosen model and display the results numerically in this section. Lastly, we conclude the chapter in Section 5.6.

5.2. Data Description

In this thesis, we use the crop yield data for the years 2004-2018, which were provided by the TUIK (Turkish Statistical Institute), for the crop wheat in the city of Ankara and Konya located in the Central Anatolia region of Turkey. Both cities have an important role in wheat production. Total wheat production for Ankara and Konya in 2018 equals to 2,367,923 tons which is %14.3 of Turkey's total wheat production. The following table represents the amounts of wheat production based on the top ten cities from the years 2016-2018.

Table 5.1: The highest ten cities according to wheat production in Turkey (2016-2018)

2016		**2017**		**2018**	
City	**Production (tons)**	**City**	**Production (tons)**	**City**	**Production (tons)**
Konya	**1,278,320**	**Konya**	**1,419,442**	**Konya**	**1,363,378**
Ankara	**1,083,670**	**Ankara**	**987,908**	**Ankara**	**1,004,545**
Diyarbakır	845,105	Tekirdağ	882,674	Diyarbakır	728,016
Tekirdağ	825,714	Diyarbakır	814,675	Adana	681,736
Adana	621,302	Adana	689,904	Tekirdağ	637,685
Eskişehir	575,430	Şanlıurfa	632,257	Sivas	607,995
Şanlıurfa	571,854	Kırklareli	552,431	Eskişehir	526,919
Sivas	563,469	Eskişehir	538,796	Kırklareli	503,107
Çorum	530,537	Sivas	531,349	Şanlıurfa	492,706
Edirne	522,970	Çorum	516,052	Edirne	482,849

As seen in Table 5.1, Konya and Ankara are the first two cities compared to the others in terms of amount of wheat production. Thus, we choose these cities for the case study in this thesis. We model the crop yield for the districts belong to these cities. These districts are convenient for the consideration of the spatiotemporal dependency since they share common borders which means that they are neighbours.

Table 5.2: The ID of the districts

District	District ID	District	District ID	District	District ID
Akören	1	Elmadağ	17	Kızılcahamam	33
Akşehir	2	Emirgazi	18	Kulu	34
Altındağ	3	Ereğli	19	Mamak	35
Altınekin	4	Evren	20	Meram	36
Ayaş	5	Gölbaşı	21	Nallıhan	37
Bala	6	Güdül	22	Polatlı	38
Beypazarı	7	Güneysınır	23	Sarayönü	39
Beyşehir	8	Halkapınar	24	Selçuklu	40
Bozkır	9	Haymana	25	Seydişehir	41
Cihanbeyli	10	Hüyük	26	Sincan	42
Çankaya	11	Ilgın	27	Şereflikoçhisar	43
Çeltik	12	Kadınhanı	28	Tuzlukçu	44
Çubuk	13	Kalecik	29	Yalıhüyük	45
Çumra	14	Karapınar	30	Yenimahalle	46
Derbent	15	Karatay	31	Yunak	47
Doğanhisar	16	Keçiören	32		

In this study, we use the IDs for the districts for convenience. The corresponding IDs of the districts are given in Table 5.2. As shown in the table, the data set in the application part of our study consists of 47 districts in Konya and Ankara. In order to handle the geographical information for these districts, the following figure is given.

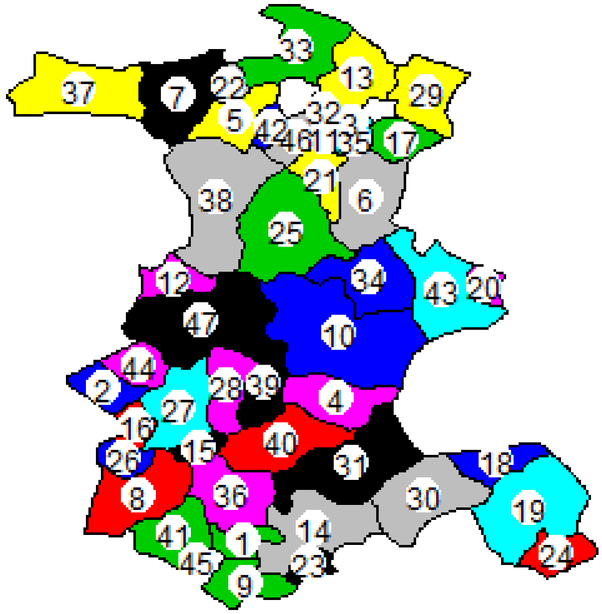

Figure 5.1: The map of the districts with ID

In Figure 5.1, the neighbourhood structures among districts can ben seen visually. For instance, if we handle the District 4 Altınekin, the neighbours from its north to its south-east are District 10 Cihanbeyli, District 39 Sarayönü, District 40 Selçuklu, and District 31 Karatay, respectively.

Having examined the neighbourhood of the districts, we represent the realized wheat yields in terms of kilos for 15 years from 2004 to 2018 in the following figure. In this figure, y-axis shows the latitude values of the districts whereas x-axis presents the longitude values of the districts. In addition, the color scales in the figure indicates the amount of wheat yields in each year where the blue color scale denotes the lowest amounts whereas the yellow color scale denotes the highest amounts.

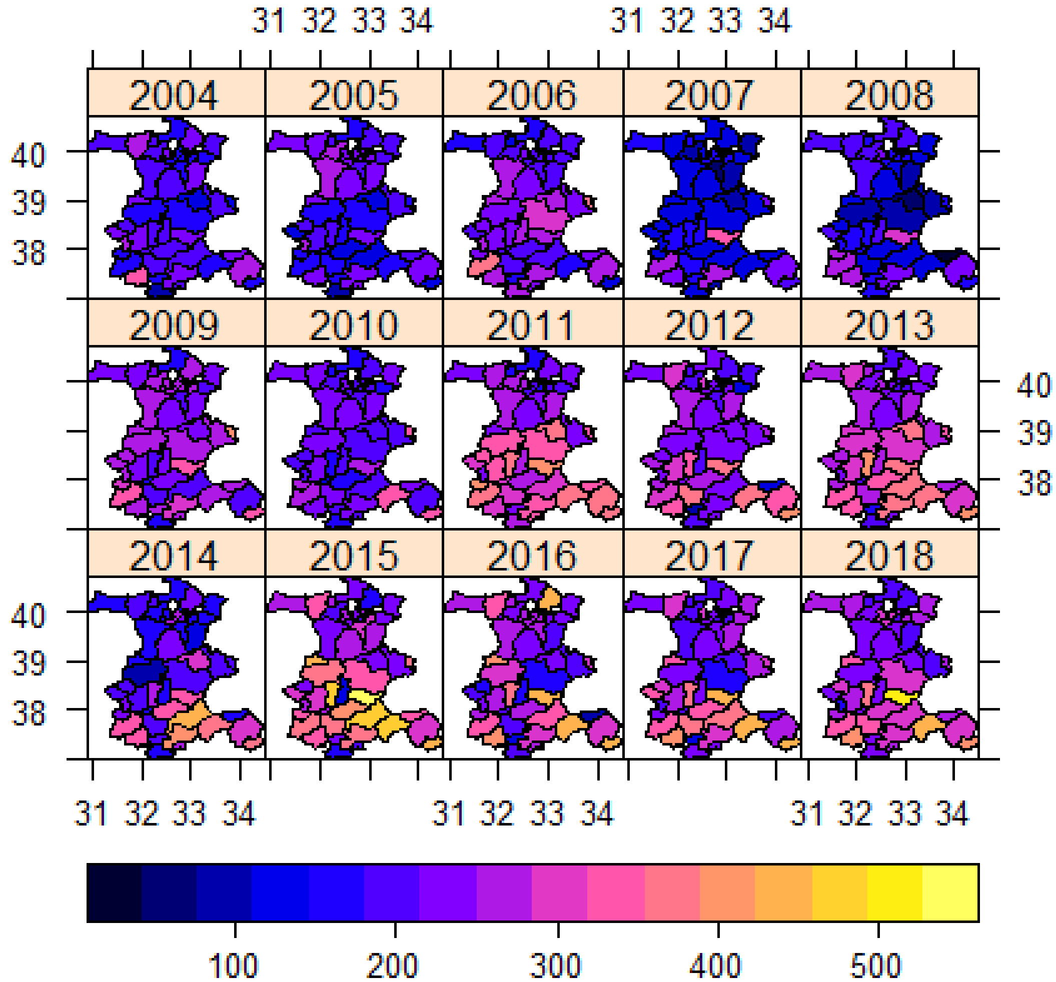

Figure 5.2: Wheat yield from the years 2004-2018

As it is seen from Figure 5.2, the overall colour seems to be changed from the blue color scale to purple-pink-yellow color scale, which means that the overall yearly wheat production has increased about 100 kilos approximately. Another inference of Figure 5.2 is that the change in the color scale seems similar for neighbour districts. This interpretation could be taken as a proof of the accuracy of our assumption in Section 4 that the spatial neighbourhood is used to examine dependence among the subregions.

In order to see the structure of the neighbourhood among districts, Figure 5.3 is given. It represents nodes and lines between nodes to visualize the adjacent districts with their IDs.

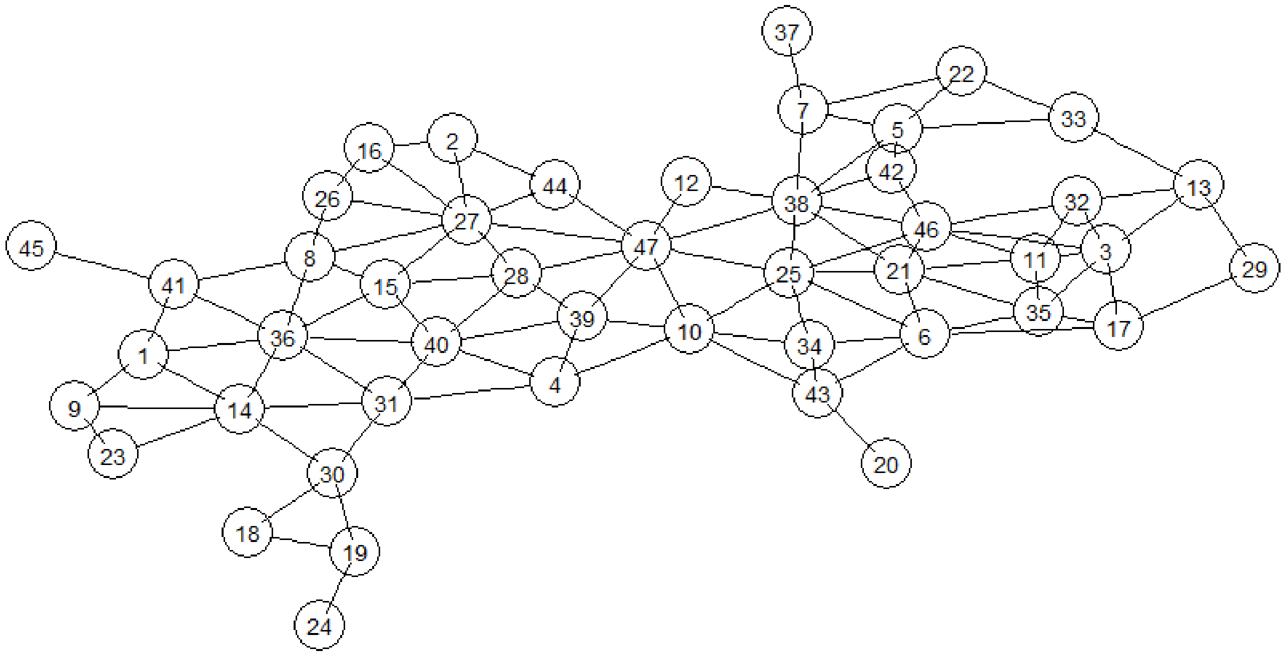

Figure 5.3: The nodes of adjacent districts with their IDs

In this figure, it is very easy to find the neighbourhood relations among districts. For instance, as it is obtained from Figure 5.1, the neighbours of District 4 from right to the left are District 10, District 39, District 40 and District 31, respectively.

Another represention of the neighbourhood is the adjanceny matrix given as follows.

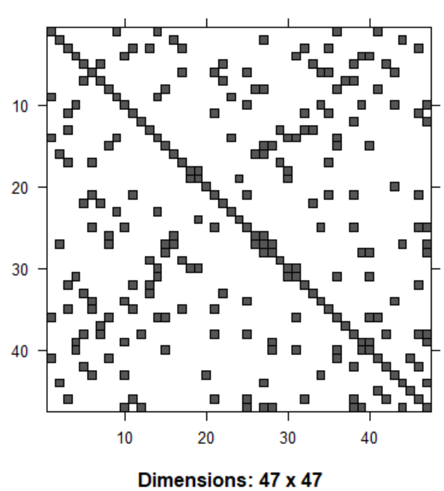

Figure 5.4: The adjanceny matrix for the districts with ID

Adjanceny matrix is a commonly used graphical tool in the literature. The numbers in the x-axis and y-axis in Figure 5.4 indicates the district IDs used for the case study.

5.3. Spatiotemporal Models

Here, we perform various models in order to examine spatial, temporal and spatiotemporal effects on the wheat yield in crop yield insurance. The theoretical considerations and INLA results of these models for the estimation of wheat yield of the districts are given in the following subsections. The general theoretical modelling of "the hierarchial Bayesian approach" is given in Equation (4.17) and the related properties given in Section 4.4. As it is given in this modelling proposed by Awondo et al. [51], the conditional distribution of $Y_s|\beta, \gamma_s, \sigma_\epsilon^2$ is assumed to be normal with the parameters $\beta X_s^T + \gamma_s$ and σ_ϵ^2, respectively.

Suppose that $Y_{s,t}$s are the overall 705 yield observations related to 47 districts for 15 years. We use logaritmic transformation to make the yield data approximate the normal distribution. In addition, using the log of the yield provides us to obtain less skewed data and it reduces heterogeneity in the data. The following figures given below explain why we use the log-scale data.

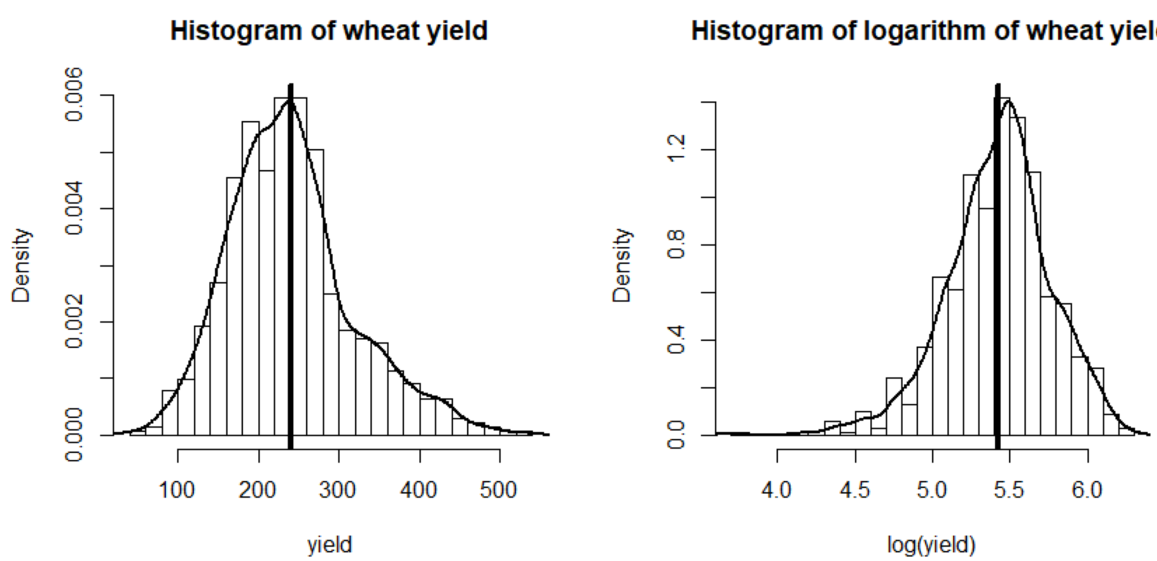

Figure 5.5: The histogram of the densities of the wheat yield observations: Y_{st} (left) and $\log(Y_{st})$ (right)

In Figure 5.5, the histogram of Y_{st} on the left shows that the densitys of the wheat yield observations is skewed which is less convenient for the use of the hierarchial modelling. On the other hand, the logaritmic transformation of the wheat yields are more suitable for our application according to the histogram of $\log(Y_{st})$ on the right.

Moreover, we use scatter plots in order to investigate the heterogeneity in the crop yield data.

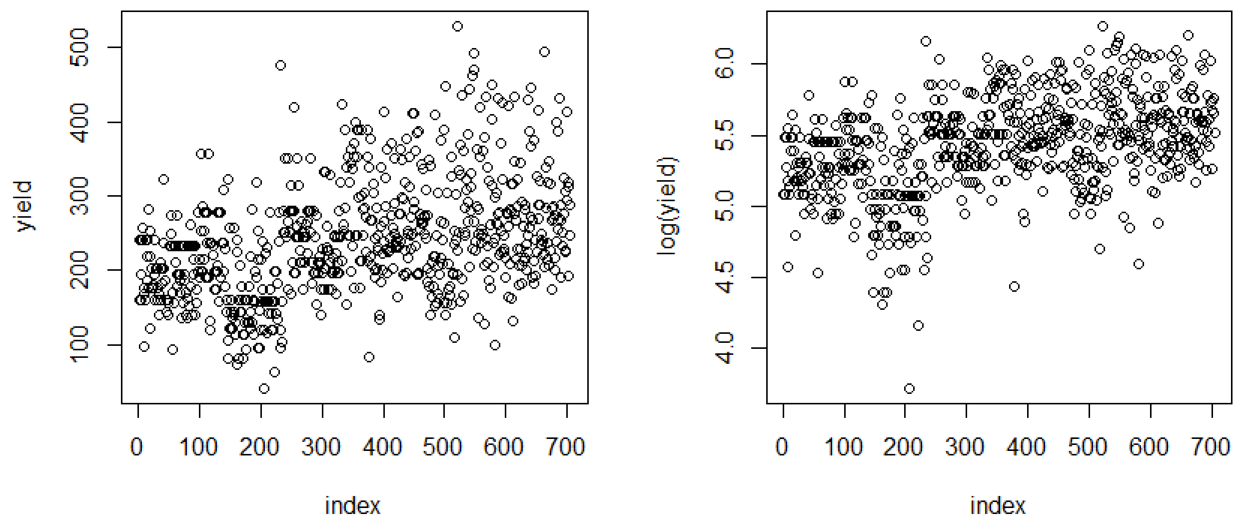

Figure 5.6: The scatter plot of Y_{st} and $\log(Y_{st})$

According to Figure 5.6, the scatter plot of Y_{st} on the left shows that the wheat yield data is a heterogeneous sample whereas the logaritmic transformation of the wheat yields apppears to be a homogeneous sample according to the scatter plot of $\log(Y_{st})$ on the right. Therefore, it is inferred from this figure that the log-yield data reduces the heterogeneity. We use the linear predictor $\eta_{s,t} = \log(Y_{s,t})$ in the forthcoming subsections.

5.3.1. Model 1: Basic error model

We start with the basic model which includes an intercept term and an unstructured spatial random effect. Although this model is less informative in comparison with the other models which are used in this thesis, we consider this model as a baseline model to see the effect of intercept and unstructured spatial effect for the spatial models.

Let $Y_{s,t}^{(i)}$ denote the wheat yield in the district i with location s for the years $t = 1, ..., 15.$ and the basic error model is given by:

$$
\begin{aligned}
&\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \gamma_{s_1}^{(i)} + \epsilon_{s,t}^{(i)}; \quad i = 1, \ldots, 47 \text{ and } t = 1, \ldots, 15 \\
&\alpha_0 \sim N(0, \sigma_{\alpha_0}^2) \\
&\epsilon_{s,t} \sim N(0, \sigma_{\epsilon_{s,t}}^2) \\
&\gamma_{s_1} \sim N(0, \sigma_{\gamma_{s_1}}^2) \\
&\log \tau_{\gamma_{s_1}} \sim \log \text{Gamma}(a_{s_1}, b_{s_1}) \\
&\log \tau_{\epsilon_{s,t}} \sim \log \text{Gamma}(a_{\epsilon_1}, b_{\epsilon_2})
\end{aligned}
\tag{5.1}
$$

where;

- α_0 represents the intercept term for the model and measures the mean of wheat yield for the districts.
- $\gamma_{s_1}^{(i)}$ denotes the spatially unstructured effect and we assume that they are independent and identically distributed (*iid*).

The INLA results of the basic error model are given in Table 5.3. The intercept term α_0 with mean 5.43 and standard deviation (sd) 0.02 is the fixed effect of this model. Also, we investigate the spatial unstructured random effect for each district.

The summary of Model 1 is given in the following table.

Table 5.3: Summary statistics of fixed and random effects for the basic error model

Model 1

$$\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \gamma_{s_1}^{(i)} + \epsilon_{s,t}^{(i)}$$

Parameter	**mean**	**sd**	**0.025quant**	**0.50quant**	**0.975quant**	**mode**
α_0	5.43	0.02	5.38	5.43	5.47	5.43
Precision						
τ_{s_1}	59.76	17.24	33.03	57.37	100.12	52.91
τ_{ϵ}	10.66	0.59	9.54	10.64	11.85	10.62

In Table 5.3, we represent the descriptive statistics (mean, sd, 2.5% quantile, 50% quantile, 97.5% quantile and mode) of the estimations of the intercept term and the precisions τ_{s_1} and τ_{ϵ}.

In addition to the numerical results, a graphical illustration of Model 1 is provided by the density plot given in the following figure.

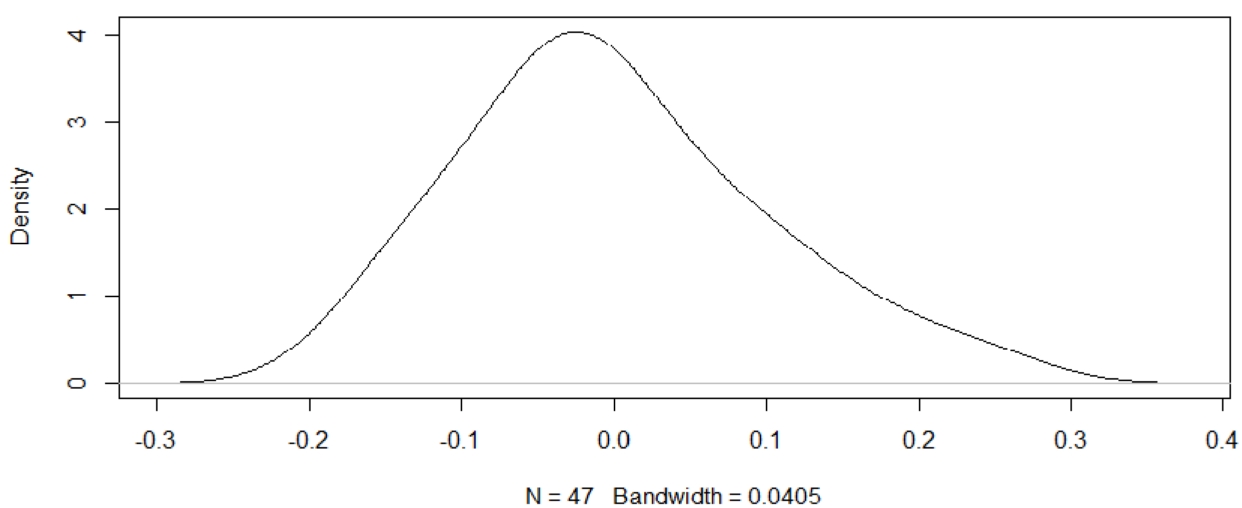

Figure 5.7: basic error model random effects term density

In Figure 5.7, the density for the spatial unstructured random effects seems to have the shape of the normal distribution and to be symmetric around 0.

Before introducing the following model, we also examine whether the estimation is reasonable or not. For this aim, it is expected that the estimated values are close to the observed values in data. Hence, we use the scatter plot of the observed vs the estimated values to check the fit of the basic error model. The scatter plot of observed vs estimated values is given in Figure 5.8.

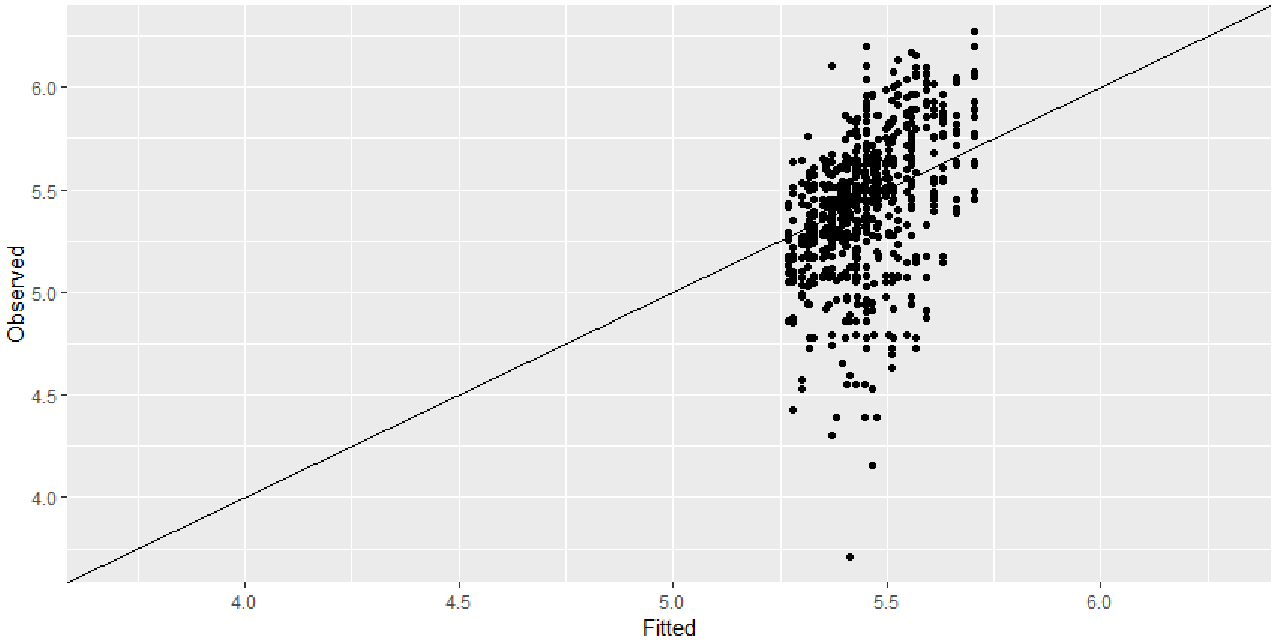

Figure 5.8: Scatter plot of observed vs estimated values

We infer from this figure that the fit of this model is not good since most of the points in the figure are far away from the $x = y$ line. In Section 5.4, we call this model as Model 1 for the comparison of the used models.

5.3.2. Model 2: Spatial model

In this model, we extend the basic error model (Model 1) to the spatial model (Model 2) in order to examine the spatial dependence. For this aim, we add the spatially structured random effect to Model 1 given in Equation (5.1). The second model is given as follows [56]:

$$
\begin{aligned}
&\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + \epsilon_{s,t}^{(i)}; \quad i = 1, \ldots, 47 \text{ and } t = 1, \ldots, 15 \\
&\alpha_0 \sim N(0, \sigma_{\alpha_0}^2) \\
&\epsilon_{s,t} \sim N(0, \sigma_{\epsilon_{s,t}}^2) \\
&\gamma_{s_1} \sim N(0, \sigma_{\gamma_{s_1}}^2) \\
&\gamma_{s_2} \sim N(0, \sigma_{\gamma_{s_2}}^2) \\
&\log \tau_{\gamma_{s_1}} \sim \log \text{Gamma}(a_{s_1}, b_{s_1}) \\
&\log \tau_{\gamma_{s_1}} \sim \log \text{Gamma}(a_{s_2}, b_{s_2}) \\
&\log \tau_{\epsilon_{s,t}} \sim \log \text{Gamma}(a_{\epsilon_1}, b_{\epsilon_2})
\end{aligned} \tag{5.2}
$$

where $\gamma_{s_2}^i$ denotes the spatial structured term. We define the properties of this term in Equation (4.19). The summary results of Model 2a are given in the following table.

Table 5.4: Summary statistics of fixed and random effects for only spatial model

Model 2a $\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + \epsilon_{s,t}^{(i)}$						
Parameter	**mean**	**sd**	**0.025quant**	**0.50quant**	**0.975quant**	**mode**
α_0	5.43	0.02	5.39	5.43	5.47	5.43
Precision						
τ_{s_1}	78.82	31.05	30.40	74.51	150.43	65.46
τ_{s_2}	850.41	2,284.31	50.23	325.88	4,941.66	111.06
τ_{ϵ}	10.65	0.59	9.54	10.63	11.85	10.60

In Table 5.4, we represent the descriptive statistics (mean, sd, 2.5% quantile, 50% quantile, 97.5% quantile and mode) of the estimations of the intercept term and the precisions τ_{s_1}, τ_{s_2} and τ_{ϵ}.

Considering the results for the fixed effect α_0, the mean 5.43 and sd 0.02 are similar with the Model 1. The precise for spatially unstructured and structured term are 78.82 and 850.41, respectively. If we consider this model with the covariates, . Therefore, we consider the city wheat yield for the districts and their elevation to estimate crop yield. But, the elevation is not meaningful for the models. So, we use only the city wheat yield as covariate. The linear regression results without the spatial effects are given below.

In this study, we firstly investigate the models which consider random effects without covariates. After that, we add the covariates such as log-scale of yields related to the city $\log(\text{yield}_{city})$ and log-scale of elevations $\log(\text{Elevation})$ to the models which are commonly used in the literature. Due to the lack of farmer-based yield data, we do not use the data related to the average meteorological variables of all districts together.

We apply a basic regression method to the data in order to examine the effects of $\log(\text{yield}_{city})$ and $\log(\text{Elevation})$ on the crop yield. The results are given in the following table.

Table 5.5: The linear regression model results

	Estimate	Std. Error	t value	p
intercept	0,98032	0,42199	2,323	0,0205
log(yield$_c$)	0,91734	0,04063	22,576	<2e-16*
log(Elevation)	-0,08041	0,05568	-1,444	0,1491

In Table 5.5, the results of the analysis show that only yearly average of the wheat yield of city-based districts, i.e. the $\log(\text{yield}_c)$ variable increase the estimation power of the model. The results also indicates that the district elevation does not have an impact on the model. Therefore, we use only the yearly average of the yield of city-based districts $\log(\text{yield}_c)$ as a covariate in our models.

When we add the covariate into the model, the spatial model with the covariate turns out to be as follows:

$$
\begin{aligned}
&\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \log(\text{yield}_c) + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + \epsilon_{s,t}^{(i)}; \quad i = 1, \ldots, 47 \text{ and } t = 1, \ldots, 15 \\
&\alpha_0 \sim N(0, \sigma_{\alpha_0}^2) \\
&\log(\text{yield}_c) \sim N(0, \sigma_{\log(\text{yield}_c)}^2) \\
&\epsilon_{s,t} \sim N(0, \sigma_{\epsilon_{s,t}}^2) \\
&\gamma_{s_1} \sim N(0, \sigma_{\gamma_{s_1}}^2) \\
&\gamma_{s_2} \sim N(0, \sigma_{\gamma_{s_2}}^2) \\
&\log \tau_{\gamma_{s_1}} \sim \log \text{Gamma}(a_{s_1}, b_{s_1}) \\
&\log \tau_{\gamma_{s_1}} \sim \log \text{Gamma}(a_{s_2}, b_{s_2}) \\
&\log \tau_{\epsilon_{s,t}} \sim \log \text{Gamma}(a_{\epsilon_1}, b_{\epsilon_2})
\end{aligned}
\tag{5.3}
$$

The model outputs related to Equation (5.3) are given in Table 5.6.

Table 5.6: Summary statistics of fixed and random effects for spatial model with the covariate

Model 2b

$\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \log(\text{yield}_c) + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + \epsilon_{s,t}^{(i)}$

Parameter	**mean**	**sd**	**0.025quant**	**0.50quant**	**0.975quant**	**mode**
α_0	0.51	0.19	0.13	0.51	0.88	0.51
$\log(\text{yield}_c)$	0.90	0.04	0.83	0.90	0.97	0.90
Precision						
τ_{s_1}	59.62	14.88	35.32	57.99	93.51	54.90
τ_{s_2}	1,879.80	1,837.59	141.86	1,344.04	710.70	396.48
τ_{ϵ}	21.21	1.17	18.98	21.18	23.58	21.14

In Table 5.6, we represent the descriptive statistics of the estimations of the fixed effects α_0 and $\log(\text{yield}_c)$ (with the mean 0.51 and 0.90, respectively) and the precisions τ_{s_1}, τ_{s_2} and τ_{ϵ}.

5.3.3. Model 3: Spatial Temporal Model

We revisit the spatial model in Equation (5.2) in order to examine temporal effects on crop yield estimation. In this section, we firstly introduce the parametric representation for the spatial-temporal modelling which was proposed by Bernardinelli et al. [65]. Model 3a is defined as follows:

$$
\begin{aligned}
&\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + \left(\alpha + \delta_{s,t}^{(i)}\right) t + \epsilon_{s,t}^{(i)}; \quad i = 1, \dots, 47 \text{ and } t = 1, \dots, 15 \\
&\alpha_0 \sim N(0, \sigma_{\alpha_0}^2) \\
&\epsilon_{s,t} \sim N(0, \sigma_{\epsilon_{s,t}}^2) \\
&\gamma_{s_1} \sim N(0, \sigma_{\gamma_{s_1}}^2) \\
&\gamma_{s_2} \sim N(0, \sigma_{\gamma_{s_2}}^2) \\
&\delta_{s,t} \sim N(0, \sigma_{\delta_{st}}^2) \\
&\log \tau_{\gamma_{s_1}} \sim \log \text{Gamma}(a_{s_1}, b_{s_1}) \\
&\log \tau_{\gamma_{s_1}} \sim \log \text{Gamma}(a_{s_2}, b_{s_2}) \\
&\log \tau_{\delta_{s,t}} \sim \log \text{Gamma}(a_{\delta_{s,t}}, b_{\delta_{s,t}}) \\
&\log \tau_{\epsilon_{s,t}} \sim \log \text{Gamma}(a_{\epsilon_1}, b_{\epsilon_2})
\end{aligned}
\tag{5.4}
$$

Here, the term $\left(\alpha + \delta_{s,t}^{(i)}\right) t$ consists of two components where:

- α denotes a main linear trend which represents the overall trend effect.
- $\delta_{s,t}^{(i)}$ represents the time trend related to the district i and it is used to define the interaction for space and time.

The summary of Model 3a outputs are presented in Table 5.7.

Table 5.7: Summary statistics of fixed and random effects for spatial-temporal model

Model 3a

$$\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + \left(\alpha + \delta_{s,t}^{(i)}\right) t + \epsilon_{s,t}^{(i)}$$

Parameter	mean	sd	0.025quant	0.50quant	0.975quant	mode
α_0	5.17	0.02	5.13	5.17	5.22	5.17
year	0.03	0.00	0.03	0.03	0.04	0.03
Precision						
τ_{s_1}	576.57	964.68	79.33	306.66	2,737.01	145.74
τ_{s_2}	1,101.22	1,199.15	41.72	716.97	4,342.66	90.84
$\tau_{\delta_{s,t}}$	5,432.48	1,843.43	2,601.55	5,179.33	9,775.09	4,691.66
τ_{ϵ}	14.23	0.80	12.70	14.21	15.85	14.19

In Table 5.7, we represent the descriptive statistics of the estimations of the fixed effects α_0 and year where year indicates the coefficient of t in the component $\left(\alpha + \delta_{s,t}^{(i)}\right) t$ and the precisions τ_{s_1}, τ_{s_2}, $\tau_{\delta_{s,t}}$ and τ_{ϵ}.

The spatial-temporal model with the covariate is given as follows:

$$\begin{aligned}
&\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \log(\text{yield}) + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + (\alpha + \delta_{s,t}^{(i)})t + \epsilon_{s,t}^{(i)}; \\
&\qquad i = 1, \ldots, 47 \text{ and } t = 1, \ldots, 15 \\
&\alpha_0 \sim N(0, \sigma_{\alpha_0}^2) \\
&\log(\text{yield}_c) \sim N(0, \sigma_{\log(\text{yield}_c)}^2) \\
&\epsilon_{s,t} \sim N(0, \sigma_{\epsilon_{s,t}}^2) \\
&\gamma_{s_1} \sim N(0, \sigma_{\gamma_{s_1}}^2) \\
&\gamma_{s_2} \sim N(0, \sigma_{\gamma_{s_2}}^2) \\
&\delta_{s,t} \sim N(0, \sigma_{\delta_{st}}^2) \\
&\log \tau_{\gamma_{s_1}} \sim \log \text{Gamma}(a_{s_1}, b_{s_1}) \\
&\log \tau_{\gamma_{s_1}} \sim \log \text{Gamma}(a_{s_2}, b_{s_2}) \\
&\log \tau_{\delta_{s,t}} \sim \log \text{Gamma}(a_{\delta_{s,t}}, b_{\delta_{s,t}}) \\
&\log \tau_{\epsilon_{s,t}} \sim \log \text{Gamma}(a_{\epsilon_1}, b_{\epsilon_2})
\end{aligned} \tag{5.5}$$

The summary of the estimation results are given in Table 5.8.

Table 5.8: Summary statistics of fixed and random effects for spatial-temporal model with covariate

Model 3b

$$\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \log(\text{yield}_c) + \gamma_{s1}^{(i)} + \gamma_{s2}^{(i)} + (\alpha + \delta_{s,t}^{(i)})t + \epsilon_{s,t}^{(i)}$$

Parameter	**mean**	**sd**	**0.025quant**	**0.50quant**	**0.975quant**	**mode**
α_0	0.75	0.22	0.32	0.75	1.18	0.75
year	0.01	0.00	0.00	0.01	0.01	0.01
$\log(\text{yield}_c)$	0.85	0.04	0.77	0.85	0.93	0.85
Precision						
τ_{s_1}	114.01	48.66	50.75	103.51	237.31	86.42
τ_{s_2}	1,643.67	1,686.97	97.49	1,134.49	6,173.57	258.15
$\tau_{\delta_{s,t}}$	6,674.77	2,236.68	3,404.32	6,300.23	12,094.70	5,623.74
τ_{ϵ}	23.43	1.34	20.87	23.40	26.16	23.35

In Table 5.8, we represent the descriptive statistics of the estimations of the fixed effects α_0, year and $\log(\text{yield}_c)$ and the precisions τ_{s_1}, τ_{s_2}, $\tau_{\delta_{s,t}}$ and τ_{ϵ}.

Now, we present a non-parametric formulation, which is borrowed from Knorr-Held [66] for the spatial-temporal modelling. This model is given by:

$$\begin{aligned}
&\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + \phi_{t_1} + \phi_{t_2} + \epsilon_{s,t}^{(i)}; \quad i = 1, \dots, 47 \text{ and } t = 1, \dots, 15 \\
&\alpha_0 \sim N(0, \sigma_{\alpha_0}^2) \\
&\epsilon_{s,t} \sim N(0, \sigma_{\epsilon_{s,t}}^2) \\
&\gamma_{s_1} \sim N(0, \sigma_{\gamma_{s_1}}^2) \\
&\gamma_{s_2} \sim N(0, \sigma_{\gamma_{s_2}}^2) \\
&\phi_{t_1} \sim N(0, \sigma_{\phi_{t_1}}^2) \\
&\phi_{t_2}^{(1)} \sim N\left(0, \frac{1}{\tau(1-\rho)}\right)
\end{aligned} \tag{5.6}$$

where the parameters of the model have the similar characteristics as in the previous models except the terms ϕ_{t_1} and ϕ_{t_2}. ϕ_{t_1} represents temporally unstructured random effect and it is modelled by using a Gaussian exchangeable prior, i.e. $\phi_{t_1} \sim N(0, \sigma_{\phi_{t_1}}^2)$. In addition to this, the term ϕ_{t_2} denotes the structured temporal effect and it is modelled as an autoregressive process as follows:

$$\phi_{t_2}^{(n)} = \rho\phi_{t_2}^{(n-1)} + \epsilon_n, \quad n = 2, \ldots, t \text{ and } \mid \rho \mid < 1$$

Here, we use the first-order autoregressive (AR1) process for temporally structured random effect $\phi_{t_2}^{(n)}$.

The summary results of the non-parametric model (Model 3c) given in Equation (5.6) are displayed in the following table.

Table 5.9: Summary statistics of fixed and random effects for the non-parametric model time

Model 3c

$$\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + \phi_{t_1} + \phi_{t_2} + \epsilon_{s,t}^{(i)}$$

Parameter	**mean**	**sd**	**0.025quant**	**0.50quant**	**0.975quant**	**mode**
α_0	5.43	0.09	5.24	5.43	5.61	5.43
Precision						
τ_{s_1}	62.57	20.59	27.92	61.03	107.10	57.19
τ_{s_2}	953.40	2,815.00	52.39	342.90	5,671.00	115.75
τ_{t_1}	17,050.00	17,810.00	1,125.91	11,710.00	64,060.00	3,042.11
τ_{t_2}	27.34	11.64	9.09	26.05	53.50	22.14
τ_{ϵ}	19.36	1.08	17.29	19.34	21.53	19.27
ρ	0.48	0.17	0.12	0.49	0.79	0.51

In Table 5.9, we represent the descriptive statistics of the estimations of the fixed effects α_0 and the precisions τ_{s_1}, τ_{s_2}, τ_{t_1}, τ_{t_2}, τ_{ϵ} and ρ.

The extended version of the previous model is obtained by adding the covariate. Non-parametric model with the covariate (Model 3d) is given in the following equation.

$$
\begin{aligned}
&\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \log(\textbf{yield}_c) + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + \phi_{t_1} + \phi_{t_2} + \epsilon_{s,t}^{(i)};\\
&\qquad i = 1, \ldots, 47 \text{ and } t = 1, \ldots, 15\\
&\alpha_0 \sim N(0, \sigma_{\alpha_0}^2)\\
&\log(\textbf{yield}_c) \sim N(0, \sigma_{\log(\text{yield}_c)}^2)\\
&\epsilon_{s,t} \sim N(0, \sigma_{\epsilon_{s,t}}^2)\\
&\gamma_{s_1} \sim N(0, \sigma_{\gamma_{s_1}}^2)\\
&\gamma_{s_2} \sim N(0, \sigma_{\gamma_{s_2}}^2)\\
&\phi_{t_1} \sim N(0, \sigma_{\phi_{t_1}}^2)\\
&\phi_{t_2}^{(1)} \sim N\left(0, \frac{1}{\tau(1-\rho)}\right)
\end{aligned}
\tag{5.7}
$$

The results related to Model 3d are given in Table 5.10.

Table 5.10: Summary statistics of fixed and random effects for the non parametric model with the covariate

Model 3d $\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \log(\textbf{yield}_c) + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + \phi_{t_1} + \phi_{t_2} + \epsilon_{s,t}^{(i)}$						
Parameter	**mean**	**sd**	**0.025quant**	**0.50quant**	**0.975quant**	**mode**
α_0	0.54	0.21	0.13	0.54	0.95	0.53
$\log(\textbf{yield}_c)$	0.90	0.04	0.82	0.90	0.97	0.90
Precision						
τ_{s_1}	59.41	14.85	35.32	57.71	93.42	54.49
τ_{s_2}	1,930.00	1,860.00	142.83	1,389.00	6,823.00	395.87
τ_{t_1}	22,050.00	20,030.00	2,390.67	16,440.00	75,250.00	6,794.21
τ_{t_2}	21,660.00	21,390.00	1,817.69	15,420.00	78,840.00	5,129.84
τ_ϵ	21.24	1.17	19.00	21.22	23.62	21.19
ρ	0.10	0.69	-0.98	0.21	0.99	0.51

In Table 5.10, we represent the descriptive statistics of the estimations of the fixed effects α_0 and $\log(\textbf{yield}_c)$ and the precisions τ_{s_1}, τ_{s_2}, τ_{t_1}, τ_{t_2}, τ_ϵ and ρ.

Another objective of this study is to examine the spatial-temporal interactions, which explain differences of yield amount according to the spatial and temporal interaction trends of different districts. In this sense, the model given in Equation (5.6) can be extended by adding an unstructured interaction term $\delta_{s,t}$ as follows:

$$
\begin{aligned}
&\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + \phi_{t_1} + \phi_{t_2} + \delta_{s,t}^{(i)} + \epsilon_{s,t}^{(i)}; \quad i = 1, \ldots, 47 \text{ and } t = 1, \ldots, 15 \\
&\alpha_0 \sim N(0, \sigma^2_{\alpha_0}) \\
&\epsilon_{s,t} \sim N(0, \sigma^2_{\epsilon_{s,t}}) \\
&\gamma_{s_1} \sim N(0, \sigma^2_{\gamma_{s_1}}) \\
&\gamma_{s_2} \sim N(0, \sigma^2_{\gamma_{s_2}}) \\
&\phi_{t_1} \sim N(0, \sigma^2_{\phi_{t_1}}) \\
&\phi_{t_2}^{(1)} \sim N\left(0, \frac{1}{\tau(1-\rho)}\right) \\
&\delta_{s,t} \sim N(0, \sigma^2_{\delta_{st}})
\end{aligned}
\tag{5.8}
$$

Here, we firstly define $\delta_{s,t}$ as the interaction term between the unstructured effects γ_{s_1} and ϕ_{t_1}. We define $\boldsymbol{\Theta_\delta}$ as a structure matrix for the unstructured effects γ_{s_1} and ϕ_{t_1}, and can be represented by the *Kronecker product* i.e, $\boldsymbol{\Theta}_\delta = \boldsymbol{\Theta}_{\gamma_{s_1}} \otimes \boldsymbol{\Theta}_{\phi_{t_1}} = \boldsymbol{I} \otimes \boldsymbol{I} = \boldsymbol{I}$ [67]. Since γ_{s_1} and ϕ_{t_1} represent the unstructured spatial and temporal effects, there is no spatial or temporal effect on the interaction term $\delta_{s,t}$. Using this result, we assume that the interaction term is normally distributed with mean 0and variance $\dfrac{1}{\tau_{\delta_{s,t}}}$ and it is assumed *iid*.

The results of the unstructured spatiotemporal interaction model (Model 3e) which includes the interaction term is given in Table 5.11.

Table 5.11: Summary statistics of fixed and random effects for the unstructured spatiotemporal interaction model

Model 3e

$$\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + \phi_{t_1} + \phi_{t_2} + \delta_{s,t}^{(i)} + \epsilon_{s,t}^{(i)}$$

Parameter	**mean**	**sd**	**0.025quant**	**0.50quant**	**0.975quant**	**mode**
α_0	5.43	0.08	5.25	5.43	5.60	5.43
Precision						
τ_{s_1}	60.45	17.89	30.14	59.08	99.70	56.23
τ_{s_2}	863.10	2026.00	57.09	362.70	4,766.00	131.69
τ_{t_1}	19,320.00	18,770.00	1,239.86	13,790.00	68,860.00	3,341.81
τ_{t_2}	27.98	10.66	11.58	26.60	52.78	23.61
$\tau_{\delta_{s,t}}$	39.10	4.20	29.79	38.61	48.14	38.51
τ_{ϵ}	39.01	4.18	29.74	38.53	48.02	38.43
ρ	0.43	0.21	0.01	0.43	0.81	0.43

Table 5.11 represents the descriptive statistics of the estimations of the fixed effects α_0 and the precisions τ_{s_1}, τ_{s_2}, τ_{t_1}, τ_{t_2}, $\tau_{\delta_{s,t}}$, τ_{ϵ} and ρ.

The unstructured spatiotemporal interaction model with covariate, which is extension of Model 3e, is given as follows:

$$\begin{aligned}
&\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \log(\text{yield}_c) + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + \phi_{t_1} + \phi_{t_2} + \delta_{s,t}^{(i)} + \epsilon_{s,t}^{(i)}; \\
&\qquad i = 1, \ldots, 47 \text{ and } t = 1, \ldots, 15 \\
&\alpha_0 \sim N(0, \sigma_{\alpha_0}^2) \\
&\log(\text{yield}_c) \sim N(0, \sigma_{\log(\text{yield}_c)}^2) \\
&\epsilon_{s,t} \sim N(0, \sigma_{\epsilon_{s,t}}^2) \\
&\gamma_{s_1} \sim N(0, \sigma_{\gamma_{s_1}}^2) \\
&\gamma_{s_2} \sim N(0, \sigma_{\gamma_{s_2}}^2) \\
&\phi_{t_1} \sim N(0, \sigma_{\phi_{t_1}}^2) \\
&\phi_{t_2}^{(1)} \sim N\left(0, \frac{1}{\tau(1-\rho)}\right) \\
&\delta_{s,t} \sim N(0, \sigma_{\delta_{st}}^2)
\end{aligned} \tag{5.9}$$

The summary of the results of the unstructured spatiotemporal interaction model with the covariate (Model 3f) is given in Table 5.12.

Table 5.12: Summary statistics of fixed and random effects for the unstructured spatiotemporal interaction model with the covariate

Model 3f

$$\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \log(\text{yield}_c) + \gamma_{s1}^{(i)} + \gamma_{s2}^{(i)} + \phi_{t_1} + \phi_{t_2} + \delta_{s,t}^{(i)} + \epsilon_{s,t}^{(i)}$$

Parameter	mean	sd	0.025quant	0.50quant	0.975quant	mode
α_0	0.54	0.21	0.13	0.54	0.95	0.53
$\log(\text{yield}_c)$	0.90	0.04	0.82	0.90	0.97	0.90
Precision						
τ_{s_1}	59.45	14.82	35.28	57.81	93.24	54.71
τ_{s_2}	1,945.00	1,866.00	140.61	1,402.00	6,861.00	391.27
τ_{t_1}	22,140.00	19,820.00	2,322.20	16,650.00	74,770.00	6,679.18
τ_{t_2}	21,440.00	21,680.00	1,818.48	15,070.00	79,430.00	5,102.02
$\tau_{\delta_{s,t}}$	18,470.00	18,330.00	1,267.57	13,060.00	66,900.00	3,456.48
τ_ϵ	21.26	1.18	19.02	21.24	23.65	21.21
ρ	0.10	0.69	-0.98	0.21	0.99	0.99

Table 5.12 represents the descriptive statistics of the estimations of the fixed effects α_0 and $\log(\text{yield}_c)$ and the precisions τ_{s_1}, τ_{s_2}, τ_{t_1}, τ_{t_2}, $\tau_{\delta_{s,t}}$, τ_ϵ and ρ.

Lastly, we consider the effect of the spatial and temporal structure on the interaction term. Therefore, the estimation of the parameters related to the term $\delta_{s,t}$ is not obtained under the assumption that $\delta_{s,t}$ is *iid* as given in Equations (5.8) and (5.9).

The aim of using Model 3g given in Equation (5.9) and Model 3h given in Equations (5.10) is to investigate the overall trend which is correlated with both spatial and temporal characteristics of the data according to their neighbours. Therefore, the interaction term $\delta_{s,t}$ is added to the model as a random effect. We employ the " *Besag-York-Mollie (BYM)*" model [52] for spatially structured random effect and *AR1* process for the structured temporal random effect.

The fixed and random effects for the structured spatiotemporal interaction model (Model 3g) is given as follows:

$$
\begin{aligned}
&\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + \phi_{t_1} + \phi_{t_2} + \delta_{s,t}^{(i)} + \epsilon_{s,t}^{(i)}; \quad i = 1, \ldots, 47 \text{ and } t = 1, \ldots, 15 \\
&\alpha_0 \sim N(0, \sigma_{\alpha_0}^2) \\
&\epsilon_{s,t} \sim N(0, \sigma_{\epsilon_{s,t}}^2) \\
&\gamma_{s_1} \sim N(0, \sigma_{\gamma_{s_1}}^2) \\
&\gamma_{s_2} \sim N(0, \sigma_{\gamma_{s_2}}^2) \\
&\phi_{t_1} \sim N(0, \sigma_{\phi_{t_1}}^2) \\
&\phi_{t_2}^{(1)} \sim N\left(0, \frac{1}{\tau(1-\rho)}\right) \\
&\delta_{s,t} \text{ is the random effect.}
\end{aligned} \tag{5.10}
$$

According to Equation (5.10), the summary statistics of the structured spatiotemporal interaction model (Model 3g) is given in Table 5.13.

Table 5.13: Summary statistics of fixed and random effects for the structured spatiotemporal interaction model

Model 3g $\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + \phi_{t_1} + \phi_{t_2} + \delta_{s,t}^{(i)} + \epsilon_{s,t}^{(i)}$						
Parameter	**mean**	**sd**	**0.025quant**	**0.50quant**	**0.975quant**	**mode**
α_0	5.43	0.11	5.20	5.43	5.64	5.43
Precision						
τ_{s_1}	110.79	49.64	48.55	99.35	238.15	81.41
τ_{s_2}	1903.80	1850.00	138.10	1363.29	6783.75	380.76
τ_{t_1}	168.17	4.89e+12	15.72	95.87	836.21	44.75
τ_{t_2}	55.93	6.62e+9	5.68	40.93	214.51	17.90
$\tau_{\delta_{s,t}}$	12.93	2.35	8.91	12.73	18.13	12.34
τ_ϵ	43.34	5.32	33.61	43.11	54.46	42.74
ρ	0.54	0.31	-0.31	0.62	0.92	0.78
$\rho_{\delta_{s,t}}$	0.66	0.07	0.51	0.66	0.79	0.66

Table 5.13 represents the descriptive statistics of the estimations of the fixed effects α_0 and the precisions τ_{s_1}, τ_{s_2}, τ_{t_1}, τ_{t_2}, $\tau_{\delta_{s,t}}$, τ_ϵ, ρ and $\rho_{\delta_{s,t}}$. Here, $\rho_{\delta_{s,t}}$ indicates the structured spatial and

temporal effect on the interaction term.

As for the final model, the fixed and random effects for the structured spatiotemporal interaction model with covariate (Model 3h) is given as follows:

$$
\begin{aligned}
&\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + \phi_{t_1} + \phi_{t_2} + \delta_{s,t}^{(i)} + \epsilon_{s,t}^{(i)}; \quad i = 1, \ldots, 47 \text{ and } t = 1, \ldots, 15 \\
&\alpha_0 \sim N(0, \sigma_{\alpha_0}^2) \\
&\log(\text{yield}_c) \sim N(0, \sigma_{\log(\text{yield}_c)}^2) \\
&\epsilon_{s,t} \sim N(0, \sigma_{\epsilon_{s,t}}^2) \\
&\gamma_{s_1} \sim N(0, \sigma_{\gamma_{s_1}}^2) \\
&\gamma_{s_2} \sim N(0, \sigma_{\gamma_{s_2}}^2) \\
&\phi_{t_1} \sim N(0, \sigma_{\phi_{t_1}}^2) \\
&\phi_{t_2}^{(1)} \sim N\left(0, \frac{1}{\tau(1-\rho)}\right) \\
&\delta_{s,t} \text{ is the random effect.}
\end{aligned} \tag{5.11}
$$

According to Equation (5.11), the summary statistics of the estimation results for the structured spatiotemporal interaction model with covariate (Model 3h) is given in Table 5.14.

Table 5.14: Summary statistics of fixed and random effects for the structured spatiotemporal interaction model with the covariate

Model 3h

$$\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \log(\text{yield}_c) + \gamma_{s1}^{(i)} + \gamma_{s2}^{(i)} + \phi_{t_1} + \phi_{t_2} + \delta_{s,t}^{(i)} + \epsilon_{s,t}^{(i)}$$

Parameter	**mean**	**sd**	**0.025quant**	**0.50quant**	**0.975quant**	**mode**
α_0	0.43	0.18	0.08	0.42	0.80	0.42
$\log(\text{yield}_c)$	0.92	0.03	0.85	0.92	0.98	0.92
Precision						
τ_{s_1}	100.20	41.20	46.66	91.25	204.70	76.58
τ_{s_2}	1,948.00	1,873.00	153.59	1,407.00	6,877.00	433.43
τ_{t_1}	21,370.00	21,370.00	2,120.29	15,150.00	78,260.00	5,964.80
τ_{t_2}	18,590.00	30,950.00	549.61	9,202.00	95,550.00	1,203.74
$\tau_{\delta_{s,t}}$	14.86	2.96	9.85	14.59	21.44	14.08
τ_ϵ	41.10	4.89	32.28	40.83	51.48	40.34
ρ	0.52	0.39	-0.44	0.62	0.98	0.94
$\rho_{\delta_{st}}$	0.67	0.07	0.52	0.68	0.81	0.68

Table 5.14 represents the descriptive statistics of the estimations of the fixed effects α_0 and $\log(\text{yield}_c)$ and the precisions τ_{s_1}, τ_{s_2}, τ_{t_1}, τ_{t_2}, $\tau_{\delta_{s,t}}$, τ_ϵ, ρ and $\rho_{\delta_{s,t}}$.

After investigating eah model separately, we compare the considered models in the following section.

5.4. Model Selection

We consider 11 models examining the district-based distribution of the wheat yield. The basic error model is considered as the first model which takes into consideration of the spatially unstructured random effect for the districts, i.e. it is assumed to be *iid*. Then, we phase in various models which consist of the spatial, temporal and spatiotemporal effects. In this section, we present the comparison criteria related to considered models given in Section 5.3. These criteria could also be used for model selection and diagnostics of the models.

We perform two criteria DIC and Lscore for the choice of the reasonable model to be used to estimate wheat yield in our case study. We choose the best model according to the smallest DIC and Lscore values. The DIC values, which is the summation of $\bar{D}$ in the first column and p_D in the second column, and Lscore values are given in the following table. The formulations of the models are also displayed in this table.

Table 5.15: $\bar{D}$, p_D , DIC and Lscore Results for Model Selection

Models	$\bar{D}$	p_D	DIC	Lscore
Model 1				$\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \gamma_{s_1}^{(i)} + \epsilon_{s,t}^{(i)}$
	336,37	36,36	372,73	0,27
Model 2a				$\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + \epsilon_{s,t}^{(i)}$
	336,38	34,96	371,34	0,26
Model 2b				$\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \log(\text{yield}_c) + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + \epsilon_{s,t}^{(i)}$
	-149,30	42,28	-107,02	-0,07
Model 3a				$\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + \left(\alpha + \delta_{s,t}^{(i)}\right) t + \epsilon_{s,t}^{(i)}$
	133,61	49,48	183,09	0,13
Model 3b				$\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \log(\text{yield}_c) + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + (\alpha + \delta_{s,t}^{(i)})t + \epsilon_{s,t}^{(i)}$
	-218,26	61,24	-157,02	-0,11
Model 3c				$\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + \phi_{t_1} + \phi_{t_2} + \epsilon_{s,t}^{(i)}$
	-84,01	54,34	-29,67	-0,02
Model 3d				$\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \log(\text{yield}_c) + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + \phi_{t_1} + \phi_{t_2} + \epsilon_{s,t}^{(i)}$
	-150,24	43,99	-106,25	-0,07
Model 3e				$\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + \phi_{t_1} + \phi_{t_2} + \delta_{s,t}^{(i)} + \epsilon_{s,t}^{(i)}$
	-582,23	371,85	-210,38	-0,02
Model 3f				$\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \log(\text{yield}_c) + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + \phi_{t_1} + \phi_{t_2} + \delta_{s,t}^{(i)} + \epsilon_{s,t}^{(i)}$
	-152,31	45,87	-106,44	-0,07
Model 3g				$\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + \phi_{t_1} + \phi_{t_2} + \delta_{s,t}^{(i)} + \epsilon_{s,t}^{(i)}$
	-648,65	306,53	-342,12	-0,16
Model 3h				$\eta_{s,t}^{(i)} = \log(Y_{s,t}^{(i)}) = \alpha_0 + \log(\text{yield}_c) + \gamma_{s_1}^{(i)} + \gamma_{s_2}^{(i)} + \phi_{t_1} + \phi_{t_2} + \delta_{s,t}^{(i)} + \epsilon_{s,t}^{(i)}$
	-614,40	270,15	**-344,25**	**-0,17**

As seen in Table 5.4, the lowest values of DIC and Lscore belong to **Model 3h**, which suggests that this model is the best model among all considered models. **Model 3h** represents the structured spatiotemporal interaction model which includes spatial, temporal and spatiotemporal effects with the covariate. Here, we also examine the PIT values related to **Model 3h** for the model adequacy test in addition to DIC and Lscore criteria. The PIT values must have uniform distribution in order to represent a good estimation of the model. The scatter plot and histogram of the PIT values are presented in the following figure.

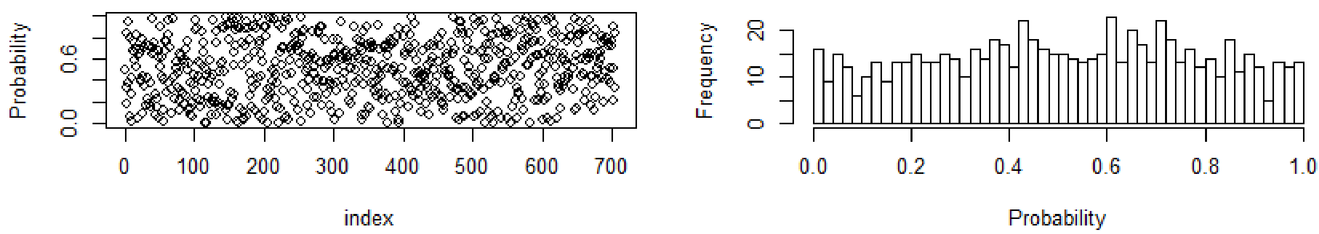

Figure 5.9: Scatter plot (left) and histogram (right) of the PIT values of the selected model

According to Figure 5.9, the values of PIT seems to be close to uniform distribution. Also, we perform goodness-of-fit test for these values for the uniform distribution (p-value $= 0,075 > 0,05$). The model results related to the fixed and random effects are given in Section 5.3. The marginal posterior of the fixed effects and the marginal posterior of the random effects are presented in Figure 5.10 and Figure 5.11, respectively.

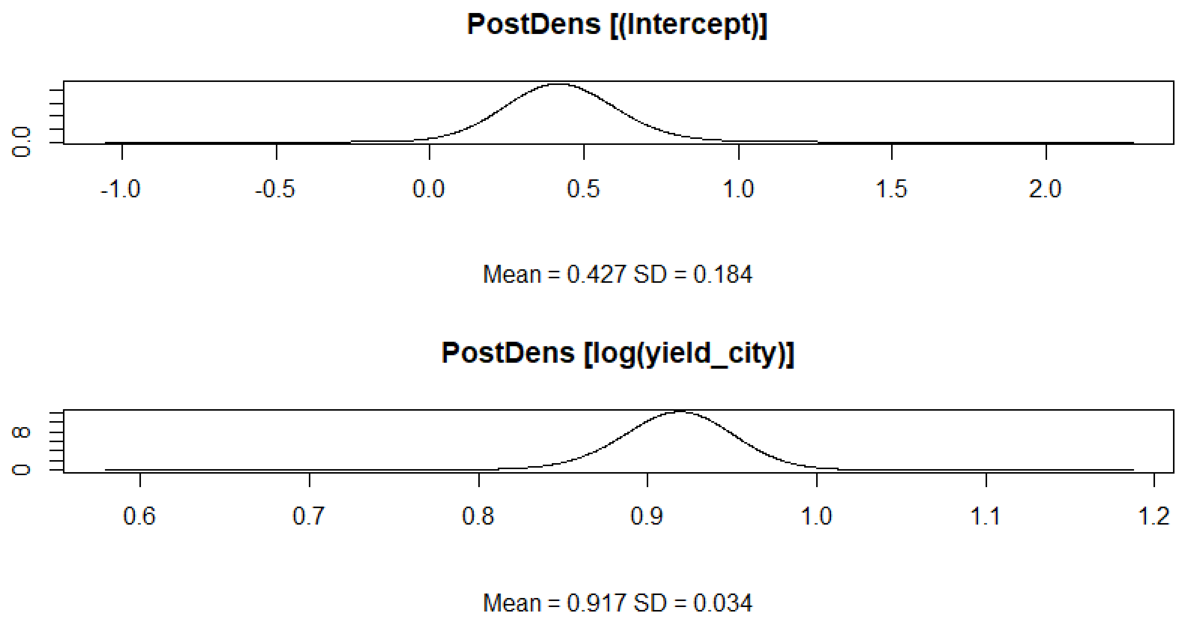

Figure 5.10: Marginal posterior of the fixed effects

According to this figure, we can infer that the normality assumptions for prior distributions are adequate.

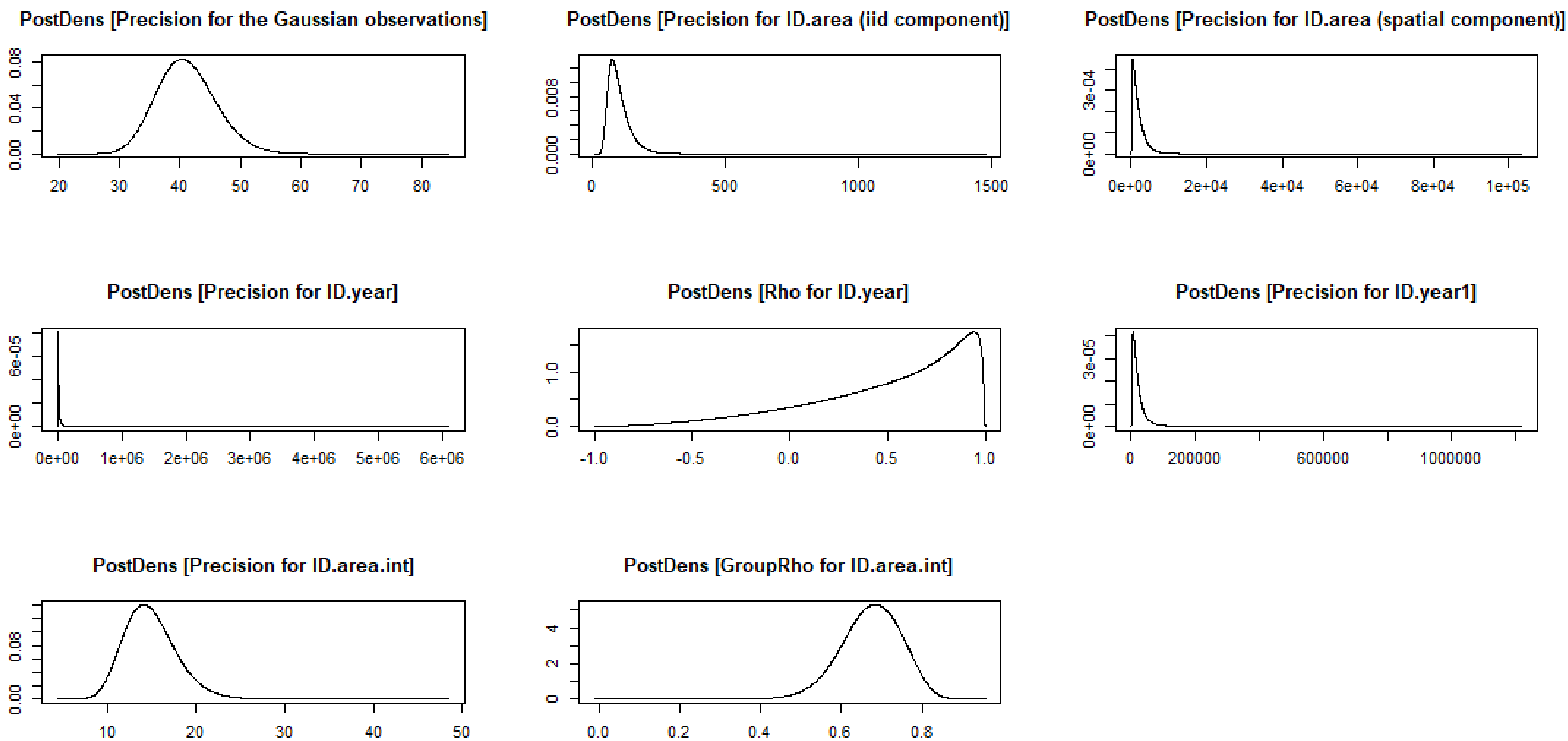

Figure 5.11: Marginal posterior of the random effects

The structured spatiotemporal interaction term $\delta_{s,t}$ for the model represents the relationship between space and time. The posterior mean of $\delta_{s,t}$ for the districts along fifteen years are given below.

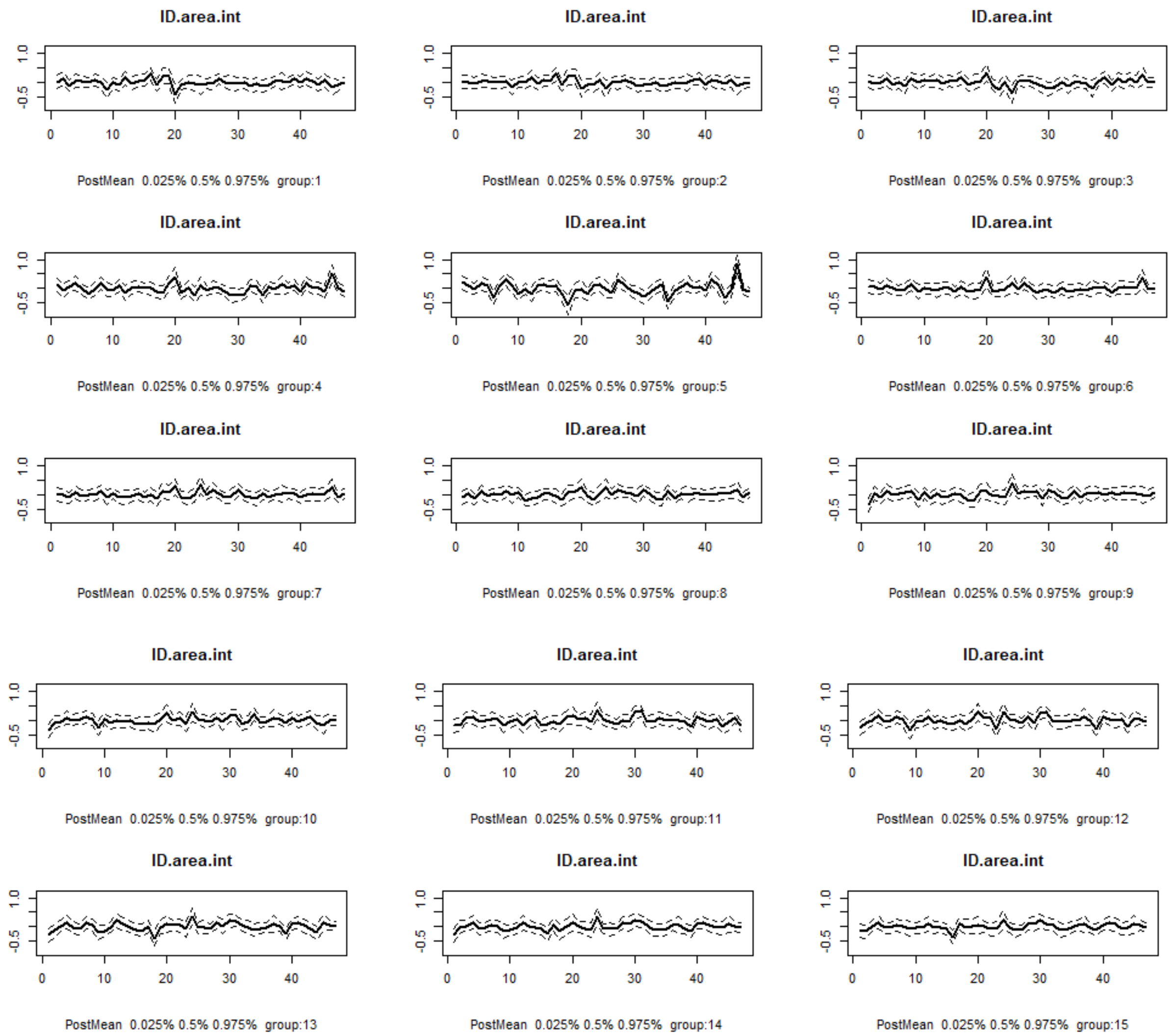

Figure 5.12: The Posterior mean of the structured spatiotemporal interaction random effect

The dashed curves in Figure 5.12 indicate the %95 confidence interval for the $\delta_{s,t}$ whereas the thick curves represent the posterior mean of $\delta_{s,t}$.

A graphical comparison of the observed and fitted values obtained by using the chosen model (Model 3h) for 15 years is provided in the following figure.

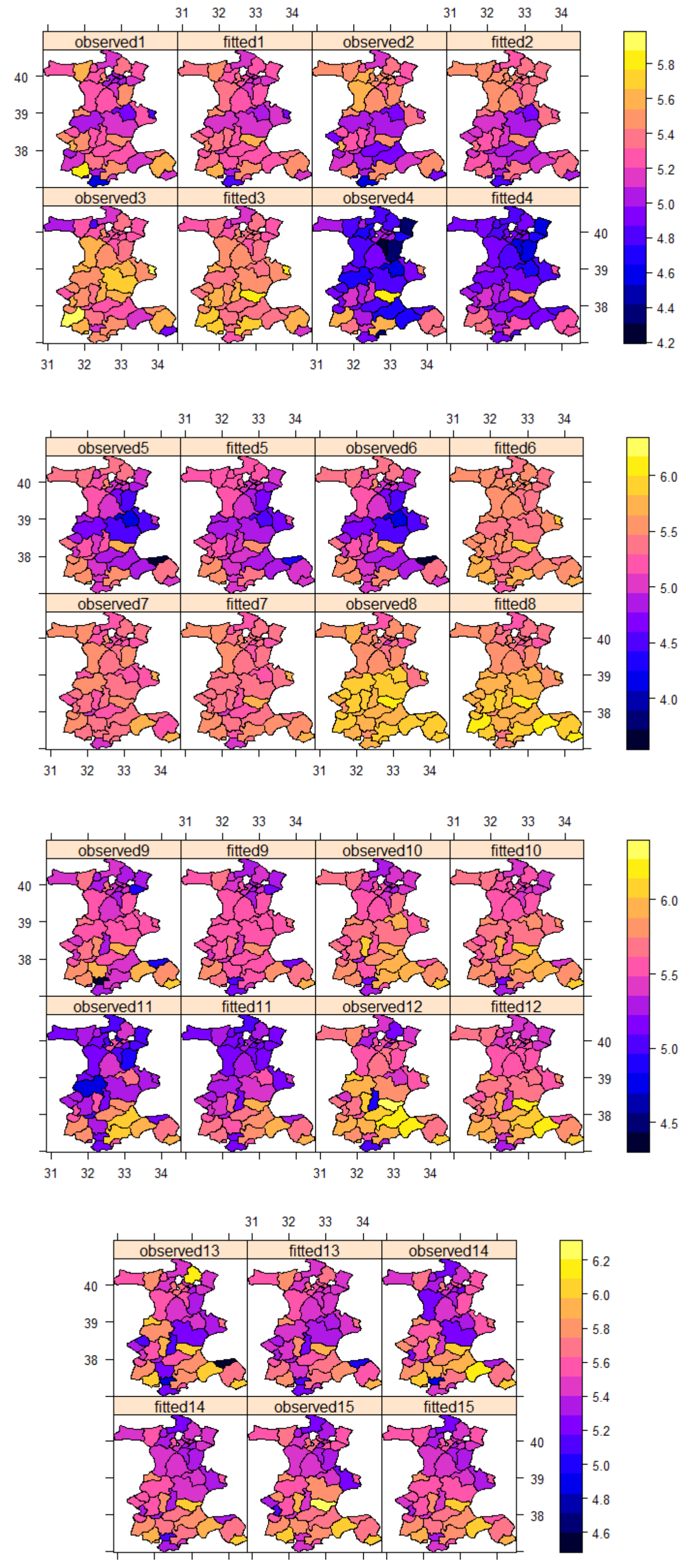

Figure 5.13: The observed and the fitted values of wheat yield according to space and time

According to Figure 5.13, it can be said that the best model fits the wheat yield observations since the color scale seems similar for 15 years.

As a risk assessment of the modelling of district-based crop yield insurance, we provide the following figure. In this figure, we represent the probability of the yield being more than a yield threshold, which is taken as the average yield for each district, for 15 years. Here, these probabilities are calculated by using the chosen model reflecting the structured spatiotemporal interaction model which includes spatial, temporal and spatiotemporal effects with the covariate.

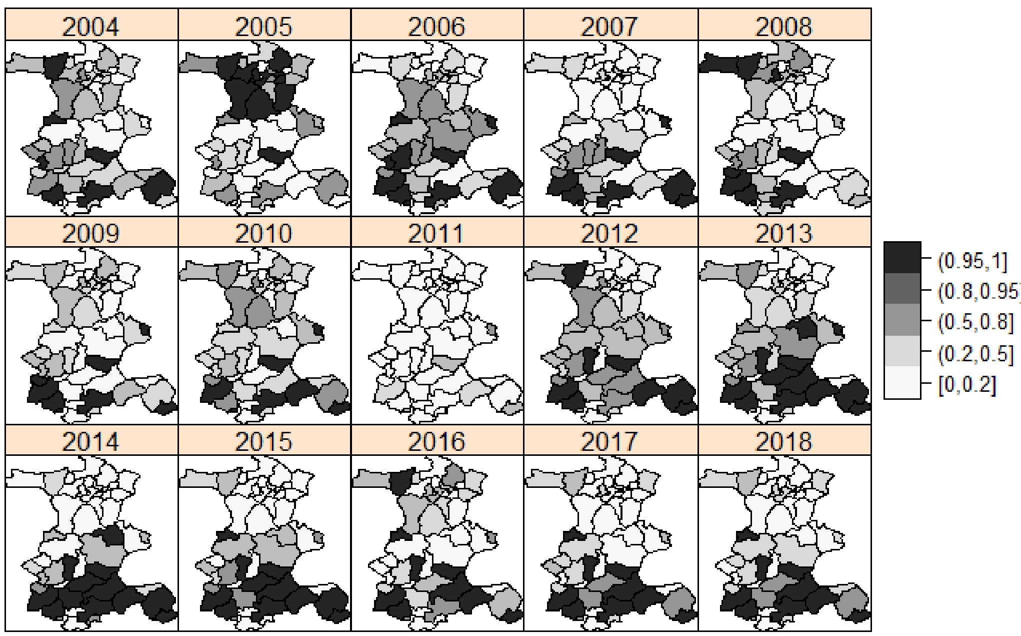

Figure 5.14: The exceedance probability of the wheat yield according to observed and estimated values of the wheat yield for the 705 observations ($705 = 47 \times 15$)

According to Figure 5.14, the districts having higher exceedance probabilities that can be seen as dark-colored regions are less risky since the estimation of the yield amount above a threshold is high. On the other hand, the districts having lower exceedance probabilities that can be seen as light-colored regions are more risky because the estimation of the yield amount above a threshold is low.

5.5. Premium Calculation of the Crop Yield Insurance

In this section, we obtain the premium rates for the crop yield insurance. Goodwin and Ker [27] propose the following equation in order to obtain the premium rate π_r.

$$\pi_r = \frac{P(y < c\bar{y})\left[c\bar{y} - E(y\,|y < c\bar{y})\right]}{c\bar{y}} \tag{5.12}$$

In this equation, $P(y < c\bar{y})$ indicates the probability that the realized yield is smaller than the prespecified threshold (yield loss level) $c\bar{y}$ where $c; 0 < c < 1$ is the coverage rate provided by the insurer and $\bar{y}$ is the expected value of the yield, i.e. $\bar{y} = \int_{-\infty}^{\infty} yf(y)dy$. The conditional expectation of realized yield given that it is smaller than the yield threshold is denoted by $E(y\,|y < c\bar{y})$. The numerator in Equation (5.12) represents the premium to be paid by the insured. The premium rate is obtained through dividing the premium by the yield threshold. The mathematical formulae for $P(y < c\bar{y})$ and $E(y\,|y < c\bar{y})$ are as follows.

$$P(y < c\bar{y}) = \int_0^{c\bar{y}} f(y)dy$$

$$E(y|y < c\bar{y}) = \frac{\int_0^{c\bar{y}} yf(y)dy}{\int_0^{c\bar{y}} f(y)dy}$$

We use numerical methods for the solution of the above integrals. Suppose that the crop yield is normally distributed. Then the probability distribution function is

$$f(y) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\dfrac{(y-\mu)^2}{2\sigma^2}}$$

where μ denotes the mean whereas σ is the standard deviation.

According to this calculation, we compute the average premium rates related to specified coverage levels $c = 0.70, 0.75, 0.80, 0.85, 0.90$ by using the selected model (Model 3h). The premium rates are obtained for all districts, city-based and the chosen districts respectively.

The following table represents the overall premium rates that is calculated by taking into account of all districts.

Table 5.16: Premium rates of all districts

Level of Coverage (%)	**Premium Rate** (%)
70	2.82
75	3.70
80	4.78
85	6.05
90	7.53

In Table 5.16, we assume that the insurance amount of a crop yield policy is the yield in kilos per decare. The premium rate is hence denoted as the percentage of the insurance amount. It is obvious that the premium rate is higher for the higher coverage levels.

Table 5.17 represents the premium rates related to Ankara and Konya that is calculated separately by taking into account of all districts of each city.

Table 5.17: Premium rates for Ankara and Konya

City	**Level of Coverage** (%)	**Premium Rate** (%)
Ankara	70	0.93
	75	1.47
	80	2.24
	85	3.27
	90	4.59
Konya	70	3.45
	75	4.40
	80	5.53
	85	6.84
	90	8.33

It is seen from Table 5.17 that the premium rates are higher in Konya in comparison with Ankara. This means that the wheat yield risk of Konya is higher than the risk of Ankara in terms of our selected model.

Lastly, we present the premium rates for 3 districts having the highest volume of wheat production in Ankara and Konya in the following table. The results are displayed separately for 3 top districts of each city.

Table 5.18: Premium rates for Ankara and Konya

District	**Level of Coverage** (%)	**Premium Rate** (%)
Polatlı	70	0.41
	75	0.76
	80	1.31
	85	2.15
	90	3.34
Haymana	70	0.57
	75	0.99
	80	1.63
	85	2.55
	90	3.79
Bala	70	1.00
	75	1.56
	80	2.35
	85	3.39
	90	4.73
Cihanbeyli	70	1.96
	75	2.73
	80	3.70
	85	4.90
	90	6.34
Karatay	70	4.78
	75	5.84
	80	7.04
	85	8.40
	90	9.90
Sarayönü	70	0.44
	75	0.79
	80	1.37
	85	2.22
	90	3.42

In Table 5.18, the first three districts, i.e. Polatlı, Haymana and Bala are districts of Ankara and they are ordered from the highest to the lowest according to the wheat production in 2018. The second three districts, i.e. Cihanbeyli, Karatay and Sarayönü are districts of Konya and they are also ordered from the highest to the lowest according to the wheat production in 2018. It is not possible to infer that the premium rates are directly related to only the volume of wheat production. Since the spatial, temporal and spatiotemporal characteristics are taken into account in the premium calculation, the relationship between the premium rates and the volume of wheat production is not the single indicator of the risk of the wheat yield. If an extreme weather event occurs in a year, the premium rates may change dramatically from that year to the following year according to our selected model.

5.6. Interim Conclusion: Modelling and Pricing in Crop Yield Insurance

Having considered spatial, temporal and spatiotemporal effects on the modelling of the crop yield, this study shows that the dependency related to both space and time effects must be taken into account for the distribution of the crop yield. Model results indicates that the structured spatiotemporal interaction model with the covariate is the best model.

In addition to the results of model performance criteria, the graphical results also show that best fitting leads us to consider the spatiotemporal dependency in addition to the spatial and temporal effects alone.

Moreover, the results of premium calculation help us to assess the risk of yield loss, i.e. the yield's being below a specified yield threshold.

6. CONCLUSION AND FURTHER STUDY

We aim to provide a summary of the main findings of our thesis as well as some suggestions for further research in this chapter. The analysis in the thesis depends on the discussion of what the insurance policy is within the economic framework. Therefore, we firstly introduce the insurance demand in a competitive market and we examine the equilibrium conditions reflecting the insureds' behaviours for a general insurance policy. Accordingly, we deal with the crop yield insurance taking into consideration of the farmers' preferences for the situation where asymmetric information exists.

One part of our study is mainly based on the impacts of the farmer's moral hazard on optimal yield insurance. For this aim, the optimal effort of the farmer is investigated by maximizing the profit of the farmer under various cases. These cases are determined according to the fact that the farmer's effort has an influence on both preventing and reducing the yield loss. For both loss-reduction and loss-prevention cases, the insurer's ability of observing the farmer's effort is handled within the framework of asymmetric information. The solutions of the optimal effort under two cases is the main contribution of this study since we utilize the farmer's optimal effort as an evidence of the non-existence of moral hazard. In addition to this, we propose a consideration of asymmetric information by considering the difference between observable and non-observable effort.

Having proposed two models, loss-prevention and loss-reduction, to analyze optimal effort under EUT, CE approach enables us to discuss these models numerically. According to various coverage rates, risk aversion coefficients, crop yield levels and effort levels, the efficiency of these models are analyzed. High-risk insureds might prefer the proposed policy based on loss-prevention model since less effort and less cost is sufficient considering the results in the numerical examples. On the other hand, loss-reduction model might be preferred as higher effort level results in a payment including both the indemnity and a bonus payment related to the farmer's effort. Consequently, higher agricultural productivity is aimed in loss-reduction model since the farmer takes a payment according to his/her yield amount regardless of whether the insured is low-risk or high-risk. These models could be proposed as an alternative to the crop yield insurance practice in Turkey as government support exists in agricultural insurance for premium payments. It could be proposed that incentive pricing models based upon farmers' effort provides agricultural and financial sustainability and protects ecosystem rather than premium subsidy applications.

The other main part of this study is the theoretical and applied analysis of the conditional yield distribution. We suggest to use hierarchical Bayesian method in order to take into consideration of not only spatial and temporal effects but also the effect of spatiotemporal dependency among geographical subregions. Model results indicates that the structured spatiotemporal interaction model with the covariate is the best model. In addition to the modelling results, we provide a methodology for the premium calculation which is a useful tool for the risk assessment of yield loss. The results of premium calculation are obtained related to the specified coverage levels by using the selected model for all districts, city-based and the chosen districts.

If we can obtain the farmer-based data related to crop yield and covariates such as demographic, socio-economic, meteorological varibles, it will be possible to extend the proposed methodology in order to estimate the farmer-based crop yield with the account of the farmer's effort by INLA model under the approaches and theoretical considerations proposed in Chapter 3.

The impacts of decision-makers' risk perceptions and spatiotemporal characteristics on the risk prioritization are investigated in agricultural insurance [68, 69]. Since the moral hazard and adverse selection are considered within the frame of preference theory, these context can be investigated by an analysis of non-observable tendencies about risk. For the future studies, it is aimed to combine these two notions, asymmetric information and risk perception, in order to obtain the optimal effort of the farmer in agricultural insurance.

Finally, we aim to estimate the farmer's yield by improving a dependent aggregate claims model under the case of the dependency between the claim frequency and the claim severity. Under the assumption that the yield is a function of the farmer's neighbourhoods and the farmer's effort, we aim to obtain a risk score related to the farmer and accordingly a moral hazard map as a future study. Therefore, we can obtain a yield insurance more efficiently as a result of farmer-based premium calculation.

REFERENCES

[1] G.A Akerlof. The market for lemons: Quality and the market mechanism. quarterly. *Journal Economics*, 84:488–500, **1970**.

[2] B. Holmstrom et al. Moral hazard and observability. *Bell journal of Economics*, 10(1):74–91, **1979**.

[3] A Raviv. The design of an optimal insurance policy. *The American Economic Review*, 69(1):84–96, **1979**.

[4] M Rothschild and J Stiglitz. Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. In *Uncertainty in economics*, pages 257–280. Elsevier, **1978**.

[5] J Stiglitz. Monopoly, non-linear pricing and imperfect information: the insurance market. *The Review of Economic Studies*, 44(3):407–430, **1977**.

[6] Jerry W Liu and Mark J Browne. First-best equilibrium in insurance markets with transaction costs and heterogeneity. *Journal of Risk and Insurance*, 74(4):739–760, **2007**.

[7] M Martin Boyer and Richard Peter. Insurance fraud in a rothschild–stiglitz world. *Journal of Risk and Insurance*, 87(1):117–142, **2020**.

[8] R. Chambers. Insurability and moral hazard in agricultural insurance markets. *American Journal of Agricultural Economics*, 71(3):604–616, **1989**.

[9] K.H Coble, T.O Knight, R.D Pope, and J.R Williams. An expected-indemnity approach to the measurement of moral hazard in crop insurance. *American journal of agricultural economics*, 79(1):216–226, **1997**.

[10] B.K Goodwin. An empirical analysis of the demand for multiple peril crop insurance. *American Journal of Agricultural Economics*, 75(2):425–434, **1993**.

[11] X. Du, J. Ifft, L. Lu, and D. Zilberman. Marketing contracts and crop insurance. *American Journal of Agricultural Economics*, 97(5):1360–1370, **2015**.

[12] Robert G Chambers and John Quiggin. Optimal producer behavior in the presence of area-yield crop insurance. *American Journal of Agricultural Economics*, 84(2):320–334, **2002**.

[13] Snaebjorn Gunnsteinsson. Experimental identification of asymmetric information: Evidence on crop insurance in the philippines. *Journal of Development Economics*, 144:102414, **2020**.

[14] V.H Smith and B.K Goodwin. Crop insurance, moral hazard, and agricultural chemical use. *American Journal of Agricultural Economics*, 78(2):428–438, **1996**.

[15] K.H Coble, R.G Heifner, and M. Zuniga. Implications of crop yield and revenue insurance for producer hedging. *Journal of Agricultural and Resource Economics*, pages 432–452, **2000**.

[16] O. Mahul and B.D Wright. Designing optimal crop revenue insurance. *American Journal of Agricultural Economics*, 85(3):580–589, **2003**.

[17] E. Ligon. Optimal risk in agricultural contracts. *Agricultural Systems*, 75(2-3):265–276, **2003**.

[18] L Eeckhoudt and C Gollier. The impact of prudence on optimal prevention. *Economic Theory*, 26(4):989–994, **2005**.

[19] J.G Jaspersen and A. Richter. The wealth effects of premium subsidies on moral hazard in insurance markets. *European Economic Review*, 77:139–153, **2015**.

[20] Michael R Carter, Francisco Galarza, and Stephen Boucher. Underwriting area-based yield insurance to crowd-in credit supply and demand. *Savings and Development*, pages 335–362, **2007**.

[21] Ernst Berg. Assessing the farm level impacts of yield and revenue insurance: an expected value-variance approach/ 10 , exploring diversity in the european agri-food system, zaragoza, spain, 28-31 august 2002. **2002**.

[22] Isaac Ehrlich and Gary S Becker. Market insurance, self-insurance, and self-protection. *Journal of political Economy*, 80(4):623–648, **1972**.

[23] Georges Dionne and Louis Eeckhoudt. Self-insurance, self-protection and increased risk aversion. *Economics Letters*, 17(1-2):39–42, **1985**.

[24] Philippe De Donder and Jean Hindriks. Adverse selection, moral hazard and propitious selection. *Journal of Risk and Uncertainty*, 38(1):73–86, **2009**.

[25] Kristin H Roll. Moral hazard: the effect of insurance on risk and efficiency. *Agricultural Economics*, 50(3):367–375, **2019**.

[26] A.P Ker and K Coble. Modeling conditional yield densities. *American Journal of Agricultural Economics*, 85(2):291–304, **2003**.

[27] B.K Goodwin and A.P Ker. Nonparametric estimation of crop yield distributions: implications for rating group-risk crop insurance contracts. *American Journal of Agricultural Economics*, 80(1):139–153, **1998**.

[28] A.P Ker and B.K Goodwin. Nonparametric estimation of crop insurance rates revisited. *American Journal of Agricultural Economics*, 82(2):463–478, **2000**.

[29] C.G Turvey and J Zhao. Parametric and non-parametric crop yield distributions and their effects on all-risk crop insurance premiums. Technical report, **1999**.

[30] Carl H Nelson and Paul V Preckel. The conditional beta distribution as a stochastic production function. *American Journal of Agricultural Economics*, 71(2):370–378, **1989**.

[31] Viswanath Tirupattur, Robert J Hauser, and Nabil M Chaherli. Crop yield and price distributional effects on revenue hedging. *OFOR Working Paper Series, no. 96-05*, **1996**.

[32] AR Jung and CA Ramezani. Valuing risk management tools as complex derivatives: An application to revenue insurance. *Journal of Financial Engineering*, 8:99–120, **1999**.

[33] Jeffrey R Stokes. A derivative security approach to setting crop revenue coverage insurance premiums. *Journal of Agricultural and Resource Economics*, pages 159–176, **2000**.

[34] V.A Ozaki, S.K Ghosh, B.K Goodwin, and R Shirota. Spatio-temporal modeling of agricultural yield data with an application to pricing crop insurance contracts. *American journal of agricultural economics*, 90(4):951–961, **2008**.

[35] P. Saengseedam and N. Kantanantha. Spatio-temporal model for crop yield forecasting. *Journal of Applied Statistics*, 44(3):427–440, **2017**.

[36] Eunchun Park, B Wade Brorsen, and Ardian Harri. Using bayesian kriging for spatial smoothing in crop insurance rating. *American Journal of Agricultural Economics*, 101(1):330–351, **2019**.

[37] H Rue, S Martino, and N Chopin. Approximate bayesian inference for latent gaussian models by using integrated nested laplace approximations. *Journal of the royal statistical society: Series b (statistical methodology)*, 71(2):319–392, **2009**.

[38] R Ruiz-Cárdenas, E.T Krainski, and H Rue. Direct fitting of dynamic models using integrated nested laplace approximations—inla. *Computational Statistics & Data Analysis*, 56(6):1808–1828, **2012**.

[39] Andrea Riebler, Leonhard Held, Håvard Rue, et al. Estimation and extrapolation of time trends in registry data—borrowing strength from related populations. *The Annals of Applied Statistics*, 6(1):304–333, **2012**.

[40] S.H S rbye, J.B Illian, D.P Simpson, D Burslem, and H Rue. Careful prior specification avoids incautious inference for log-gaussian cox point processes. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 68(3):543–564, **2019**.

[41] Ramiro Ruiz-Cárdenas and Elias Teixeira Krainski. Evaluating spatio-temporal models for crop yield forecasting using inla: implications to pricing area yield crop insurance contracts. *URL http://www. ime. unicamp. br/sinape/sites/default/files/Ruiz-Krainski_paper_SINAPE2010. pdf*, **2010**.

[42] R. Rees, A. Wambach, et al. The microeconomics of insurance. *Foundations and Trends® in Microeconomics*, 4(1–2):1–163, **2008**.

[43] Bruce J Sherrick, Peter J Barry, Paul N Ellinger, and Gary D Schnitkey. Factors influencing farmers' crop insurance decisions. *American journal of agricultural economics*, 86(1):103–114, **2004**.

[44] Yoshihiro Asai and Mahito Okura. How do cost and regulation change loss control activities and insurers' monitoring? *Journal of Insurance issues*, pages 172–188, **2011**.

[45] Zhiwei Shen and Martin Odening. Coping with systemic risk in index-based crop insurance. *Agricultural Economics*, 44(1):1–13, **2013**.

[46] Li Zhang. Three essays on agricultural risk and insurance. **2008**.

[47] Jeff Luckstead and Stephen Devadoss. Implications of commodity programs and crop insurance policies for wheat producers. *Journal of Agricultural and Applied Economics*, 51(2):267–285, **2019**.

[48] Simon Mesfin Kidane. Comparison of agricultural technologies and risk analysis of smallholder farmers in mozambique. **2015**.

[49] V.A Ozaki and R.S Silva. Bayesian ratemaking procedure of crop insurance contracts with skewed distribution. *Journal of Applied Statistics*, 36(4):443–452, **2009**.

[50] E. Park, B.W Brorsen, and A Harri. Using bayesian spatial smoothing and extreme value theory to develop area-yield crop insurance rating. Technical report, **2016**.

[51] S.N Awondo, G.S Datta, O.A Ramirez, and E.G Fonsah. Estimation of crop yield distribution and insurance premium using shrinkage estimator: A hierarchical bayes and small area estimation approach. Technical report, **2012**.

[52] J Besag, J York, and A Mollié. Bayesian image restoration, with two applications in spatial statistics. *Annals of the institute of statistical mathematics*, 43(1):1–20, **1991**.

[53] S.A.G Pereira. Spatio-temporal methods and models for unemployment estimation. **2018**.

[54] T.G Martins, D Simpson, F Lindgren, and H Rue. Bayesian computing with inla: new features. *Computational Statistics & Data Analysis*, 67:68–83, **2013**.

[55] Â.A Rodrigues. *Spatio-temporal modelling of tornados with R-INLA, at the county-level in Texas and Ocklahona*. PhD thesis, **2017**.

[56] M Blangiardo, M Cameletti, G Baio, and H Rue. Spatial and spatio-temporal models with r-inla. *Spatial and spatio-temporal epidemiology*, 4:33–49, **2013**.

[57] H Rue and S Martino. Approximate bayesian inference for hierarchical gaussian markov random field models. *Journal of statistical planning and inference*, 137(10):3177–3192, **2007**.

[58] M Blangiardo and M Cameletti. *Spatial and spatio-temporal Bayesian models with R-INLA*. John Wiley & Sons, **2015**.

[59] L Held, B Schrödle, and H Rue. Posterior and cross-validatory predictive checks: a comparison of mcmc and inla. In *Statistical modelling and regression structures*, pages 91–110. Springer, **2010**.

[60] H.S Stern and N Cressie. Posterior predictive model checks for disease mapping models. *Statistics in medicine*, 19(17-18):2377–2397, **2000**.

[61] L Pettit. The conditional predictive ordinate for the normal distribution. *Journal of the Royal Statistical Society: Series B (Methodological)*, 52(1):175–184, **1990**.

[62] A.P Dawid. Present position and potential developments: Some personal views statistical theory the prequential approach. *Journal of the Royal Statistical Society: Series A (General)*, 147(2):278–290, **1984**.

[63] M Stone. An asymptotic equivalence of choice of model by cross-validation and akaike's criterion. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1):44–47, **1977**.

[64] D.J Spiegelhalter, N.G Best, B.P Carlin, and A Van Der Linde. Bayesian measures of model complexity and fit. *Journal of the royal statistical society: Series b (statistical methodology)*, 64(4):583–639, **2002**.

[65] L Bernardinelli, D Clayton, C Pascutto, C Montomoli, M Ghislandi, and M Songini. Bayesian analysis of space—time variation in disease risk. *Statistics in medicine*, 14(21-22):2433–2443, **1995**.

[66] L Knorr-Held. Bayesian modelling of inseparable space-time variation in disease risk. *Statistics in medicine*, 19(17-18):2555–2567, **2000**.

[67] W.R Clayton, D. (1996). Generalized linear mixed models. In: Gilks, S Richardson, and D Spiegelhalter. *Markov chain Monte Carlo in practice*. Chapman and Hall/CRC, **1995**.

[68] E. Nevruz. *Multivariate Stochastic Prioritization of Dependent Actuarial Risks under Uncertainty*. PhD thesis, Department of Actuarial Sciences, Graduate School of Science and Engineering, Hacettepe University, Ankara, Turkey, **2018**.

[69] E. Nevruz and S.K. Yildirak. Spatiotemporal interpolation through an extension of differential evolution algorithm for agricultural insurance claims. *Journal of Computational and Applied Mathematics*, 352:278–292, **2019**.