SPACE-TIME MODELLING OF UNEMPLOYMENT RATE IN NAMIBIA

A THESIS

SUBMITTED IN PARTIAL FULFILMENT

OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE (IN APPLIED STATISTICS & DEMOGRAPHY)

OF

THE UNIVERSITY OF NAMIBIA

BY

FRANSINA AMUTENYA

(200818139)

APRIL 2021

SUPERVISOR: DR. ISAK NEEMA

ABSTRACT

Among major and burning issues in the developing world is the problem of unemployment. Not being employed does not only have a negative effect on the well-being of an individual, but also pose a great concern to policy makers at all levels of government. Namibia like any other developing country especially in Africa, continues to carry the burden of a high unemployment rate, varying across regions.

This study aimed at determining whether variation in unemployment rate is due to regional clustering and if covariates such as education, sex, age, population density and time have effects on the regional unemployment rate variations. To achieve these objectives, the study used a Fully Bayesian (FB) spatial smoothing approach with temporal trends to assess regional variations in the unemployment rates obtained from the Namibia Labour Force Survey data of 2014, 2016, and 2018.

Results indicated that most of the variation in the average unemployment rate was due to the temporally unstructured effect and not regional clustering, while the effect of age, sex, and population density were insignificant. Educational level on the other hand was found to have a significant negative effect on the variation of unemployment rate. Furthermore, the probability map showed a long-term increasing risk of unemployment rate in regions such as Kunene, Omaheke and Kavango West.

In conclusion, results showed that the variations in the unemployment rate in Namibia are not due to regional clustering, nor do covariates such as sex, age, population density have a significant effect on them. However, introducing the the unstructured time trend component in the model significantly explains the variation in the unemployment rate over time and space. Also, educational level has a significant negative effect on the variation of unemployment rate. Lastly, the thesis showed that the best response distribution for modelling rates and proportions is a beta distribution.

Keywords: Unemployment rate, Bayesian analysis, Space-time, Labour force, Beta distribution

LIST OF PUBLICATION(S)/COFERENCE(S) PROCEEDINGS

None.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
BYM	Besag, York and Mollie model
FB	Full Bayesian model
DIC	Deviance Information Criterion
GIS	Geographical Information System
GMRF	Gaussian Markov Random Field
GAMM	Generalized Additive Mixed Models
HCT	Human Capital Theory
ILO	International Labour Organization
INLA	Integrated Nested Laplace Approximation
LFS	Labour Force Survey
RR	Relative Risk
SPSS	Statistical Package for Social Sciences
SVAR	Structural Vector Autoregressive
MDGs	Millennium Development Goal(s)
NIDS	Namibia Intercensal & Demographic Survey
NLFS	Namibia Labour Force Survey
NPC	National Planning Commission
NSA	Namibia Statistics Agency
NSFAF	Namibia Students Financial Assistance Fund

NSS	National Statistical System
NUTS	Nomenclature of Territorial Units
UR	Unemployment Rate
VIF	Variance Inflation Factor

ACKNOWLEDGEMENTS

Firstly, I give thanks to The Lord Almighty for presenting me with the opportunity to pursue this degree.

Secondly, I owe thanks to Dr. Isak Neema for supervising this thesis, especially his contributions towards the study methodology chapter. To Dr. Dismas Ntirampeba, thank you for providing guidance on statistical programming with R-INLA. To my writing buddy and friend Angelika, I can only say thank you for the word of encouragement and phone calls as early as 03.00 am. To John and Ndilimeke, thanks for the constant reminder that I needed to finish what I started. As for my colleague Linda, thanks for availing the necessary data sources and information.

Lastly and probably most importantly, a great thanks to the Namibia Students Financial Assistance Fund (NSFAF) and the Namibia Statistics Agency (NSA) for sponsorship.

DEDICATION

I bestow this paper in loving memory of my late close friend Luamuhezi Leon Liswaniso who passed away in the early hours of 14 September 2013.

DECLARATIONS

I, Fransina Amutenya, hereby declare that this study is my own work and is a true reflection of my research, and that this work, or any part thereof has not been submitted for a degree at any other institution.

I, Fransina Amutenya, grant The University of Namibia the right to reproduce this thesis in whole or in part, in any manner or format, which The University of Namibia may deem fit.

……………………   ……………………   ……………………

Fransina Amutenya   Signature   Date

CHAPTER 1

INTRODUCTION

1.1 Background of the Study

The Namibia Labour Force Surveys define a person aged 15 years and above as employed when he/she has worked for pay, profit, or family gain for at least one hour during the last 7 days preceding the survey interview. In line with the international statistical standards, particularly the International Labour Organization (ILO), the Namibian Government's unemployment definition is characterised as either broad or strict. In exacting, the broad definition of unemployment is based on three criteria, namely being without work, being available for work, and seeking work. On the other hand, the "strict" definition of unemployment excludes from the ranks of the unemployed, those who were not actively seeking for work during the reference period (the week preceding the interview).

Since the first Namibia Labour Force Survey (NLFS) conducted in 1997, the broad unemployment rate has varied over time, as well as across regions. At initial stages, the Labour Force Surveys were conducted every 4 years namely during 2000, 2004, 2008, and 2012 periods. However, after the establishment of the Namibia Statistics Agency (NSA) through the Statistics Act, No.9 of 2011 in 2012, Labour Force Surveys (LFSs') were then conducted annually as from 2012 onwards. To date, NSA has conducted five NLFS's (2012, 2013, 2014, 2016, and 2018) of which the latter were biennial due to funding effect.

The analysis of unemployment rate presented in this study which is the number of unemployed individuals divided by the total labour force (the sum of employment and unemployment) was based on data obtained from the NLFSs produced by the NSA. The LFS data provided information on the geographical location where the person was found, and a number of demographic characteristics such as age, sex, highest level of educational attainment and relative population density of that particular region. To allow for Beta modelling, proportions of these individual characteristics were created at regional level.

From the reviewed literature (Livanos, 2009; Ashipala & Eita, 2010; Misra & Singh, 2011; Mwinga, 2012; Baah-Boateng, 2013; Bigenius, 2014; Cheema & Atta, 2014; Namibia Statistics Agency, 2015; Sam, 2016; Sunde & Akanbi, 2016), it was evident that most unemployment studies focused on the impact of these personal characteristics on unemployment, neglecting effects of temporal trends and regional heterogeneity at all or considering them only a nuisance (Muller-Fraqczek & Pietrzak, 2011). Careful incorporation and estimation of temporal and spatial effects into the unemployment models is not only an exciting element on its own, but it is also a necessity in preventing biasness in estimating the effect of personal characteristics, future employment status and unemployment benefits it has on the unemployment rates.

This study aimed at modelling unemployment rates in Namibia using the Bayesian framework taking into account the spatial aspect and temporal trends. The response variable included regional unemployment rates calculated using the broad definition instead of strict, because in many developing economies like that of Namibia, work opportunities are often limited and potential workers most especially those residing in rural areas may well give up after an unsuccessful period of looking for work (Namibia Statistics Agency, 2013). In addition, the Southern African Development Community (SADC) of which Namibia is part recommends member states to use the unemployment broad definition because it produces a better reflection of the current economic activities in the country. Contrary, Strict Unemployment is mostly used in studies that are aimed at comparing the labour markets of country-to-country (international comparisons) which is surely not the prime focus of this research.

1.2 Statement of the problem

Namibia like any other developing country especially in Africa, continues to carry the burden of a high unemployment rate. According to the Namibia Statistics Agency (2018), the unemployment rate in Namibia increased to 33.4 percent in 2018 from 27.4 percent in 2012. In principle, whether unemployment rate in the country is at 51 percent, 15 percent, 10 percent or even 5 percent, it is still unemployment which denies Namibians food on their tables. As such, even a 1 percent unemployment rate warrants a full attention from the authority.

The government of Namibia together with its developmental partners has taken and continue to take note of recent developments in the country relating to high unemployment rate, and its impact on the population. It had initiated various programmes, including rural developmental projects through the responsible line ministries such as the "Food-cash-for work; Rural Sanitation, micro-financing, Food bank etc. all geared towards the Sustainable Development Goal number one (1) and eight (8) respectively which are to "*end poverty in all its forms everywhere"* and to *"promote decent work and economic growth for all*".

Despite these efforts, such programmes are often slow to get underway and sometimes fail to accomplish the intended purpose as Namibia continues to face the challenges of high levels of poverty and inequality amongst plenty (National Planning Commission, 2017). This situation constrains economic growth and reinforces itself through high levels of unemployment. While the economy has grown substantially, this has not reduced the level unemployment in the country.

Since its establishment in 2011, the NSA has carried out five comprehensive Labour Force Surveys in 2012, 2013, 2014, 2016 and 2018 respectively. All surveys cover all aspects of people's work, including types of jobs, occupation, income and benefits from work, and as well as the education and training needed to equip them for work.

These surveys therefore provided baseline statistics for dealing with the crisis of unemployment in Namibia. However, although such data exist, it remains unexplored and often ignored in determining the bases of regional unemployment rates' variation in Namibia. Previous work in the research area of unemployment in Namibia, particularly that of Ashipala and Eita (2010); Mwinga (2012) and Sunde and Akanbi (2016) has largely focused on the causes and impact of the economic crisis on unemployment rate, but failed to identify and monitor the spatial and temporal structure of the problem, an aspect that would be beneficial when hosting labour-related discussions and providing policy recommendations to cabinet. It is therefore for this reason that this study applied a full Bayesian model approach in the quest to understanding the underlying factors contributing to the unusual regional variations of unemployment rates over time in Namibia so that appropriate and timely interventions can be initiated.

1.3 Objectives of the study

The main objective of the study was to assess regional variations of unemployment rates in Namibia for the period 2014-2018.

Specific Objectives were:

- To fit a space-time model to estimate the true posterior mean of the unemployment rates in Namibia.
- To evaluate whether variation in the odds ratio of unemployment rates in Namibia for the period of (2014-2018) was due to regional clustering, age, sex, education, population density and time.

1.4 Significance of the study

The study is significant to policy makers in the labour market for two main reasons. Firstly, the study outcome will allow for targeted interventions for the areas identified as hotspots of unemployment. Secondly, the study is aimed at assisting policy makers in understanding and tackling the causes of variations in the ever fluctuating regional Unemployment rates of Namibia for better planning and decision making.

1.5 Limitations of the Study

Three limitations were observed during the study. Firstly, for comparative purposes, the data used in this study covers only a period of three years (2014, 2016 & 2018). This is because; annual LFSs were only conducted from 2012 and the attempt to standardize earlier versions of the NLFS 1997, 2000, 2004 and 2008 data for analysis was going to have a huge effect on the estimates due to a different sampling methodology used particularly in the 2008 LFS. Also, the NLFS 2012 was excluded from the analysis because it still had the two Kavango regions combined.

Secondly, the sample design used for data collection for the period under review does not guarantee adequate coverage of all industries, because NLFS's are household-based surveys in nature and not industrially stratified. Thirdly, the NLFS coverage also omits dwelling units that are non-residential areas (such as public or school hostels, army/police barracks, etc.). Thus, household members residing in these institutions are only included if they lived in their own private accommodation or households.

1.6 Delimitation of the Study

With the NLFS data, the period of 1997-2013 was excluded from the study period because administrative boundaries have changed since 2014, thus this study had used the new administrative boundaries with the 14 regions to match with the NLFS of 2014, 2016 and 2018 respectively.

1.7 Research Ethics

The NLFS's for the years under review were conducted by the NSA under the Statistics Act, 2011 (Act No.9 of 2011), which mandates the agency among others, to constitute the central statistical authority of the country and to collect, produce, analyse, and disseminate official and other statistics in Namibia. Thus, by virtue of this Act, all information collected that could be linked to identified individuals or households were removed from the published microdata sets.

CHAPTER 2

LITERATURE REVIEW

This chapter reviews relevant statistical methods used in Bayesian analysis, model selection criteria, parameter estimation methods as well as a review of relevant literature on modeling unemployment.

2.1 Review of Statistical Models

In this section, the study reviews two types of models commonly used in disease mapping namely, spatial models as well as the spatio-temporal models available from the literature at the time of the study.

2.1.1 Spatial Models

This sub-section reviews the spatial model developed by Besag York and Mollie (1991).

2.1.1.1 The BYM Model

Besag, York and Mollie (1991), proposed a convolution Bayesian model which controls for variation in estimating mortality risks in disease mapping (Aregay, Lawson, Faes & Kirby,2017). This was supported by Tu and Greenwood (2012) who stated that Bayesian methods were proposed to deal with sparse data arising, for instance, through small incidence or mortality rates within the context of an ecological analysis. Further, Tu and Greenwood (2012) emphasized that this approach improves the precision and stability of risk estimates. In the same vein, the method also provides a framework to model simultaneously the spatial and non-spatial (or heterogeneity) effects on disease risk.

The model split information on area specific relative risk into two groups of random effects. The first group represents spatial dependence heterogeneity or clustering heterogeneity and the second group consists of uncorrelated heterogeneity. The spatial dependence effects are responsible for controlling of the spatial auto-correlation that may be present in the data, whereas the unstructured heterogeneity assumes that estimates are independent of each other.

For instance, let S be the study area of interest which is grouped into geographical boundaries or areas. Assuming the response variable follows a Poisson distribution (count data), such that together the observed cases, Y_i and the expected cases E_i for each area *i*, (for *i*=1,2,…s). Let X_i be an explanatory variable of interest. Then the notation of the Poisson model will follow this form:

$$log(Ri) = log(\beta_0) + \beta X_i \qquad [2.1]$$

Where

R_i is the relative risk in area i

$log\ \beta_0$ is the overall level of the relative risk of mortality

and β is a parameter coefficient

It is evident that the model above does not include any clustering or unstructured heterogeneity. These two variabilities need to be accounted for, thus the BYM was proposed as follows:

$$log(R_i\ ) = log(\beta_0) + \ \beta X_i \ \ + U_i + V_i \qquad [2.2]$$

Where

U_i models the spatially structured area specific random effects based on the Conditional Autoregressive approach (CAR)

V_i models the unstructured random effects also known as heterogeneous effects.

The canonical link function log (R_i) in the equation 2.2 contains area-level error terms which are responsible for spatial variability or dependence.

In addition, the BYM is the benchmark parametric model and it is widely used in disease mapping studies due to its flexibility on residuals. For real-life applications refer to Besag, York and Mollie (1991) as they provided two examples: one for mapping risk from multiple myeloma in 94 departments of mainland France and the second for cancers in 1218 wards in the North England.

2.1.2 Spatio-temporal models

This sub-section outlined a general framework for spatio-temporal models, breaking them up into 3 stages: the *probabilistic model for observations*, the *components of the linear predictor*, and the *structures of the effects*. For each stage, the most commonly used alternatives in the literature are discussed. Lastly, the actual review is presented, classifying each model according to the *structure* of the temporal trends that may arise and discussing the relative *advantages* and *disadvantages* of the different approaches.

2.1.2.1 General Framework and Notation

Throughout the review of this section (spatio-temporal models), let us assume that the area under study is divided into I regions indexed by $i = 1, \ldots,$ I. The temporal dimension will be indexed by $j = 1, \ldots,$ J representing each period of time under study, usually years. Let n_{ij} denote the number of persons–time at risk in area i and period j, and y_{ij} the corresponding observed cases. In some cases, an additional categorization can be used apart from region and period; for instance, when considering one or more covariates such as *age*, *race*, *sex* or risk factors. In these cases, an additional sub index k = 1, . . ., K will identify each combination of existing categories. For example, if s*ex* and *race* are used as covariates, with levels {male, female} and levels {black, white} respectively, then there would be four group, namely: black males, black females, white males and white females.

2.1.2.1.1 Probabilistic Model for Observations

Conceptually, the observations y_{ijk} are assumed to be a conditionally independent random sample from a given probability distribution from the exponential family (Lopez-Quilez & Munoz, 2009). Typically, a Poisson distribution will be preferred when the observed values in each region and period are expected to be low. With some non-rare diseases, a Binomial distribution could be more appropriate (Knorr-Held and Besag, 1998). In any case, the observed data y_{ijk} depends mainly on the number of persons–time at risk, n_{ijk}.

For this reason, the expected value for y_{ijk} is factorized as $n_{ijk} \cdot r_{ijk}$, where r_{ijk} denotes the risk in region i, period j and group k. Commonly, the number of persons–time at risk is standardized by *age* such that, for each combination of *region*, *period* and *risk,* the expected number of cases E_{ijk} is computed, and the relative risk is modelled (Lopez-Quilez & Munoz, 2009).

2.1.2.1.2 Components of the linear predictor, η_{ijk}

Depending on the model chosen for observations, the logarithm or the log-odds of the risk r_{ijk} is called the linear predictor η_{ijk}. According to Knorr-Held and Besag (1998), the linear predictor is usually expressed additively as the sum of some components or effects that can be interpreted as individual and independent contributions to the risk in that region and period. Therefore, the linear predictor may have all of the following terms or, more commonly, a subset of them:

$$\eta_{ijk} = \text{Intercept} + C_k + S_i + T_j + CS_{ik} + CT_{jk} + ST_{ij} + CST_{ijk} + \varepsilon_{ijk} \qquad [2.3]$$

It is worth mentioning that the contribution of any of these given terms can increase or decrease the risk, and that the structure (notations) of the effects in equation (2.3) will be discussed in the next sub-section i.e. 2.1.2.1.3. For now, the study briefly describes each one of the possible effects as mostly defined by Knorr-Held and Besag (1998); Lopez-Quilez and Munoz (2009) below.

Intercept: The intercept term gives a starting amount of risk that is shared by all regions, periods and groups. It is usually included in every model, but is sometimes included within the mean value of a random effect, in one of the following terms:

Main effects C_k , S_i and T_j: These components represent the additional risk of belonging to group *k*, living in region *i* and period *j* respectively.

Second order interaction terms CS_{ik} , CT_{jk} and ST_{ij}: These components represent the contribution to the risk due to a combination of the effects that cannot be explained additively by the main effects.

Third order interaction term CST_{ijk}: This effect represents an additional risk affecting one specific group of people *k*, living in a specific region *i* and in a specific period of time *j*. According to Lopez-Quilez and Munoz (2009), this component is rarely used because it greatly increases the complexity of the model.

Extra variability term ε_{ijk}: Frequently, an explicit unstructured extra variability term is included in order to capture the overall effect of other minor factors. It is implemented as a white noise random effect, and can introduce noise either globally or into a specific subspace. Whereby,

$$\varepsilon_d | \lambda_\varepsilon \sim \text{No}\,(0, \lambda_\varepsilon), \qquad\qquad d \in \{i, j, k, ij, ik, jk, ijk\} \quad [2.4]$$

2.1.2.1.3 Structures of the effects

Covariates effect C_k: Usually the effect of the covariates is expressed as a linear model on them, or on a transformation of them. Lopez-Quilez and Munoz (2009) stated that, depending on the type of covariate, it can also be stratified into several categories and included as a fixed effect or as a structured random effect. A special case is the modelling of the age effect as a covariate, instead of standardizing by age. Usually age is stratified into age groups and included as a fixed effect.

Spatial effect S_i: When the region-specific data are scarce, the classical fixed effects model with maximum-likelihood estimation often leads to unsatisfactory estimates of the spatial effects in each area (Knorr-Held & Besag, (1998). In disease mapping, Lagazio, Biggeri and Dreassi (2003) stated that this problem has been overcome by a Bayesian approach which models the spatial effects as *random effects*, through a prior distribution. Specifically, the spatial effects may give rise to a *spatially unstructured* variation (*heterogeneity*) or to a *spatially structured variation* (*clustering*). In the model for *unstructured heterogeneity*, the spatial effects ϕ_i are assumed to be sampled from a normal distribution with mean 0 and precision λ_ϕ , such that,

$$S_i = \phi_i \mid \lambda_\phi \sim \text{No}\,(0, \lambda_\phi), \qquad\qquad \text{(heterogeneity effect) [2.5]}$$

In the *clustering* model, the mean of the structured effect θ_i is allowed to depend on the neighbouring θ_{j^s} through the Gaussian Conditionally Autoregressive (CAR) distribution (see, for example, Besag and Kooperberg, 1995). Formally, the joint distribution of the vector $\theta = (\theta_1, \theta_2, \ldots, \theta_I)'$ is denoted by,

$$S_i = \theta \mid \lambda_\theta \sim \mathrm{CAR}(\lambda_\theta), \qquad \text{(clustering effect) [2.6]}$$

The choice between the *clustering* and the *heterogeneity* model depends upon our prior belief about the scope of dominant risk determinants (Lopez-Quilez & Munoz, 2009). Risk determinants exceeding the limits of one or more regions leads to a *clustering* model since they induce similar risk values in neighbouring regions. On the contrary, when the scope of risk determinants is smaller than a region's size it leads to a *heterogeneity* model.

Among all the proposals for performing risk smoothing which have appeared in the literature, the one stated by Besag, York and Mollie (1991) has had a particular impact. Most of the models reviewed in the upcoming sub-section 2.1.2.2 based their spatial modelling on that approach, which this study will refer to as the *BYM specification*. The risk associated with a region is broken down as the sum of a *heterogeneity* and a *clustering* effect,

$$S_i = \phi_i + \theta_i \qquad \text{(BYM specification) [2.7]}$$

where vectors ϕ and θ are distributed as in equations 2.5 and 2.6 above.

Temporal effect T_j: In contrast to restrictive evolution models such as linear or polynomial parametric models, most of the time, a smooth and flexible evolution is preferred (Lagazio, Biggeri & Dreassi, 2003). So, it is frequent to model this term as a structured random effect, ensuring that contiguous periods are likely to be similar, but allowing for flexible shapes in the evolution curve, especially when long periods of time are being considered. First or second order random walks (RW1, RW2), autoregressive processes (AR), or splines are the models that have been used within the reviewed papers. Sometimes time has also been stratified into a few blocks of time and modelled as a fixed effect, thus estimating the effect of each block independently from the others (Lopez-Quilez & Munoz, 2009).

Covariate interactions CS_{ik} , CT_{jk} and CST_{ijk} : These are rare effects according to Lopez-Quilez and Munoz (2009). It is more usual to assume independence between the covariates and the spatial and temporal effects. An example of such an interaction can be found in the paper by Sun, Tsutakawa, Kim and He (2000). These authors stratified *age* into four groups and assumed that each age-group could present a different evolution pattern in mortality rates. So, they modelled this interaction as a linear parametric function of time, and found that mortality rates showed a decreasing trend for the youngest group but an increasing trend for the other three over the period under study.

Spatio-temporal interaction ST_{ij} : This is the key aspect of spatio-temporal models, and possibly the most difficult one, because of the many possibilities and the lack of an accepted standard that functions well (Lopez-Quilez & Munoz, 2009). Knorr-Held (2000) established a classification of four possible types of interaction (Type I, Type II, Type III and Type IV) between spatial and temporal random effects. However, in this study the author presented a generalization, using the same abstract types of interaction but not restricted to any specific form of spatial and temporal effects. This enabled the researcher to classify all of the reviewed papers in sub-section 2.1.2.2

into one of these types, according to conceptual aspects –instead of technical– of the spatio-temporal modelling.

i) **Type I interaction:** This is some independent unobserved covariates for each combination of region and period (*i, j*), thus without any structure (Knorr-Held, 2000). However, note that if spatial and temporal main effects are present in the model, then this interaction effect only implies independence in the *deviations* from them. Contribution to risk in neighbouring regions or in consecutive periods of time can still be highly correlated, due to the main effects. This is a global space-time heterogeneity effect and is usually modelled as white noise. Moreover, it is a simple way of implementing a spatio-temporal interaction, allowing data to show if there is anything worthy of further investigation.

ii) **Type II interaction:** Here each region has a specific evolution structure that is independent of that in the neighbouring regions (Blangiardo & Cameletti, 2015). The evolution structure for each region may have as many forms as the temporal main effect itself. In the same way as before, this does not mean that each region has an evolution independent of the neighbouring ones, since they may share a common temporal main effect. Independence only affects the deviation from the global trend. According to Lopez-Quilez and Munoz (2009), this is suited to modelling factors affecting specific regions and inducing deviations from the global trend. It was the type of interaction selected in the example studied in Knorr-Held (2000) because it provided a good balance between fit and complexity.

iii) Type III interaction: The interaction can be assumed to have a spatial structure for each period, independent of adjacent periods like the spatial clustering effect, this is typically modelled with a CAR distribution for each period (Blangiardo & Cameletti, 2015). For this interaction, the inclusion of an additional heterogeneity spatial term (such as in the BYM specification) would produce a Type I interaction term. Implicitly, here it is assumed that each specific region may have a slight deviation from the global trend, but that this deviation is likely to be similar to that in the neighbouring regions while, at the same time, independent of that in the previous or subsequent period of time. Lopez-Quilez and Munoz (2009) mentioned that such an interaction could represent situations where an unobserved regional factor is affecting an area containing two or more adjacent regions, but not persistent in time.

iv) Type IV interaction: From a theoretical point of view, the most interesting form of interaction according to Lopez-Quilez and Munoz (2009) arises when deviations from the global trends are assumed to be correlated with their neighbours both in space and time. This can model hidden factors whose effects exceed the limits of one or more regions and also persist for more than one period of time. They further stated that, this type of interaction is also the most efficient way of extracting information from data, especially in the case of rare-diseases or less populated regions, since the risk estimation for a region-period is performed not only based on the locally observed data but also on that in neighbouring regions and periods.

2.1.2.2. Spatio-temporal models in the literature

In this sub-section, the study presents a brief review of the most prominent papers in the spatio-temporal modelling of diseases. The focus is kept on the spatio-temporal modelling approach, keeping to one side other technical aspects like the centering of the covariates, or whether a Poisson or a Binomial distribution is preferred for the observations, or the prior specification for hyper-parameters, etc. Notation for all models reviewed was homogenized to that established in all sub-sections under 2.1.2.1 above. Models were classified into three categories according to the structure of the temporal evolution of the estimated risk for each area. Namely, *parametric models* have a predefined shape (linear, quadratic, ...), while on the contrary, *temporally independent models* estimate the risks for each period independently of those from previous periods, and finally *smooth temporal evolution models* allow for structured trends without restricting to any predefined shape (Lopez-Quilez & Munoz, 2009).

2.1.2.2.1 Parametric models

In brief, the log risk in a given region of a parametric model is assumed to be a linear or quadratic function of time. The coefficients in the function are region-specific and are spatially structured so that neighbouring regions have similar evolutions. The advantage of these type of models (parametric) according to Bernardinelli, Clayton, Pascutto, Montomoli, Ghislandi, and Songini (1995) is that information is shared in both space and time and that the parametric formulation is straightforward. The only disadvantage is that parametric evolution in time seems to be inappropriate for longer periods of time, as it is too restrictive.

As examples, Bernardinelli, Clayton, Pascutto, Montomoli, Ghislandi, and Songini (1995) as well as Assuncao, Reis and Oliveira (2001) proposed the following parametric model:

$$\eta_{ij} = \text{Intercept} + S_i + T_j + ST_{ij} \qquad [2.8]$$

where S_i is modelled as a *heterogeneity* (2.5) or a *clustering* (2.6) effect depending on the problem, $T_j = \beta t_j$ and $ST_{ij} = \delta_i t_j$ where δ_i is again a heterogeneity or a clustering effect.

In this way, the log risk is a linear function of time, with region-specific intercepts and slopes. When δ_i is a heterogeneity effect the model has an interaction of Type II, but of Type IV if δ_i is a clustering effect. The authors of the model also proposed an extension that allows for the *a* priori correlation between the spatial and spatio-temporal random effects, using an additional level in the hierarchical model. They performed Bayesian inference using MCMC methods.

Assuncao, Reis and Oliveira (2001) followed the same line, incorporating an additional quadratic term in time that allows for curved trends with convex or concave shapes. Random effects were modelled as *clustering* effects (2.6), thus producing spatio-temporal interactions of Type IV.

Sun, Tsutakawa, Kim and He (2000) expanded equation 2.8 by incorporating the *age group* as a fixed-effect covariate C_k, and model the spatial effect S_i with a *clustering* effect (2.6). After the adjustments the model took the form:

$$\eta_{ijk} = \text{Intercept} + C_k + S_i + CT_{jk} + ST_{ij} + \varepsilon_{ijk} \qquad [2.9]$$

The *covariate-time* and *space-time* interaction terms were modelled as linear functions of time, with a slope depending on the *age group* and region respectively: $CT_{jk} = age_k t_j$, and $ST_{ij} = \theta_i t_j$. Furthermore, while age_k is a fixed effect, θ_i is a spatial clustering effect. Hence, the spatio-temporal interaction is of Type IV. Finally, an overall *heterogeneity* term ε_{ijk} (2.4) was added.

2.1.2.2.2. Temporally independent spatial models

In summary, temporally independent spatial models can be seen simply as a set of spatial models, one for each period of time, with almost no relation between them, except possibly for some restrictions in their precision parameters (Lopez-Quilez & Munoz, 2009). The advantage of these type of models is that Temporal evolution is not restricted to any specific shape, information is shared in space. The only disadvantage is that for each period of time one spatial model is estimated, therefore not sharing information in time.

As examples, Waller, Carlin, Xia and Gelfand (1997) discussed the inclusion of spatial and temporal main effects together with *sex* and *race* covariates into a temporally independent spatial model to take the following form:

$$\eta_{ijk} = \text{Intercept} + C_k + ST_{ij} \qquad [2.10]$$

Initially, the authors assumed additivity, meaning that *sex* and *race* effects are not affected by region and year, therefore there are no CS_{ik} or CT_{jk} interaction terms. The covariate component C_k is that of a linear model with *race* and *sex* as fixed effects plus a *sex-race* interaction effect. The spatio-temporal interaction term has the form $ST_{ij} = \phi^j{}_i + \theta^j{}_i$, where for each period of time

j, the vector ϕ^j and θ^j follows the BYM specification (2.7), with different precision parameters $\lambda_\phi{}^j$ and $\lambda_\theta{}^j$ for each period of time, therefore yielding an interaction of a Type III.

Xia and Carlin (1998) on the other hand, extended the model by Waller, Carlin , Xia and Gelfand (1997) by introducing another covariate (*smoking prevalence*) in a more sophisticated way, namely introducing a way to account for errors in covariates.

In addition to the above, Nobre, Schmidt and Lopes (2005) followed the same approach and refined the model in equation (2.10), by fitting a spatial model nested within time with some modifications and took the form:

$$\eta_{ijk} = CT_{jk} + ST_{ij} \qquad [2.11]$$

The authors generalized the covariate effect, allowing its coefficient to evolve over time through a RW1 prior. Regarding the spatio-temporal modelling, $ST_{ij} = \delta^{(j)}{}_i$, where, for each period j, $\delta^{(j)}{}_i$ is a *heterogeneity* effect (2.5) or a *clustering* effect (2.6), but not both, therefore representing spatio-temporal interactions of Types II or IV respectively.

2.1.2.2.3 Smooth temporal evolution models

When it comes to smooth temporal evolution models, the progression of the estimated risk in each region according to Lopez-Quilez and Munoz (2009) is a smooth function of time with an advantage that the temporal evolution is not restricted to any predefined shape. Information is actually shared in time. The only setback with this type of models is that there are many possible alternatives for the interaction term, which may necessitate model selection criteria.

As examples, Knorr-Held (2000) considered a spatio-temporal model of the form,

$$\eta_{ijk} = \text{Intercept} + S_i + T_j + ST_{ij}, \qquad [2.12]$$

where the spatial main effect S_i follows the BYM specification (2.7), and the temporal main effect T_j is specified as the sum of a RW1 structured effect and an unstructured random effect. Knorr-Held (2000) tried five different alternatives for the spatio-temporal interaction ST_{ij} , the first of which was not having any interaction term at all. The interaction terms of the other four alternatives were modelled as random effects with precision matrices λK, where λ is an unknown scalar to be estimated from the data and *K* is a precision matrix computed as the Kronecker product of the structure matrices of either the structured or unstructured spatial effect with either the structured or unstructured temporal effect, following a rationale by Gilks (2005). In this way, if both unstructured effects are combined then a Type I interaction results, while on the contrary the combination of both structured effects produces an interaction of Type IV.

Lagazio, Biggeri and Dreassi (2003) extended the Knorr-Held (2000) model (e.q 2.12) by turning it into an *Age-Period-Cohort* model. While keeping the BYM specification for the spatial effect S_i, the temporal effect T_j is split into three random effects corresponding to *age group*, *calendar period* and *birth cohort*, all of them with first order random walk priors. Regarding the interaction term, they considered interactions between the spatial *clustering* effect with both period and cohort effects. Hence, two possible interactions of Type IV. They performed model comparison based on Expected Predictive Deviance, and also by measuring the differences based on the Kullback-

Leibler divergence. Finally, they selected the model with the three temporal main effects plus space-cohort interaction, which seemed to play a very important role.

Schmid and Held (2004) presented the same model as Lagazio, Biggeri and Dreassi (2003), with slight modifications. Namely, the observations are assumed to be drawn from a binomial distribution, the spatial effect is modelled as a clustering effect (4), and a heterogeneity effect scoping age, period and space dimensions is added. They investigated whether a period-space interaction, a cohort-space interaction, or no interaction at all is more appropriate for their data. In either case of spatio-temporal interaction, they also determined whether a Type II or a Type IV interaction produces the best results, and finally whether a first or second order random walk prior was more adequate for the temporal main effects.

Their computations were based on simulation. Finally, they showed how to make future predictions from these models. The authors used the DIC to compare the 10 models that arose from the alternative specifications and found almost no differences between RW1 and RW2 priors for the temporal main effects. They also found Type II interaction giving the best result.

MacNab and Dean (2001) and MacNab and Dean (2002) followed an empirical Bayes approach, using Generalized Additive Mixed Models (GAMM). Specifically, they modelled the linear predictor as having both temporal and spatial main effects and an interaction term; and a second model incorporating the effect of *age*, without any interaction with other terms. The brackets in their equation (2.13) below indicated that the term was present in the second model only.

$$\eta_{ijk} = \text{Intercept} + [C_k] + S_i + T_j + ST_{ij} \qquad [2.13]$$

The temporal main effect was modelled with a cubic B-spline, allowing for a smooth and quite flexible evolution. For the spatial main effect, the authors used a *clustering* effect (2.6).

Lastly, all the spatio-temporal models reviewed above focused on a single response/outcome of interest. However, when two or more phenomena (outcomes) are related, better results typically arise from a joint study than from an individual study. Richardson, Abellan and Best (2006) have treated *male* and *female* lung cancer as two related diseases/outcomes and developed a methodology to study their spatio-temporal behaviour jointly. Instead of considering *sex* as a covariate, they considered *male* and *female* lung cancer as two different but correlated diseases. For each disease they fit a spatio-temporal model, having one common component and one specific component that calibrates the differential between the two diseases. However, the spatio-temporal behaviour of each disease follows a rather simple model:

$$\eta_{ijk} = \text{Intercept} + S_i + T_j + ST_{ij} \qquad [2.14]$$

Where S_i is a *clustering* effect, T_j is a RW1, and ST_{ij} is a global heterogeneity effect, thus producing an interaction of Type I.

2.2 Review of Model Selection criteria

Model selection criteria are important application tools in Statistical modelling. In an ideal situation, several possible candidate models will be fitted to the data, however finding the model that best fits the data among these possible candidate models is not always easy. Questions like which main effects to include in the model and what possible interactions that contributes significantly to the response, becomes evident. The model selection thus seeks to simplify this task by identifying those candidate models that are reasonably simpler or unnecessarily complex to accommodate the data

There are many model selection criteria used in statistics. Sclove (2011) undertook a review of statistical model selection criteria, with special emphasis on the application to prediction in regression; histograms and finite mixture models. In particular, he discussed the following selection criteria: the Bayesian Information Criterion (BIC), Akaike's Information Criterion (AIC) as well as the Kashyap's Information Criterion (KIC).

Due to the prime focus of this study i.e. Bayesian analysis, the study briefly discussed the three penalized likelihood model-selection criteria that are used within the framework of the maximum likelihood method namely the BIC, AIC and DIC (Deviance Information Criterion). Ntzoufras (2009) stated that these are the three most popular information criteria used in Bayesian statistics, hence discussed below.

2.2.1 Bayesian Information Criterion (BIC)

The BIC also sometimes referred to as the Schwarz Information Criterion (SIC) or Schwarz BIC was proposed by Schwarz (1978). It is a model selection criterion that is based on the empirical log-likelihood. According to Sclove (2011), the model to be fitted would not require the specifications of priors, thus favoured in situations where the priors are often difficult to set. The BIC is therefore computed using the following equation:

$$BIC = -2\ln(L) + k ln(n) \,, \qquad [2.15]$$

where

L represents the maximum value of the likelihood function of the model, k gives the number of free parameters to be estimated in the model, while n represent the number of observations, or the sample size for the data modelled.

For the candidate models being fitted, the best model is thus the candidate model that best minimizes the BIC.

2.2.2 Akaike Information Criterion (AIC)

Just like the BIC, the AIC is a model selection criterion that has its roots on the likelihood functions. The criterion is calculated using the following equation:

$$AIC = 2k - 2\ln(L) \,, \qquad [2.16]$$

where k represents the number of parameters in the statistical model, and L is the maximized value of the likelihood function for the estimated model. The term $2k$ represents the size of the penalty to be paid for over fitting and this discourages adding too many variables in the models which always leads to a smaller likelihood. The candidate model that best minimizes the value of the AIC is deemed to be the best fit.

2.2.3 Deviance Information Criterion (DIC)

The DIC is a hierarchical modelling generalization of the AIC and the BIC. It is particularly useful in Bayesian model selection problems where the posterior distributions of the models have been obtained by Markov chain Monte Carlo (MCMC) simulation. The advantage of DIC over other criteria in the case of Bayesian model selection is that the DIC is easily calculated from the samples generated by a MCMC simulation. AIC and BIC require calculating the likelihood at its maximum over, which is often not readily available from the MCMC simulation.

The deviance is given as:

$$D(\theta) = -2\log\big(p(y|\theta)\big) + C \quad [2.17]$$

Where y are the data, θ are the unknown parameters of the model and $p(y|\theta)$ is the likelihood function. C is a constant that cancels out in all calculations that compare different models, and which therefore does not need to be known.

The effective number of parameters of the model is computed as:

$$pD = \overline{D} - D(\bar{\theta}) \quad [2.18]$$

Where $\bar{\theta}$ is the expectation of θ and pD is the effective number of parameters in the model.

The DIC is thus calculated using the following equation:

$$DIC = pD + \overline{D} \quad [2.19]$$

As rule of thumb, a model with smaller DIC is preferred over models with larger DICs.

2.3 Review of parameter estimation methods

Amongst the possible candidate models, once the best model fit is selected using the selection criteria discussed above, the next stage is to fully estimate the model parameters. In a Bayesian framework, this is done in terms of assigning the likelihood and posterior distributions. For the model given in 3.14 of Chapter 3, the likelihood function is given by:

$$L\left(y_{rt} \middle| \beta_0,\ldots,\beta_r, u_{rt}, v_{rt}, tt_t\right) \qquad [2.20]$$

and the log – likelihood would then be denoted by $\log(L(y|\theta)) = l(y|\theta)$ where now

$$\theta = (\beta_0,\ldots,\beta_r, u_{rt}, v_{rt}, tt_{rt}) \qquad [2.21]$$

The product of the likelihood and the prior distributions is called the posterior distribution. According to Lawson, Browne and Rodeiro (2003), the posterior distribution describes the behaviour of the parameters after the data are observed and prior assumptions are made. In this case the posterior distribution takes the form

$$p(\theta|y) \;\; \alpha \;\; L(y|\theta)g(\theta), \qquad [2.22]$$

where $g(\theta)$ is the joint prior distribution of the vector θ.

Therefore, for the estimation of the model parameters; prior and hyper-prior distribution must be assigned. In principle, the complexity of the posterior distribution of parameters that results from hierarchical levels of models often requires the use of advanced sampling algorithms. In this thesis, two major classes, namely Markov Chain Monte Carlo (MCMC) and Integrated Nested Laplace Approximation (INLA) methods are reviewed in the following subsequent sections.

2.3.1 Markov Chain Monte Carlo (MCMC) method

Estimation of models in the Bayesian framework can be achieved through the MCMC simulation technique. Lawson, Browne and Rodeiro (2003) describe MCMC as a set of methods which use iterative simulation of parameter values within a Markov chain. They further indicated that this simulation is necessitated by the fact that in a Bayesian approach, models are hierarchical in nature, implying a series of at least two levels and the resulting complexity of the posterior distribution of the parameters requires the use of sampling algorithms.

The convergence of these chains to a stationary distribution, which is assumed to be the posterior distribution must therefore be assessed fully (Lawson, Browne & Rodeiro, 2003). The two of the most important algorithms used in this case are the Metropolis algorithm and its extension the Metropolis-Hastings algorithm as well as the Gibbs sampler algorithm. These two are discussed in detail as follows:

2.3.1.1 The Metropolis-Hastings Algorithm

According to Lawson, Browne and Rodeiro (2003), the notion behind MCMC is to mimic a Markov Chain whose equilibrium distribution is $P(x)$.

The Metropolis-Hastings (M-H) algorithm provides a better way to correct a fairly arbitrary transition kernel $q(x'|x)$,which typically won't have $P(x)$as its equilibrium, but to rather give it a chain which does have the desired target. Catelan, Lagazio and Biggeri (2010) stated that in M-H, the transition kernel is used to generate a *proposed* new value, x' for the chain, which is then accepted as the new state at random with a particular probability $a(x'|x) = \min(1, A)$

Where $$A = P(x')q\left(x\middle|x'\right)/[P(x)q\left(x'\middle|x\right)] \qquad [2.23]$$

Wilkinson (2010) interpreted the above to say that if the value is accepted, then the chain moves on to the new proposed state, x'. Also, if the value is not accepted, the chain still advances to the next step, but the new state is given by the previous state of the chain, x.

From the above, it is clear that this algorithm is also a Markov chain, and that for $x' != x$, the transition kernel of this chain is $q(x'|x)\ a(x'|x)$ So, under some regularity conditions which are usually irreducibility and aperiodicity (Besag, York & Mollie., 1991), this chain will have $P(x)$ as its balance, and this gives a convenient way of simulating values from this target. In other words, for an equilibrium distribution to be reached, the proposal density must allow all states to communicate in a Markov chain environment.

The Metropolis–Hastings algorithm works by generating a sequence of sample values in such a way that, as more and more sample values are produced, the distribution of values more closely approximates the desired distribution $P(x)$.

Wilkinson (2010) mentioned that these sample values are produced iteratively, with the distribution of the next sample being dependent only on the current sample value (thus making the sequence of samples into a Markov chain). Specifically, at each iteration, the algorithm picks a candidate for the next sample value based on the current sample value. Then, with some probability, the candidate is either accepted (in which case the candidate value is used in the next iteration) or rejected (in which case the candidate value is discarded, and current value is reused in the next iteration). The probability of acceptance is thus, determined by comparing the values of the function $f(x)$ of the current and candidate sample values with respect to the desired distribution $P(x)$.

To illustrate how the algorithm works, Let $f(x)$ be a function that is proportional to the desired probability distribution $P(x)$.

Step 1 – Initialization:

Choose an arbitrary point x_t to be the first sample and choose an arbitrary probability density $q(x|y)$ that suggests a candidate for the next sample value x , given the previous sample value y. Assuming that q is symmetric i.e. it satisfies

$$q(x|y) = q(y|x) \quad [2.24]$$

A usual choice will be to let $q(x|y)$ be a Gaussian distribution centred at y, so that the points closer to y are more likely to be visited next, making the sequence of samples into a random walk.

Step 2 – Iteration:

For each iteration t , generate a candidate x' for the next sample by picking from the distribution $q(x'|x_t)$.

An acceptance ratio $\alpha = f(x')/f(x_t)$ which will be used to decide whether to accept or reject the candidate will be calculated. Now since f is proportional to the density of P, we have that

$$\alpha = f(x')/f(x_t)= P(x')/P(x_t) \quad [2.25]$$

The algorithm will then accept or reject the sample value by generating a uniform random number $u \in [0,1]$. The decision is that, If $u \leq \alpha$, then accept the candidate by setting $x_{t+1}= x'$,

If $u > \alpha$ then *reject* the candidate and set $x_{t+1}=x_t$ instead.

In summary, this algorithm proceeds by randomly attempting to move about the sample space, sometimes accepting the moves and sometimes remaining in place.

2.3.1.2 The Gibbs Sampler Algorithm

The Gibbs sampler is another MCMC method which updates a single parameter at a time and sample from a conditional distribution when other parameters are fixed (Yildirim, 2012). In other words, unlike in the M-H, here the analyst needs to know the conditional distributions. The idea in Gibbs sampling is to generate posterior samples by sweeping through each variable (or block of variables) to sample from its conditional distribution with the remaining variables fixed to their current values.

To illustrate how this algorithm works, let X_1, X_2 and X_3 be three random variables. The first step will be to set up the random variables to their initial values ${x_1}^{(0)}$, ${x_2}^{(0)}$, ${x_3}^{(0)}$

(often values sampled from a prior distribution *q*). At iteration *i,* the algorithm will sample

$$ {x_1}^{(i)} \sim p(X_1 = x_1 | X_2 = {x_2}^{(i-1)} , X_3 = {x_3}^{(i-1)}), \qquad [2.26]$$

sample

$$ x_2 \;\sim\; p(X_2 = x_2 | X_1 = {x_1}^{(i)} , X_3 = {x_3}^{(i-1)}), \text{ and} \qquad [2.27]$$

and sample

$$x_3 \sim p(X_3 = x_3 | X_1 = x_1{}^{(i)}, X_2 = x_2{}^{(i)}). \qquad [2.28]$$

This process continues until "convergence" i.e. until the sample values have the same distribution as if they were sampled from the true posterior joint distribution. In brief, as with other MCMC algorithms, Gibbs sampling generates a Markov chain of samples, each of which is correlated with nearby samples. As a result, care must be taken if independent samples are desired (typically by *thinning* the resulting chain of samples by only taking every n^{th} value, e.g. every 100$^{\text{th}}$ value). In addition, samples from the beginning of the chain (the *burn-in period*) may not accurately represent the desired distribution.

2.3.2 Integrated Nested Laplace approximation (INLA) method

In addition to the Markov Chain Monte Carlo (MCMC) methods described in the section above, an analytic approximation based on the Laplace method called INLA, has been recently developed as an alternative to MCMC. The INLA approach is a new tool for Bayesian inference on latent Gaussian models when the focus is on posterior marginal distributions (Blangiardo, Cameletti, Baio, & Rue, 2013). Bonat, Ribeiro Jr and Shimakura (2014) showed that INLA approach was more efficient to fit beta mixed models compared to other algorithms.

INLA substitutes MCMC simulations with accurate, and deterministic approximations to posterior marginal distributions. Literature suggests that INLA is mostly used in implementing Approximate Bayesian Inference (Martino & Rue, 2010) as a new approach to statistical inference for latent Gaussian Markov Random Field (GMRF) models. According to Martino and Rue (2010) it provides a fast, deterministic alternative to MCMC that is at the moment the standard tool for inference in GMRF models.

A latent GMRF model is a hierarchical model where, at the first stage a distributional for the observables y usually assumed to be conditionally independent given some latent parameters η and, possibly, some additional parameters $\theta_1$is found.

$$\pi(y|\eta, \theta_i) = \Pi_j \pi(y_j|\eta_j, \theta_1) \quad [2.29]$$

The latent parameters η are part of a larger latent random field x, which constitutes the second stage of our hierarchical model. The latent field x is modelled as a GMRF with precision matrix Q depending on some hyper parameters θ_2.

$$\pi(x|\theta_2) \propto \exp\{-\frac{1}{2}(x-u)^T Q(x-u)\} \quad [2.30]$$

The third, and last, stage of the model consists of the prior distribution for the hyper parameters $(\theta = \theta_1, \theta_2)$. The INLA approach provides a recipe for fast Bayesian Inference using accurate approximations *to* $\pi(\theta|y)$ and $\pi(x_i|y), i = 0, \dots, n-1,$ i.e. the marginal posterior density for the hyper parameters and the posterior marginal densities for the latent variables. Martino and Riebler (2019) added that different types of approximations are available and the approximate posterior marginal can be used to compute summary statistics of interest, such as posterior means, variances or quantiles.

This is the algorithm used in this study and an in-depth description of the Laplace approximation method and the INLA algorithm can be obtained in Martino & Rue, (2010); Blangiardo et al. (2013); Blangiardo & Cameletti (2015).

2.4 Review of relevant literature

When it comes to providing employment in particular countries, the international economic system is often failing. Also, offer and demand barely coincide in any country, leading to high rates of unemployment or underemployment in some regions. A high unemployment rate has a negative socio-economic impact such as elevated crime rates (Neema & Bohning, 2012).

2.4.1 Factors influencing unemployment

The review below is on factors that are considered to have an effect on unemployment. Such factors included education, gender, wage, age, and spatio-temporal effects. The studies reviewed were grouped into three perspectives i.e. International or global, continental and regional as presented below.

2.4.1.1. International perspective

From a global perspective, Fahrmeir, Lang, Wolf and Bender (2000) analysed unemployment duration in Germany with official data from the German Federal Employment office for the period of 1980 to1995 through an application of a Semi-Parametric hierarchical Bayesian modelling approach. Their results demonstrated that an approach which incorporates temporal and spatial effects into models is a useful and flexible tool for estimating realistically complex models of unemployment. The authors indicated that, even though the Semi-Parametric approach is useful when modelling unemployment duration data, a more refined approach in estimating unemployment risks that is based on Markov random fields in *time* and *space* could be the most appropriate.

Similarly, Ganjali and Baghfalaki (2012) used a Bayesian method and WINBUGS software to analyse unemployment duration data when the assumption of proportionality did not hold. They used a random sample of people aged 14 years and above who were unemployed in spring and summer season of 2009 from a follow-up study conducted by the Statistical centre of Iran. The authors used three models: Bayesian log-logistic; log-normal and Weibull Accelerated Failure Time (AFT) models. For each of these models, the Markov chain Monte Carlo (MCMC) method was used as a sampling technique from the joint posterior distribution of unknown quantities and models were compared using the Deviance Information Criterion (DIC).

A sensitivity test was done for model validation using Kullback-Leibler divergence. Explanatory variables included current marital status, gender, age, educational status, number of household members and place of residence. Results indicated that gender, current marital status, education level and living area were effective factors on unemployment duration in Iran, such that married persons had shorter unemployment duration than single and widow or divorced people; also unemployment duration for females were longer than that for males. Moreover, families with one or more household members had shorter unemployment durations as opposed to those with larger household members, and lastly; rural people had shorter unemployment duration than those living in Iran’s urban areas.

Rios (2014) investigated the evolution of the geographical distribution of unemployment rates in a sample of 241 NUTS-2 regions belonging to 23 European countries between 2000 and 2001. He based his study on the framework that labour market performance of a particular region is affected not only by its own labour market characteristics but also by the unemployment rates experienced by the remaining regions. He applied the recently developed space-time econometric modelling approach to study changes over time in the distribution of unemployment rates, and spatially augmented the labour market model with interdependence among regional economics. The empirical model also included spatial and time effects to control for unobserved heterogeneity and a set of regional equilibrium and dis equilibrium factors together with national labour market covariates.

Results showed that regional disparities have decreased due to the catch-up process experienced by eastern European regions with relatively high unemployment rates at the beginning of the period. Correspondingly, the spatial distribution of unemployment rates indicates that spatial effects have been relevant in shaping the evolution of unemployment variances. Based on these results, the author offers isolated actions that are aimed at fostering the reduction of regional unemployment in regions facing high-unemployment to consider the possibility of important spill overs into the neighbouring regions. He also calls for future research to increase the understanding of behaviour of underlying spill-over factors and spatio-temporal trends which plays a vital role in the observed regional unemployment disparities.

Pereira, Turknan and Correia (2018) explored the methodologies that can be used to estimate the total number of unemployed persons as well as the unemployment rates for 28 regions of Portugal, designated as NUTS III regions, using model based approaches as compared to the direct estimation methods that are currently employed by many National Statistical offices. The authors focused on the application of Bayesian hierarchical models to the Portuguese Labour Force Survey data from the 1st quarter of 2011 to the 4th quarter of 2013.

They considered and compared three different data modelling strategies namely, modelling of the total unemployed through Poisson, Binomial and Negative Binomial models; modelling of rates using a Beta model; and modelling of the three states of the labor market (employed, unemployed and inactive) by a Multinomial model using Integrated Nested Laplace Approximation (INLA) approach, except for the Multinomial model which is implemented based on the method of Monte Carlo Markov Chain (MCMC). Of relevance was that, the study could not report on the time dynamics of unemployment, because there were no sufficient quarterly sample survey data with geo-referenced sampling units. An element they considered to be of paramount in spatio-temporal analysis.

The Human Capital Theory (HCT) on the other hand also plays a vital role in understanding the labour market better. The HCT states that the more educated an individual is, the higher the probability of him finding a job. In testing this theory, Livanos (2009) modelled the incidence of unemployment in Greece and assessed whether personal attributes of individuals, such as education level, affect the probability of being employed. He used micro data from the Greece 2004 LFS and constructed a logit regression model. Results indicated that actually it is not the type of qualification (PhD, Masters, etc.) that an individual hold which affects the probability of being employed. Instead, it is the subject studied that has an effect on the incidence of unemployment.
He attributed this finding in particular to the fact that, the nature of the Greek higher educational system keeps on producing graduates whose subject areas such as Humanities and Arts do not respond to the needs of the labour market, hence the low probabilities of employment. Other attributes that were found to affect the probability of an individual unemployment are gender,

marital status and region of residence. In view of the above, it would be ideal to test the Namibian higher educational system in relation to the Namibian labour market demand; however, this would require a sole study as the current NLFS does not collect information on the individual's subject studied.

Rodokanakis (2012) also tested the HCT by investigating the probability of employment in Greece and established the types of social and demographic characteristics that increase the chances of someone in the examined population to be employed, and also determined whether a university graduate faces a greater difficulty in finding a job than a person that never went to school. He used the individual anonymised records of the Greek LFS data in 2006 for both 15-64 years employed and unemployed for the two largest regions in Greece in terms of population density (Attica & Central Macedonia) and the two biggest urban cities in the country (Athens and Thessaloniki).

They fitted a logistic regression econometric model for unemployment in SPSS version 18.0 to test the impact that various social and demographic characteristics have on people's job prospects in the whole of Greece. Explanatory variables included gender, marital status, age-groups, levels of education, residence location registered in the Manpower Employment Organization and Immigrant status. Analysis illustrated the impact of gender, age, marital status, area of residence, level of education and immigrant status on finding a job in Greece as a whole and the two most populated Greek regions, Attica and Central Macedonia. The findings of the logit model confirm the HCT as university graduates were more likely to be employed compared to primary school graduates and Masters or Doctorate holders. The study merits attention of a wider international readership, since the paper does offer evidence that could be useful for comparative research among European countries and regions. From this study, it is evident that population density did not play a role in unemployment incidence in Greece.

2.4.1.2. African perspective

Through the African lenses, Baah-Boateng (2013) investigated whether employment growth in Ghana continues to stream economic growth due to high growth of low employment generating sectors against slow-moving growth of labour absorption sectors. He used three nationally representative cross-sectional datasets from the last three rounds of the Ghana Living Standards Surveys ((GLSS) i.e. GLSS 1991/92; 1998/99; and 2005/06) and applied a cross-sectional estimation of a Probit regression model.

Results showed a strong effect of demand factors on unemployment, implying a weak employment generating impact of economic; a higher vulnerability of youth and urban dwellers to unemployment with education and gender being the main determinants; and lastly reservation wage (the expected wage or earnings by job seekers vis-à-vis the actual wage or earnings of those in employment) to have an increasing effect on unemployment. He suggested a need for policies that promote investment in agriculture and manufacturing which is associated with higher employment elasticity of output; target interventions such as support for entrepreneurial training & start-up capital to attract young school leavers to become job creators instead of job seekers; and lastly a need for a downward review of expectations on the part of job seekers in their wage reservation.

Akeju and Olanipekun (2014) explored the role of structured and unstructured heterogeneity which he refers to as the observed and unobserved heterogeneity, in explaining both cross-sectional differences across individuals as well as changes in the average duration of unemployment over time. He remarked that in order to identify the source of rise in unemployment duration and its slow recovery, economists have looked at the relationship between characteristics of unemployed individuals and their duration of unemployment. He demonstrated that, when one looks at the observable characteristics of the unemployed such as *gender, age, education* and *reason for unemployment* for instance, it becomes apparent that the duration of unemployment rises significantly for groups based on almost every observable characteristic. This observation had led to many market researchers concluding that the rise in unemployment is driven by aggregate factors that affect many workers in a similar way, not knowing that in addition to these observed factors, there could be other factors associated with the unemployed individuals that we may not directly observe that may also contribute to the variation within the relative risk of unemployment duration.

For example, employers in Namibia may statistically discriminate against individuals with the same educational qualifications for instance or living in the same geographical area who have been unemployed for longer periods. Or another common possibility could be attributed to the fact that the longer a person has been unemployed, the more likely they are to accept a low-paying job or simply drop out of the labour force (Akeju & Olanipekun, 2014).

2.4.1.3. Namibian perspective

In actual sense, there are various studies that investigated the area of unemployment in economic research in Namibia (Ashipala & Eita, 2010; Mwinga, 2012; Sunde & Akanbi, 2016; Frindt & Engelbrecht, n.d). Some analysed the causes of unemployment from a microeconomic perspective, while others investigated the macroeconomic determinants of unemployment in both developed and developing countries. Considering some of the theoretical models that are relevant for modelling unemployment as an example, Ashipala and Eita (2010) explained a commonly chosen framework job search model which states that when people become unemployed, the expected duration of their unemployment depends on the probability of receiving job offers and accepting the offers. The two authors articulated the job offer to be determined by factors such as education; skill; experience and local demand condition, all which make a specific person attractive to employers.

In detail, the Ashipala and Eita (2010) investigated the causes of unemployment in Namibia for the period of 1971 to 2007 through an extensive review of literature, microeconomic and macroeconomic models of unemployment. Due to data constraints, they could not use the macroeconomic model, instead, they applied the microeconomic model, from which results showed a negative relationship between unemployment and inflation in the country. Also revealed, was the fact that unemployment responds positively if actual output is below potential output and if wages increase. They resolved by stating that since an increase in the cost of labour causes unemployment to increase, there is a need for wage flexibility, thus, workers must reduce their wage demands if we are to see unemployment decreasing in Namibia.

With respect to assessing the current measures in place aimed at capacitating the Namibian population particularly the youth, Frindt and Engelbrecht (n.d) indicated that the Municipality Assisted Training Scheme (MATS) in Keetmanshoop is one of the employment and training programmes that have been developed and implemented over the past years to provide on-the-job training to out-of-school leavers. In their study titled "*Preparing out-of-school youth for work: A*

Namibian case study", they applied a qualitative research design approach to determine whether the Keetmanshoop MATS programme provide sufficient skill-development opportunities for young people to be employable. The study focused on 13 out-of-school leavers who applied for jobs between the periods of 2005-2007 at the Keetmanshoop municipality and investigated the perceptions of the participants and their employers regarding the effectiveness of the MATS experience in developing employability skills. Results found the programme to be useful in some respects. However, it lacks the potential to develop the necessary employability skills over a broad spectrum in a quick and qualitative manner. The authors recommended that the programme incorporates formal training workshops to allow the faster development of skills more efficiently and also advocate for more extensive research about employability skills in Namibia, from which training workshop materials could be developed. Lastly, they suggested that all municipalities in Namibia implement and adopt a revised MATS programme that would benefit the youth, employers and the country at large.

On a serious note, unemployment rates have remained high for the last three years. When for instance the NLFS 2013 produced an unemployment rate of 29.6%, which is higher than the rate of 27.4% reported in the NLFS 2012 report. The NLFS 2014 produced an unemployment rate of 28.1%, a decrease of about 1.5 points from the NLFS 2013. This trend is also coherent with Mwinga (2012) in his study on unemployment in Namibia which points to the fact that, more than half of Namibia's economically active population is unemployed, consequently posing major social, political and economic risks such as social exclusion, crime, economic welfare, erosion human capital, death, misery and social instability.

Furthermore, Sunde and Akanbi (2016) attempted to establish the causes of unemployment in Namibia for the period 1980 to 2013 using the Structural Vector Autoregressive (SVAR) methodology. Empirical results indicated that a persistently high unemployment is the result of a combination of various shocks as well as the hysteresis mechanism. The authors recommended policy makers to aggregate demand policies; deregulate policies and make structural Labour market reforms in tackling unemployment in Namibia.

Realistically, the issue and seriousness of unemployment situation in Namibia over the years cannot be highly emphasized. Although, government and other relevant sectors such as the Ministry of Trade and Industry, Ministry of Youth, Sports and Culture, Ministry of Local Government, Housing and Rural development, National Planning Commission, and other developmental partners have made considerable efforts and resources on measures pertaining to the reduction of unemployment in the country, less effort on understanding the contributing factors to the increase in the relative risk of unemployment has been achieved.

In brief, judging from the literature reviewed above, particularly in Namibia, to the best of the author's knowledge there has been no study published that critically looked at the dilemma of unemployment from a Spatio-temporal point of view. The reported rate of unemployment per capita which ranks Namibia among the highest nation in Africa needs a further understanding if it is to be reduced or managed to an acceptable level. Responding to this need, this study adopted an application of a Full Bayesian approach, which comprehensively evaluates uncertainty in the unobserved random effects contributing to variations in regional relative risk of unemployment (Kolly, 2014). These random effects often include regional clustering, unstructured heterogeneity, time, space, and time interaction, as well as population density (Neema & Bohning, 2012).

CHAPTER 3

RESEARCH METHODS

3.1 Research design

This study follows a quantitative cross-sectional research design. The study relied on already existing national microdata from the Namibia Statistics Agency. This falls under the category of document analysis and analysis of existing data. The data covers from 2014 to 2018.

3.2 Population

The target population of the NLFS's were members of private households in Namibia who lived in Namibia on the reference night of 14 October 2014; 29 September 2016; and 28 September 2018 respectively. Population living in institutions such as hospitals, hostels, police barracks and prisons were not covered in the surveys. However, private households within institutional settings were covered (e.g. teachers' houses on school premises etc.).

3.3 Sample

The sample design for all NLFS's for the 2014-2018 period was a stratified two-stage probability sample, where the first stage units were geographical areas designated as the Primary Sampling Units (PSUs) and the second stage units were the households. The Primary Sampling Units (PSUs) were based and selected from the National enumeration areas of the 2011 Namibia housing and population Census. An up-to-date listing of all households that were found in the selected PSUs was done by the interviewer on the tablet during the field work, and households were then selected using systematic sampling in each PSU.

At the field level, 8 906 of 9 108 sampled households were visited and interviewed resulting in 98% coverage in the 2014 LFS; 8 524 of 9 108 sampled households resulting in 93% coverage for the 2016 LFS, and 9 728 of 10 296 sampled households resulting in 94.8% coverage for the 2018 LFS.

3.4 Research instruments

The study used secondary data sources from the Namibia Statistics Agency, namely the LFS 2014; 2016 and 2018. The questionnaires used in LFS had five main sections, namely*, demographic characteristics of the population; labour force* and *inactive population; status in employment; status on unemployment;* and *youth unemployment*.

3.5 Procedure

For all surveys and Censuses that the NSA conducts, it publishes all the data sets (also known as microdata) on its website. All microdata sets for the NLFS 2014-2018 period were downloaded for free from the NSA website together with their micro-data documentation.

3.6 Data analysis

There were two stages of analysis considered in the study, namely the exploratory and the Fully Bayesian analysis. Covariates that were considered in the study included those presented in Table 3.1. No centering of covariates was done, rather, they were aggregated at regional level for each corresponding time period. The outcome variable (unemployment rate) was extracted from the Labour force reports (NLFS 2014; 2016 and 2018) for the respective years.

The boundary maps for all regions in Namibia were readily available from the NSA Geographical Information System (GIS) unit, and were used to visualize results at regional level.

3.6.1. Exploratory analysis

A preliminary descriptive analysis of unemployment rates observed cases was performed to gain insight about the periods of unemployment biennial incidence rates in Namibia. To test for multicollinearity between independent variables, correlation analysis was made and a value of 10 and more of the Variance Inflation Factor (VIF) was used as a cut-off point to signal severe multicolinearity (Alin, 2010).

3.6.2. Bayesian statistical modelling

The study adopted a Fully Bayesian model by Neema and Bohning (2012) based on the Besag, York and Mollie framework (Besag, York & Mollie, 1991). The regional unemployment rates were examined using an expanded fully Bayesian model that was adjusted for the smooth integration of prior information into the posterior distribution of the regional probability risk of unemployment rate. A total of eight (8) candidate models were fitted in a Bayesian modelling framework using the integrated nested Laplace approximation (INLA) in R statistical software with the aim of generating Probability Risk maps at regional level. The full hierarchical space-time model is discussed in details below following the Beta distribution.

3.6.2.1 Beta Distribution

In modelling the unemployment rate in Namibia over the period 2014-2018, the following sets of variables and their respective descriptions were considered:

Table 3.1 Description of variables considered for the analysis

Variable	***Notation***	***Description***	***Min, max***
mean_age	MA	Average regional age of persons per LFS wave	21.3; 27.8
prop_female	PF	Regional proportion of females per LFS wave	46.9; 54.9
prop_male	PM	Regional proportion of males per LFS wave	45.1; 53.1
prop_none	PN	Regional proportion of persons with no formal education per LFS wave	6.1; 45
prop_primary	PP	Regional proportion of persons with primary education per LFS wave	20.7; 54.7
prop_secondary	PS	Regional proportion of persons with secondary education per LFS wave	20.8; 61.6
prop_tertiary	PT	Regional proportion of persons with tertiary education per LFS wave	0.9; 18.1
Primary_up	PU	Regional proportion of persons primary, secondary and tertiary education per LFS wave.	55.1; 93.8
pop_density	PD	Regional population density per LFS wave	0.5; 24.3
pop_risk	PR	Regional population in the labour force per LFS wave	28 250; 241 321
UR	Y	Regional unemployment rate per LFS wave	0.219; 0.522
Exp_UR	EY	Expected regional unemployment rate per LFS wave	0.216; 0.48

Now, let y_{rt} represent the unemployment rate in region r $(r=1,2,...,R)$ during the time period t $(t=1,2,...,T)$. The distribution function assigned to modelling rates in a natural setting is the beta distribution which takes the form

$$y_{rt} | \mu_{rt}, \phi_{rt} \sim Beta(\mu_{rt}, \phi_{rt}), \tag{3.1}$$

where the conditional density function is written as:

$$p(y_{rt} | p_{rt}, q_{rt}) = \frac{\Gamma(p_{rt} + q_{rt})}{\Gamma(p_{rt})\Gamma(q_{rt})} y_{rt}^{p-1} (1 - y_{rt})^{q_{rt}-1}, \qquad 0 < y_{rt} < 1 \tag{3.2}$$

In the model, $0 < p_{rt} < 1$, $q_{rt} > 0$ and $\Gamma(.)$ is the gamma function. The above model as discussed in Cribari-Neto and Zeileis (2010), in reference to Ferrari, Cordeiro and Cribari-Neto (2001), showed that the density function can be re-parameterized by allowing and representing $\mu_{rt} = \frac{p_{rt}}{(p_{rt} + q_{rt})}$ and $\phi = p_{rt} + q_{rt}$ to take the following form:

$$y_{rt} | \mu_{rt}, \phi_{rt} \sim \frac{\Gamma(\phi_{rt})}{\Gamma(\mu_{rt}\phi_{rt})\Gamma((1-\mu_{rt})\phi_{rt})} y_{rt}^{\mu_{rt}\phi_{rt}-1} (1 - y_{rt})^{(1-\mu_{rt})\phi-1} \tag{3.3}$$

Another dimension worth pursuing is to assume that the regional unemployment rates follow a normal distribution as parameterized:

$$y_{rt} | \mu_{rt}, \sigma_{rt}^2 \sim N(\mu_{rt}, \sigma_{rt}^2) \tag{3.4}$$

3.6.2.2 Hierarchical Bayesian Approach

The unemployment rate was modelled through a hierarchical Bayesian approach using flexible prior distributions. The Hierarchical approach takes cognizant of the many layers found in the natural structures of the data and modelled them in such a way that these layers are taken into perspective to reach a certain conclusion (Neema & Bohning, 2012). The general reconstruction of the model takes the following form:

$$\log(\mu_{rt}) = \log(n_{rt}) + \beta_0 + \theta_{rt} \qquad [3.5]$$

In the model, μ_{rt} represent the average unemployment rate in region r during the time period t, while β_0 is the odds ratio before adjusting for the effects of covariates on unemployment rate, $\log(n_{rt})$ is the model offset in the form of the regional expected unemployment rate during the time period t. On the other hand, the parameter θ_{rt} measures the specific regional random effects during the time period t, which is unknown and thus requires estimation. The distribution assigned to the above random effects takes the following shape:

$$\theta_{rt} \sim N(0, \sigma^2) \qquad [3.6]$$

The parameter σ^2 is presented to measure the effects of variations between the space (represented by regional distribution) and the time period (represented by years).

According to Lawson Browne, and Vidal Rodeiro (2003), the random effects θ_{rt} can be expanded to constitute two important components in the realm of the space-time time modelling. Such expansion follows through the Besag, York and Mollie (BYM) model concepts, which seek to decompose the random effects into the component that represents the structured heterogeneity in the unemployment rate, while the second component represent the unstructured heterogeneity in the unemployment rate. In the context of the BYM concept, two parameters are therefore introduced in the model such that model (3.5) is rewritten as:

$$\log(\mu_{rt}) = \log(n_{rt}) + \beta_0 + u_r + v_r \qquad [3.7]$$

In, this case $\theta_{rt} = u_{rt} + v_{rt}$,($\theta_r = u_r + v_r$) where now u_{rt} is the random effect measuring the space-time structured heterogeneity, while v_{rt} measure the unstructured heterogeneity. In the estimation of these random effects, a conditional autoregressive (CAR) prior distribution is considered for u_{rt} as proposed by Besag, York and Mollie (1991) , considering the regional proximities. The CAR takes on the form:

$$u_{rt} \left| \bar{\mu}_{rt}, \tau_u^{-1} = \left[u_{rt} \middle| u_{lt}, for\, r \neq l, \tau_u^{-1} \right] \sim N\left(\bar{u}_{rt}, \tau_u^{-1}\right) \right. \qquad [3.8]$$

Now, the

$u_{lt} \sim N(\bar{u}_{lt,}\, \tau_u{}^{-1}$). The corresponding mean and variance are given as follows:

$$\bar{u}_{rt} = \frac{\sum_l u_{lt} w_{rlt}}{n_{rt}} \text{ and } \operatorname{var}\left(\bar{u}_{rt} \middle| \log(u_{rt}), \tau_u^{-1}\right) = \tau_u^{-1} \qquad [3.9]$$

Where n_{rt} is the number of regions in the proximity of region r (neighbouring regions) and w_{rlt} is the regional adjusted weight matrix pitting region r during a particular time period t with its neighbouring regions. The neighbours are defined in terms of regions sharing at least one point on the boundary of a polygon (Besag & Green, 1993). The view is that regions in the same neighbourhood will have similar characteristics as their neighbouring regions and thus will appear to congregate.

Similarly, for the v_{rt}, an independent normal prior of the following form is proposed:

$$v_{rt} \left| \tau_v^{-1} \sim N\left(0, \tau_v^{-1}\right)\right. \qquad [3.10]$$

Furthermore, since the study is done over a period spanning three years, it is important to assess the impact the time period has on the variations of the unemployment rates. This time effect is therefore incorporated into the model as a random component. And hence, model [3.7] was adjusted for the time component as follows:

$$log(\mu_{rt}) = log(n_{rt}) + \beta_0 + u_r + v_r + \gamma_t \qquad [3.11]$$

Where now, n_{rt}, $\beta_{0,}$ u_r and v_r are as described before above, while γ_t represent the temporal structured random effect, modelled dynamically through a neighbouring structure. In many cases a random walk approach is used which takes on the form:

$$\gamma_t | \gamma - t \sim N(\gamma_{t+1}, \tau_\gamma), \qquad t = 1$$

$$\gamma_t \left| \gamma - t \sim N\left(\frac{\gamma_{t-1} + \gamma_{t+1}}{2}, \frac{\tau_\gamma}{2}\right), \qquad t = 2,3, \dots, T - 1 \gamma_t \right| \gamma -$$

$$t \sim N(\gamma_{t-1}, \tau_\gamma) \qquad t = T \qquad [3.12]$$

Similarly, since both spatial and temporal aspects of the model are considered, it is only befitting to evaluate if the interaction between these two aspects is playing a significant role in the variation of unemployment rate. This interaction is therefore incorporated using a differential trend δ_{rt} as follows:

$$log(\mu_{rt}) = log(n_{rt}) + \beta_0 + u_r + v_r + \gamma_t + \delta_{rt} \qquad [3.13]$$

In the above model the interaction is assigned the following distribution $\delta_{rt} \sim N(0, \tau_\delta)$. In both models 3.11 and 3.13, the hyper priors τ_γ and τ_δ are to be estimated.

3.6.2.3 Hierarchical model after adjusting for covariates

The following covariates were considered for modelling in this study, namely: average regional age of persons per LFS wave, regional proportion of females per LFS wave, and regional proportion of males per LFS wave. Other covariates include those related to the average level of education such as regional proportion of persons with no formal education per LFS wave, regional proportion of persons with primary education per LFS wave, regional proportion of persons with secondary education per LFS wave, regional proportion of persons with tertiary education per LFS wave as well as Regional population density per LFS wave. The fully covariate adjusted model is thus written as:

$$log(\mu_{rt}) = log(n_{rt}) + \beta_0 + \sum_i \beta_i x_{rt} + u_r + v_r + \gamma_t + \delta_{rt}, \quad i = 1, 2, ..., I \qquad [3.14]$$

The variation attributable to each of the parameters in the models can be quantified by simply measuring the effect through the precision parameters such as: $\tau_u, \tau_{v,} \tau_{\beta_1}, \tau_{\beta_2}, ..., \tau_{\beta_I}$ also τ_γ and τ_δ. For all cases of the precision parameters, gamma hyper-prior distributions are assigned as follows:

$$\tau_u \sim \Gamma\left(a_u, b_u\right)$$
$$\tau_v \sim \Gamma\left(a_v, b_v\right)$$

.

.

.

$$\tau_{\beta_I} \sim \Gamma\left(a_{\beta_1}, b_{\beta_I}\right)$$

$\tau_\gamma \sim \Gamma\left(a_\gamma, b_\gamma\right)$ and

$$\tau_\delta \sim \Gamma(a_\delta, b_\delta) \quad [3.15]$$

The gamma hyper-parameters $a_u, b_u, a_v, b_v, a_{\beta_1}, b_{\beta_1}, \ldots, b_{\beta_I}$ also a_γ, b_γ and a_δ, b_δ are to be estimated.

It is important to note that the formulation of the above model was done assuming normal distribution of the unemployment rate within the interval of [0,1]. As such a link function implemented is the log function. However, in the case of the Beta distributed random variable representing the unemployment rate, few link functions can be adopted namely: the logit, probit and the complementary log-log link function, from which a logit function was adopted in the modelling process. In both cases i.e. modelling the response variable under the normal or beta distribution, model selection criterion was used to determine the best model amongst these candidates' models. To wrap up, table 3.2 below presents a summary of the number of candidate models to be fitted in chapter 4.

Table 3.2: The full hierarchical Bayesian model set up

Model	Variables
1.	$(u_r,\ v_r)$
2.	$(u_r,\ v_r,\ \gamma_t$, *(parametric))*
3.	$(u_r,\ v_r,\ \gamma_t,$ *(non-parametric))*
4.	$(u_r,\ v_r,\ \gamma_t,\ \delta_{rt},$ *(non-parametric with Type I space-time interaction))*
5.	$(u_r,\ v_r,\ \gamma_t,\ \delta_{rt}, x_1$ *(non-parametric with Type I space-time interaction and covariate: Primary-up))*
6.	$(u_r,\ v_r,\ \gamma_t,\ \delta_{rt}, x_1,\ x_2$ *(non-parametric with Type I space-time interaction and covariates: Primary-up & population density))*
7.	$(u_r,\ v_r,\ \gamma_t,\ \delta_{rt}, x_1,\ x_2,\ x_3$ *(non-parametric with Type I space-time interaction and covariates: Primary-up, population density & mean-age))*
8.	$(u_r,\ v_r,\ \gamma_t,\ \delta_{rt}, x_1,\ x_2,\ x_3$, x_4 *(non-parametric, time effects, with Type I space-time interaction and covariates: Primary-up, population density, mean-age & sex))*

The model selection procedure was conducted using the subsequent reduction in DIC values and the change in DIC. The best model was thus the one with the smallest DIC. This is more explained in Chapter 4.

CHAPTER 4

RESULTS

4.1 Introduction

This chapter presents results of the models fitted and the model selection criterion used. In addition, the parameter estimates for the best model fit is displayed and the interpretation of the results and the subsequent presentation of the probability risk of Unemployment rate at regional level into the maps are undertaken.

4.2 Exploratory analysis

Unemployment Rates ranged from 0.22 to 0.52 for the period 2014-2018 across the14 regions of Namibia and lower rates of unemployment were observed in the first year of study (i. e. year). Figure 4.1 presents the spatio-temporal distribution of unemployment rates in Namibia for years 2014, 2016 and2018. From this figure, it appeared that there existed variations in the unemployed rates over the 14 regions as well as over the study period. In four regions, specifically Kunene, Kavango East, Zambezi and Ohangwena, high unemployment rates were constantly observed throughout the study period. The regional and temporal variability in unemployment rates portrayed by this figure has motivated the spatial-temporal analysis undertaken in this study.

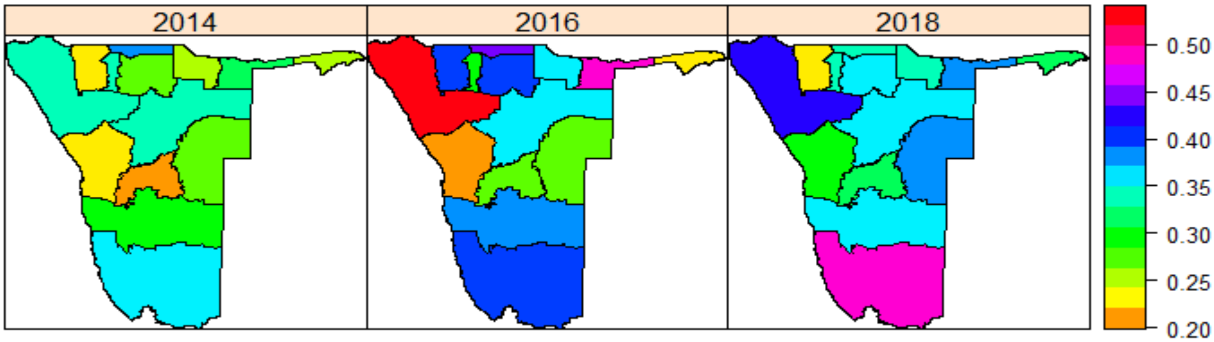


Figure 4.1: Unemployment Rates in Namibia for the period 2014-2018

Before fitting the models to the data, the researcher explored the issues of multicollinearity that may be present in the data. Firstly, a multicollinearity is a condition that occurs when independent variables are highly correlated. This condition affects the estimated regression coefficients of independent variables, as their sampling errors tend to be large (Ntirampeba, Kazembe & Neema, 2017). Many scholars still insist that there is no clear critical value of correlation among independent variables to signal multicollinearity, however, literature suggests that the absolute value of a correlation value greater than or equals to 0.7 suggest severe linear dependency between independent variables

To test for any possible linear dependency, the researcher ran a correlation analysis for the variables in Table 3.1 of chapter 3. It was found that the variables *mean-age* of the regional population & the proportion of regional population with *secondary education* (r=0.8); regional proportions of *males* & regional *population density* (r=-0.706); proportions of regional population with *no formal education* & proportions of regional population with *secondary education* (r=-

0.791); and proportions of regional population with *primary education* & proportions of regional population with *secondary education* (r=-0.694) were highly correlated. The reader may refer to Appendix C for further visualization of the correlation matrix.

Consequently, only the variables regional *mean-age*, regional proportion of persons with education level greater than *primary* (i.e. primary secondary and tertiary education), regional proportion of females, as well as the regional *population density* were used in the model fitting. A Total of eight candidate models which are summarized in Table 4.1 were then fitted in a Bayesian modelling framework using the Integrated nested Laplace approximation (INLA) in R statistical software with the aim of generating the probability risk maps of being unemployed at regional level. Eight candidate models for each assumed-distribution for the dependent variable were fitted (i.e. Gaussian and Beta models).

The first model (Model 1) assumed spatial random components as the only sources of variability in the risk of unemployment. In this model, structured & unstructured random effects were considered. This model is also known as the Besag-York-Mollie (BYM) model.

Model 2 combines a convolution model that assumes for each region two components of random effect, namely, specific-region random effect (specific region heterogeneity) and structured random effect (random effect due to clustering) as well as a temporal effect.

These models were extended by adding covariates and spatio-temporal component in a parametric formulation fashion that assumes linearity in the global time effect and the differential trend for region and time (Models 3-8). To release the assumption of linearity constraint on the differential region-time component trend, a non-parametric model was employed. Furthermore, to allow for an interaction between space and time which would explain the differences in the time trend of unemployment rate for different regions, a Type I interaction (which assumes that the two unstructured effects interact) was employed. From model 5 onwards, covariates were added in a hierarchical manner to determine whether the model improves every time a new covariate is added.

All models fitted in this study are summarized in Table 4.1. The best model (Model 5 of the beta distribution) was then used to generate probability risk maps for unemployment in order to assess regions with higher risks of unemployment rate over the study period.

4.3 The Full Bayesian Analysis

Before the researcher embarks on the full Bayesian aspect, a provision of the standard notation used throughout the course of this chapter was necessary. Thus, the following was assigned:

a) β_0 to represent the log of odds ratio before adjusting for the effects of covariates.

b) β_1 to represent the likelihood of being unemployed if you have primary, secondary or tertiary education.

c) β_2 to represent the likelihood of being unemployed given the regional population density.

d) β_3 to represent the likelihood of being unemployed given the average age of the population.

e) β_4 to represent the likelihood of females being unemployed relative to being males.

f) σ parameter measuring the mean standard deviation from the space-time regional risk of unemployment rate.

g) u random effect representing the spatial structured heterogeneity.

h) v random effect representing the spatial unstructured heterogeneity.

i) γ fixed effect representing the effect due to time period i.e. temporally structured ($u\gamma$) and temporally unstructured ($v\gamma$) effects.

j) δ space-time interaction term, to account for any spatiotemporal variation residual that was not captured by the spatial or temporal main effects (Lawson, 2015), assumed to be independently and identically distributed.

4.4 Model selection

The Deviance Information Criterion (DIC) discussed in Chapter 2 was used to select the best fit. ΔDIC is the change in the deviance with respect to the deviance of the preceding model, with a – ΔDIC indicating an improved fit on the current model. Generally, the rule of thumb indicates that the best model is one with the smallest value of DIC; however, literature suggests that a model is only considered to be significant if the ΔDIC ≥ 4 (Neema & Bohning, 2012). Models with a ΔDIC of less than 4 implies no significant difference between the models and therefore any model can be selected as a best fit model for the data.

As seen from Table 4.1, the fitting of the models begun with a simpler Bayesian model with the random component of structured and unstructured space-time heterogeneity effects (u_r, v_r) as our reference model since the space-time component form an integral part in which the study is based and then expand the model by hierarchically incorporating the rest of the variables.

Table 4.1: Comparison of the model fit and their corresponding DICs and Change in DIC

Model	Gaussian model		Beta model	
	DIC	**ΔDIC**	**DIC**	**ΔDIC**
1 ($u_{\mathrm{r}},\ v_r$)	54.87	-	38.76	-
2 ($u_r,\ v_r, \gamma_t$, *(parametric))*	54.66	-0.21	49.69	10.93
3 ($u_r,\ v_r, \gamma_t$, *(non-parametric))*	61.31	6.65	38.12	-11.57
4 ($u_{\mathrm{r}},\ v_r, \gamma_t, \delta_{rt}$, *(non-parametric with Type I space-time interaction))*	85.95	24.64	38.63	0.51
5 ($u_r,\ v_r, \gamma_t, \delta_{rt}, x_1$ *(non-parametric with Type I space-time interaction and covariate: Primary-up))*	95.03	9.08	36.99	-1.64
6 ($u_r,\ v_r, \gamma_t, \delta_{rt}, x_1, x_2$ *(non-parametric with Type I space-time interaction and covariates: Primary-up & population density))*	60.53	-34.5	37.42	0.43
7 ($u_r,\ v_r, \gamma_t, \delta_{rt}, x_1, x_2, x_3$ *(non-parametric with Type I space-time interaction and covariates: Primary-up, population density & mean-age))*	60.97	0.44	38.47	1.05
8 ($u_r,\ v_r, \gamma_t, \delta_{rt}, x_1, x_2, x_3$, x_4 *(non-parametric, time effects, with Type I space-time interaction and covariates: Primary-up, population density, mean-age & sex))*	62.60	1.63	39.78	1.31

Table 4.1 shows that overall, the model performs better when a beta distribution of unemployment rate is assumed. Under the Gaussian distribution, the model improves when the time component effect is added to the reference model (model 2, DIC: 54.66); however, changes in the deviance indicates that incorporating other variables such as the space-time interaction term (model 4) and covariates (model 5-8) to the model that already include the space-time temporal random effects does not significantly further improve the model fit. These results confirm the findings presented in table 4.3 that most of the variation in the average unemployment rate in Namibia is sufficiently explained by the temporal time trend. Contrary, when a Beta distribution of unemployment rate is

assumed, the model improves even much better when the space-time interaction term; as well as the covariate: proportion of people with *primary education and up* is added (model 5, DIC: 36.99).

Overall, based on the above results it is reasonable to fit model 5 of the Beta distribution alone in order to fully evaluate the effect of the temporal random effects; space-time interaction term as well as the educational attainment effect in the variation of unemployment rate in Namibia. The posterior mean estimate of the parameters obtained from fitting model 5 and their 95% credible intervals are presented in Table 4.2 below.

Table 4.2: Posterior mean estimates of the parameters from Model 5

Variables	**Beta model**	
	Mean	**(95% CI)**
σ_u	2131.09	(156.53; 7775.13)
σ_v	1774.01	(95.42; 6473.95)
$\sigma_{u\gamma}$	18302.44	(1223.34; 66956.85)
$\sigma_{v\gamma}$	23194.11	(2055.39; 83629.98)
σ_δ	18398.74	(1288.94; 67319.09)
$Exp_{(\beta 1)}$	0.977	(0.967; 0.988)

It is therefore evident from table 4.2 that the temporary unstructured effect ($\sigma_{v\gamma}$) has absorbed up much of the space-time variation and as such, it provides more information on the average probability of the regional unemployment rates. Likewise, model 5 provides tighter credible intervals for β_1 which could be attributed to the fact that, the proportion of people with an educational level from Primary and above have a significant negative effect on the average unemployment rate.

The spatio-temporal interaction term which assumes that the two unstructured effects v_r and γ_t interact (i.e. to say there is no spatial and/or temporal structure on the interaction either) is plotted in figure 4.2. A non-spatial pattern is clear on the interaction; this is likely because it is the variable that is taking up much of the variation. The darker the colour, the more the variation. In simple terms, when a color changes from blue to red, this means that the variation absorption of that particular parameter out of the to the total variation increases, as such, Figure 4.2 evidently shows an increase in the number of regions becoming blue and consequently red (especially in the case of !Karas region) as the years passed, which is in line with the main temporal trend observed in figure 4.4.

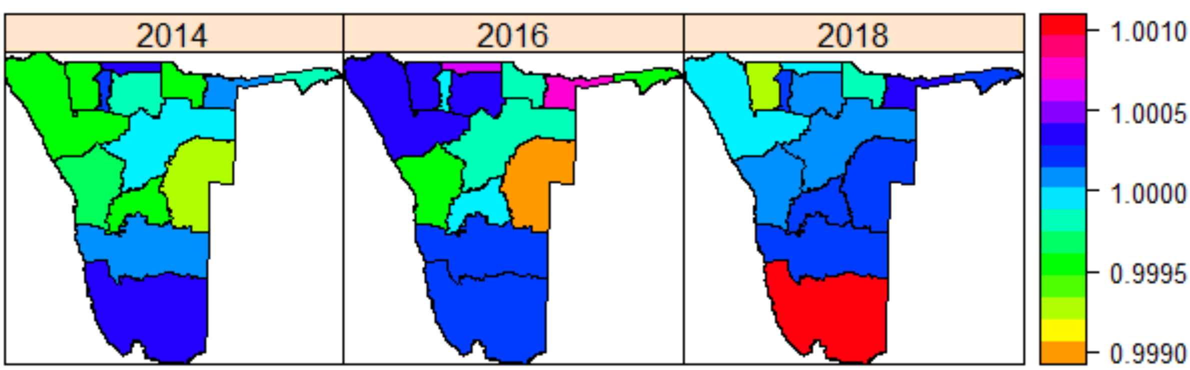


Figure 4.2 posterior mean estimates of the spatio-temporal interaction δ_{rt} (non-spatially or temporally structured interaction)

The results presented next are based on the application of the full hierarchical Bayesian model discussed in Chapter 3 to the space-time unemployment data over the time period under study. Therefore, for each of the assumed distribution of the dependent variable (Unemployment rate), the analysis begins by first fitting the full hierarchical Bayesian model of equation (3.14). In the model set up, the prior distributions for the parameters of interest are as summarized in the Methodology chapter.

The resulting posterior mean estimates for the parameters and hyper parameters of interest which are β_0, β_1 , β_2, β_3, β_4, σ_u, σ_v, σ_γ, σ_δ and their 95% credible intervals are presented in Table 4.3. Note also that here we are reporting the standard deviations σ of the mean parameters as they are more easily interpretable.

Table 4.3: Posterior mean estimates of parameters from the FB model (for all distributions)

	Gaussian model			Beta model		
Variables	**Mean**	**S.D**	(95%CI)	**Mean**	**S.D**	**(95%CI)**
exp (β_0)	1.809	1.699	(0.611; 5.021)	2.881	7.519	(0.054; 155.060)
exp (β_1)	0.994	1.002	(0.990; 0.997)	0.978	1.008	(0.963; 0.993)
exp (β_2)	1.002	1.002	(0.997; 1.007)	1.007	1.009	(0.988; 1.025)
exp (β_3)	1.002	1.011	(0.981; 1.024)	0.978	1.042	(0.902; 1.062)
exp (β_4)	1.004	1.008	(0.988; 1.021)	1.011	1.032	(0.950; 1.075)

σ_u	1413.04	1343.14	(156.49; 5042.11)	1775.24	1815.91	(103.50; 6615.31)
σ_v	2325.98	2123.39	(447.84; 7911.72)	1826.11	1846.12	(114.58; 6728.48)
$\sigma_{u\gamma}$	19827.77	20301.68	(1561.51; 74152.98)	19360	19069.69	(1399.76; 69511.04)
$\sigma_{v\gamma}$	23367.94	23313.16	(2389.76; 85125.21)	19993.29	19761.53	(1537.06; 72568.33)
σ_δ	15986.02	17019.95	(825.6; 61771.31)	19157.23	18969.63	(1369.89; 69107.8)

The result from Table 4.3 shows that the odds of being unemployed is higher when a beta distribution of the dependent variable is assumed (2.881) as opposed to a normal distribution (1.809), when covariate effects are included in the model.

It is also observed that in both distributions most of the variation in the average unemployment rate is due to the temporally unstructured time effect, since $\sigma_{v\gamma}$ contributed the highest standard deviations. Further, it is only the proportions of people with an educational level of Primary and above ($log\ \beta_1$) which have a significant negative effect on unemployment rate.

4.5 Spatio-Temporal distribution of Unemployment

The distribution of the posterior mean estimate of the risk of Unemployment rate (fitted values) is shown in Figure 4.3 over time and space. Results indicate that the regional unemployment rate ranged from 0.25 to 0.50 between the year 2014 and year 2018. Comparing this result to the observed Unemployment rate mapped in figure 4.1, both figures indicate an increase in the unemployment rate.

For interpretation purposes, a classification on average, namely, low (light to dark green), medium (blue to purplish), and high (reddish) probability classes were used to assess areas that may be prone to a higher unemployment rate in figure 4.3. The effect of regional unstructured time component is particularly well articulated in the maps where a clear regional block of increasing north-west to north-east as well as a little-bit of south gradient in the risk of unemployment is visible from 2016 to 2018. In particular, regions of Kunene, Ohangwena, Omaheke and Kavango West have been consistently classified as medium to high-risk areas in 2016 and 2018 respectively. Regions such as Otjozondjupa and Kavango East although having observed a slight decline of less than 0.2% in the average unemployment rate in 2018 from 2016 also need to be closely monitored.

The effect of the temporally unstructured effects is visible in the maps where a clear slight increase on the unemployment rate from 2014 to 2016 and then a slight decrease in 2018 is observed. In particular, Kunene region has an elevated probability of Unemployment rate higher than 0.40. Areas with a lower probability risk of Unemployment over the study period included regions such as Erongo, Khomas and Oshana.

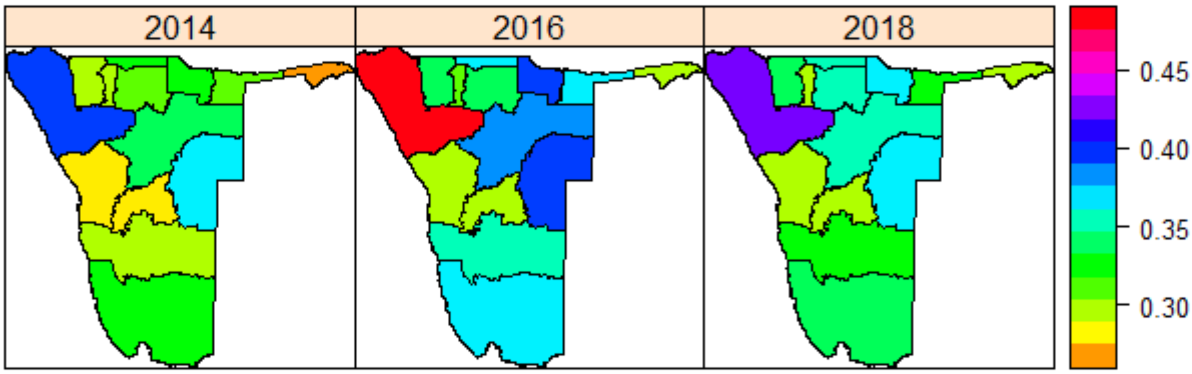
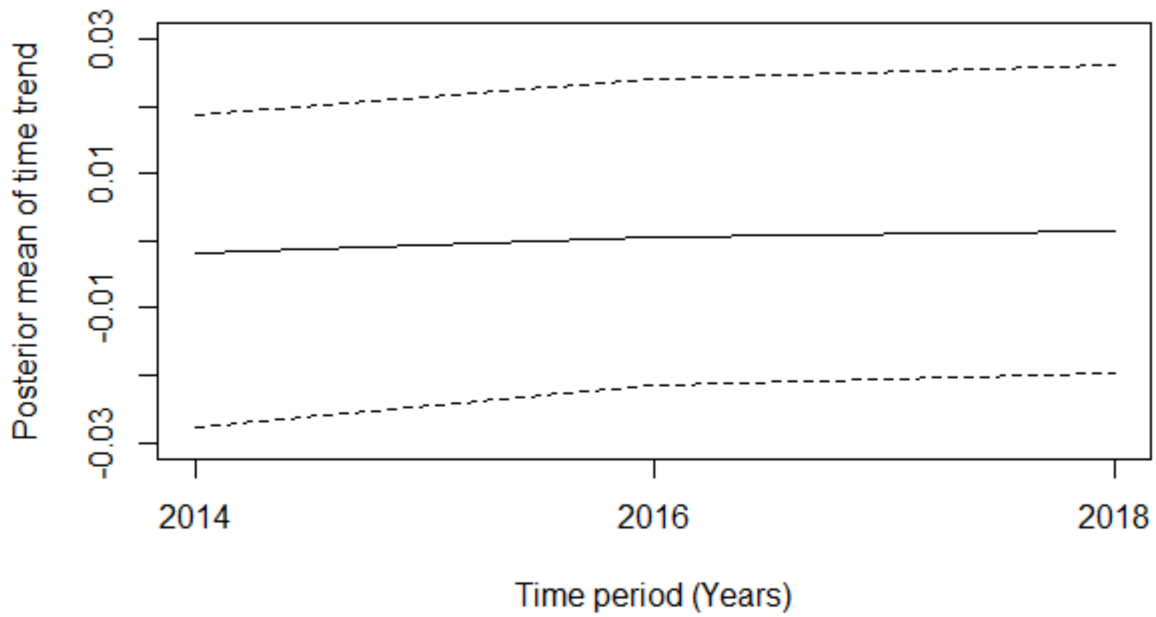


Figure 4.3 Maps of the fitted values (μ_{rt}) for the Unemployment rate for 2014-2018

In addition, the time-series plot for the posterior mean estimates of the unstructured time trend for average unemployment rate in Namibia as displayed in figure 4.4 indicates that the temporal pattern is quite evident, although with a slightly smooth increasing trend over the study period.

Figure 4.4 Time Series plot for the posterior mean estimates of the unstructured time trend with 95% C.I.



4.6 Discussions

The issue and seriousness of the unemployment situation in Namibia over the years cannot be over emphasized. Although the government of Namibia together with its developmental partners has taken and continues to take note of recent developments in the country relating to high unemployment rate and its impact on the population, less effort on understanding the contributing factors to the variation and increase in the average regional rate of unemployment has been made.

In the quest to understand the contributing factors, the thesis adopted an application of the Fully Bayesian approach, which comprehensively evaluates uncertainty in the unobserved random effects contributing to the total variation in the space-time regional average of unemployment rate. This study contributed to the existing literature by applying an existing methodology, namely hierarchical Bayesian spatio-temporal models in R-INLA to examine and assess the spatiotemporal variation in regional unemployment rates of Namibia. It is worth noting that although there are a variety of alternative models with different assumptions that we did not explicitly explore, this study is the first to incorporate spatiotemporal random effects to estimate annual regional level estimates for unemployment rates for the years 2014, 2016, and 2018 in Namibia.

A number of candidate models were fitted. The model selection procedure was conducted using the subsequent reduction in DIC values and the change in DIC. The best model was thus the one with the smallest DIC.

The modelling comparison of two distributions of the unemployment rate namely the Gaussian and Beta distribution showed that overall, the models performs better when a Beta distribution is assumed. This could be attributed to the fact that the distributions of rates and proportions are typically asymmetric, and thus Gaussian-based approximations for interval estimation and hypothesis testing can be quite inaccurate in small samples (Cribari-Neto & Zeileis, 2010). In addition, the Gaussian distribution assumes a distribution of negative infinity to positive infinity on the real number line, which is not the case for proportions or rates. A proportion lies between 0 and 1.

The evaluated random effects included regional clustering (space-time structured heterogeneity), unstructured heterogeneity, time, space and time interaction, educational level, mean-age, population density, and sex. The results indicated that most of the variation in the average regional unemployment rate in Namibia during the year 2014, 2016 and 2018 was due to the temporally unstructured effect (time trend). This could be justified by the fact that the local economy has been in a recession since the second quarter of 2016 (Namibia Statistics Agency, 2018).

Generally, a continued recession means that opportunities for unemployed people to find work become slimmer and unemployment rate rise significantly (African Development Bank, 2019). The most notable signs of a recession in an economy are factors such as high unemployment rates due to job losses and lack of jobs being available in the market as a result of a reduced investment and spend. This sign was confirmed when the Namibia Statistics Agency announced in 2018 through its Annual National Accounts report that the country has been in a recession since the second quarter of 2016, whereby the local economy had recorded a 10th straight consecutive quarter of negative growth during the period of 2016 throughout 2018 (Namibia Statistics Agency, 2018).

Moreover, the space-time regional distribution of the posterior mean estimate of the temporally unstructured time effect showed an estimated block of the unemployment risk with an increasing trend in the average unemployment rate in Kunene, Kavango West and Omaheke regions and a lower risk in regions such as Erongo, Khomas and Omusati.

In addition, the resulting space-time regional classification of the estimated average rate of unemployment identified regions such as Zambezi, Erongo Omusati, Hardap and Khomas being low risk areas for unemployment in Namibia especially during the period of 2014. In contrast, areas such as Kunene, Omaheke and Kavango West showed long-term increase in the risk of Unemployment over the study period. These results speak to the high poverty rates reported by the NSA in 2018 for the respective regions, whereby it was found that, there are very high levels of poverty in mostly Kunene and Omaheke whose poverty level is above the national average of 17.4 percent. In addition, the distribution of the severely poor households across the country was highly concentrated in Kunene and Omaheke with a severely poverty rate of above the national average of 10.7 percent (Namibia Statistics Agency, 2019).

On fixed effects that were considered in the study, results indicated that the influence of covariates such as mean-age, population density and sex (proportions of regional females) were insignificant as their overall contribution to the total variation in the average unemployment rate was minimal. On the contrary, the proportion of people with an educational level of primary and above was found to be inversely associated with unemployment rate. This shows that the population with an educational level of primary and above was able to reduce the unemployment rate in Namibia during 2014 - 2018 compared to the proportion of people with no level of education attained. This could be explained by the fact that the higher the education, the better the knowledge and skills one has, and as such, it is easier to get a job, either work for other people or open up a business alone, and in the end the unemployment rate is reduced.

The above finding is consistent with previous studies conducted in the area of unemployment that showed a significant relationship of education on unemployment rate (Jauch, 2012; Samiullah, 2014; Puspadjuita, 2018; Zimmer, 2016; Mpendulo & Mang'unyi, 2018; and Hindun, 2019). It is therefore vital for Namibia to maintain existing policies on education (e.g. free pre-primary to secondary education levels in Namibia).

In summary, improving unemployment control efforts could help reduce the average level of unemployment rate and change its geographic distribution in the years to come. In our study, despite different decentralized intervention programs aimed at reducing unemployment rate over the years, the unusually high rates of the unemployment persisted in the same places, most especially in the Kunene region, with the most likely spatial clusters showing a stable pattern in the preceding years during the study.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

The study aimed at fitting a comprehensive space-time model to estimate the true posterior mean of the unemployment rates in Namibia overtime using the LFS data set of 2014, 2016 and 2018 respectively. The study adopted a fully Bayesian model by Neema and Bohning (2012) based on the Besag, York and Mollie framework (Besag, York & Mollie; 1991). The comprehensive evaluation of the uncertainty in the unobserved random effects that are also contributing to the variation in the average probability of regional unemployment rate through the full posterior inference approach was undertaken in Chapter 4. The process was achieved by incorporating the unobservable random effects of interest into the full hierarchical Bayesian model.

The evaluated random effects included the structured space-time heterogeneity measuring the effect of regional clustering, unstructured heterogeneity, time trend effect representing the three respective LFS periods, interaction between space and time as well as the covariates effect of education measuring the regional average level of education; sex; age measuring the regional average age, as well as the density that represented the regional number of persons per kilometre square.

The model set up consisted of prior and hyper-prior distribution specifications for the parameter and hyper parameters alike while inference and programming was conducted using R-INLA. Two distribution of the prior namely Gaussian and Beta distribution were assumed from which the Beta model proved to be the best in modelling rates and proportions. The findings of the posterior mean estimates of the parameters from the full Bayesian model consolidated by the DIC as a model selection criterion, indicated that the unstructured time trend contributed to most of the variation in the average probability of unemployment rate for all distributions and not regional clustering as the study assumed in its' objectives. The effect of education was also found to have a significant negative impact on unemployment rate, while, the effect of other different risk factors such as age, sex, and population density were insignificant as their overall contribution to the total variation in the average unemployment rate was minimal.

Furthermore, the resulting mapping of the average probability of unemployment rate regional classification identified regions such as Kunene, Omaheke and Kavango West to have a long-term increase in the risk of Unemployment over the study period.

A number of significant weaknesses of this study are acknowledged. Firstly, although a large number of models were implemented, incorporating different covariates from other data sources and space and time components might have improved the predictions. Secondly, R-INLA software can implement a variety of traditional models that are built-in, however there are a class of models such as latent mixture models (e.g. the Spatial autoregressive model) that still need to be implemented. Moreover, the prior specifications that are not built-in in R-INLA need to be programmed. Thirdly, although this study incorporated a number of covariates to account for fixed effects, it acknowledged that inflation, investment and wages flexibility known to affect the unemployment (Ashipala & Eita, 2010) were not considered. Thus, future studies can look at unemployment rates in relation to these mechanisms.

Lastly, the sample design used for data collection for the period under review does not guarantee adequate coverage of all industries, because NLFS's are household-based surveys in nature and not industrially stratified, thus results should be used in view of these limitations. Also, future researchers should try and use more than 3 data-points to adequately explore the temporal trends. Nonetheless, despite the above limitations, generally, this study being the first of its kind in the area of using Bayesian analysis to model unemployment rate in Namibia, has demonstrated the usefulness of spatio-temporal analysis in describing the geographical distribution and variation of unemployment rate in different regions of the country.

5.2 Recommendations

In view of the findings, the study recommends the following on:

i) The effect of the unstructured time trend component on average Unemployment rate

From the study, it is evident that time is a factor to the risk of unemployment, in particular the unstructured time component. The assumption was that as you move from one period to another, the risk of unemployment rate changes across regions, hence the objective of assessing time as a contributing factor. The recommendation that can be drawn from this outcome is that, future studies should look deeper into the causes/ explanations of why time contributes greatly and significantly to the odds of unemployment.

ii) The effect of education on regional average unemployment rate over the study period

The recommendation that can be conveyed is that the proportion of people that have an educational level of primary and above can explain the negative effect it has on unemployment rate as opposed to people with no formal education. As such, education can then be used as a tool by the Namibian government to reduce economic problems, especially the problem of a higher unemployment rate. To build a better country, the government should continue to improve policies through education, so that the problem of a higher unemployment rate in regions such as Kunene, Omaheke and Kavango West may be reduced over time.

iii) The increasing trend of a high probability of unemployment rate in Kunene, Omaheke and Kavango West

It will be highly significant for the responsible line ministries and agencies i.e. Ministry of Poverty Eradication and Social welfare; Ministry of Labour, Industrial Relations and Employment Creation; Ministry of Urban and Rural Development; National Planning Commission among others, to continue giving a special focus on the socio-economic status (i.e. poverty and unemployment rates) of Kunene, Omaheke and Kavango West regions.

The special focus to these three (3) regions might also lead to Namibia attaining its fifth National Development Plan (NDP 5) goal number 2 which is to "*achieve Inclusive, Sustainable and Equitable Economic Growth*; as well as work towards Sustainable Development Goal number one (1) and eight (8) respectively which are to "*end poverty in all its forms everywhere"* and to *"promote decent work and economic growth for all*". As such, there is a need to have strong targeted interventions in these regions.

Lastly, the study did not include all possible explanatory variables that might have significant effects on the total variation of unemployment rate such as the Economic factors (e.g. GDP); Labour market indicators, companies (e.g. number of enterprises per 100 inhabitants in a region) and type of economic activities (e.g. Proportion of population employed in the primary sector of activity or in the secondary activity) as suggested by Pereira, Turknan and Correia (2018). Hence, future researchers should consider including these as possible fixed effects in assessing variations in the regional unemployment rates over time.

REFERENCES

African Development Bank. (2019). *African Economic Outlook 2019.* African Development Bank.

Akeju, K. F., & Olanipekun, D. B. (2014). Unemployment and economic growth in Nigeria. *Journal of Economics and Sustainable Development*, *5*(4), 138-144.

Alin, A. (2010). Multicollinearity. *Wiley Interdisciplinary Reviews: Computational Statistics*, *2*(3), 370-374.

Aregay, M., Lawson, A. B., Faes, C., & Kirby, R. S. (2017). Bayesian multi-scale modeling for aggregated disease mapping data. *Statistical methods in medical research*, *26*(6), 2726-2742.

Ashipala, J. M., & Eita, J. H. (2010). *Determinants of Unemployment in Namibia.* International Journal of Business and Management.

Assuncao, R. M., Reis, I. A., & Oliveira, C. D. L. (2001). Diffusion and prediction of Leishmaniasis in a large metropolitan area in Brazil with a Bayesian space–time model. *Statistics in Medicine*, *20*(15), 2319-2335.

Baah-Boateng, W. (2013). Determinants of unemployment in Ghana. African Development Review, 25(4), 385-399.

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

Besag, J., & Green, P. J. (1993). Spatial statistics and Bayesian computation. *Journal of the Royal Statistical Society. Series B (Methodological), 25-37.*

Besag, J., York, J., & Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the institute of statistical mathematics,* 43(1), 1-20.

Besag, J., & Kooperberg, C. (1995). On conditional and intrinsic autoregressions. Biometrika, 82(4), 733-746.

Bigenius, T. (2014). Youth unemployment in Namibia: A qualitative study among young unemployed people in Katutura.

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

Blangiardo, M., Cameletti, M., Baio, G., & Rue, H. (2013). Spatial and spatio-temporal models with R-INLA. *Spatial and spatio-temporal epidemiology*, *4*, 33-49.

Bonat, W. H., Ribeiro Jr, P. J., & Shimakura, S. E. . (2014). Bayesian analysis for a class of Beta mixed models. *Research gate*.

Catelan, D., Lagazio, C., & Biggeri, A. (2010, July 10). A hierarchical Bayesian approach to multiple testing in disease mapping. *Biometrical Journal, 52 ,*(2010), 784–797 . doi:10.1002/bimj.200900209

Cheema, A. R., & Atta, A. (2014). Economic Determinants of Unemployment in Pakistan: Co-Intergration Analysis. *International Journal of Business aand Social Science, 5*.

Cribari-Neto, F., & Zeilesis, A. (2010). Beta Distribution in R. *Journal of Statistical Software*, Volume 34, issue 2.

Fahrmeir, L., Lang, S., Wolff, J., & Bender, S. (2000). Semiparametric Bayesian time-space analysis of unemployment duration.

Ferrari, S. L., Cordeiro, G. M., & Cribari-Neto, F. (2001). Higher-order asymptotic refinements for score tests in proper dispersion models. *Journal of statistical planning and inference*, *97*(1), 177-190.

Frindt, T., Engelbrecht, F. (n.d). Preparing out-of-school youth for work: A Namibian case study. University of Namibia.

Ganjali, M., & Baghfalaki, T. (2012). *BAYESIAN ANALYSIS OF UNEMPLOYMENT DURATION DATA IN THE PRESENCE OF RIGHT AND INTERVAL CENSORING.* Journal of Reliability and Statistical Studies;.

Gilks, W. R. (2005). M arkov Chain M onte C arlo. *Encyclopedia of biostatistics*, *4*.

Hindun. (2019). Impact of Education Level on Unemployment Rate in Indonesia. *International Journal of Educational Research Review*, 0000-0001-7696-1762.

Jauch, H. (2012). Poverty, Unemployment and Inequality in Namibia . *TEMTI Series of Economic Perspectives on Global Sustainability*.

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

Knorr-Held, L., & Besag, J. (1998). Modelling risk from a disease in time and space. *Statistics in medicine*, *17*(18), 2045-2060.

Kolly, J. (2014). *Predicting the US Unemployment Rate Using Bayesian Model Averaging (Thesis)*. University of Fribourg.

Lagazio, C., Biggeri, A., & Dreassi, E. (2003). Age–period–cohort models and disease mapping. *Environmetrics: The official journal of the International Environmetrics Society*, *14*(5), 475-490.

Lawson B, A., Browne J, W., Rodeiro C, L, V. (2003). *Disease Mapping with WINBUGS and MLwiN.* Chichester: John Wiley & Sons Ltd.

Lawson, A. (2015). Bayesian Disease Mapping: Hierarchical Modelling in Spatial Epidemiology. Workshop notes(unpublished).

Livanos, I. (2009). Modelling the incidence of unemployment: the case of Greece. Applied Economics Letters, 16(16), 1607-1611.

Lopez-Quılez, A., & Munoz, F. (2009). Review of spatio-temporal models for disease mapping. *Final Report for the EUROHEIS*, *2*.

MacNab, Y. C., & Dean, C. B. (2001). Autoregressive spatial smoothing and temporal spline smoothing for mapping rates. *Biometrics*, *57*(3), 949-956.

MacNab, Y. C., & Dean, C. B. (2002). Spatio-temporal modelling of rates for the construction of disease maps. *Statistics in medicine*, *21*(3), 347-358.

Martino, S., & Riebler, A. (2019). Integrated Nested Laplace Approach (INLA). *ArXiv*.

Martino, S., & Rue, H. (2010). *Implementing approximate Bayesian inference using Integrated Nested Laplace Approximation: A manual for the INLA program*. Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway. Compiled on April, 8, 2010.

Misra, A.K., & Singh, A. K. (2010). A mathematical model for Unemployment. *Elsevier*, 128-136.

Mpendulo, G., & Mang'unyi, E. E. (2018). Exploring Relationships between Education Level and Unemployment. *Journal of Social Sciences*, 2305-9494 .

Muller-Fraqczek, I., & Pietrzak, M. B. (2011). Space-Time modelling of thee Unemployment Rate in Polish Poviats. *Dynamic Econometric models, 11*.

Mwinga, M. S. (2012). Unemployment in Namibia: Measurement problems, causes and policies. *Windhoek: First Capital Treasury Solutions. URL: http://www. firstcapitalnam. com/First% 20Capital% 20Reports/First% 20Capital% 20Namibia*, *20*.Namibia Statistics Agency. (2012). *Namibia Household Income & Expenditure Survey (NHIES) 2009/2010 Report.* Windhoek.

Namibia Statistics Agency. (2013). *Republic of Namibia 2011 Population and Housing Census, main report.* Windhoek.

Namibia Statistics Agency. (2013). Namibia Labour Force Survey 2013 Report. Namibia Statistics Agency, Windhoek.

Namibia Statistics Agency. (2015). *Analysis of Youth Employment and Unemployment in Namibia 2012-2013.* Windhoek: Namibia Statistics Agency.

Namibia Statistics Agency. (2015). *Namibia Labour Force Survey Report 2014 Report.* Windhoek: Namibia Statistics Agency.

Namibia Statistics Agency. (2018). *Annual National Accounts 2018 Report.* Windhoek.

Namibia Statistics Agency. (2018). *Namibia Household Income and Expenditure Survey (NHIES) 2015/16 Report.* Windhoek: Namibia Statistics Agency.

Namibia Statistics Agency. (2019). *Namibia Labour Force Survey Report 2018 Report.* Windhoek: Namibia Statistics Agency.

National Planning Commission. (2017). Namibia's 5th National Development Plan (NDP5).

Neema, I., & B¨ohning, D. (2012). *Monitoring Murder Crime in Namibia Using Bayesian Space-Time Models.* Hindawi Publishing Corporation.

Nobre, A. A., Schmidt, A. M., & Lopes, H. F. (2005). Spatio-temporal models for mapping the incidence of malaria in Pará. *Environmetrics: The official journal of the International Environmetrics Society*, *16*(3), 291-304.

Ntirampeba, D., Kazembe, L. N., & Neema, I. (2017). Modelling spatial patterns of misaligned disease data: An application on measles incidence in Namibia. *Clinical Epidemiology and Global Health*, 5(4), 190-195.

Ntzoufras, I. (2009). Bayesian Modeling Using Winbugs. John Wiley& Sons. *Inc, Canada*.

Pereira, S., Turknan, F., Correia, L. (2018). Spatio-temporal analysis of regional unemployment rates: A comparison of model based approaches. *Revstat - Statistical Journal, 16*.

Puspadjuita, E. A. (2018). Factors that Influence the rate of unemployment in Indonesia. *International Journal of Economics and Finance*, *10*(1), 140-147.

Richardson, S., Abellan, J. J., & Best, N. (2006). Bayesian spatio-temporal analysis of joint patterns of male and female lung cancer risks in Yorkshire (UK). *Statistical methods in medical research*, *15*(4), 385-407.

Rios, V. (2014). What drives Regional Unemployment Rate disparities in European Regions?

Rodokanakis, S., & Vlachos, V. A. (2012). Measuring the unemployment risk in Northern Greece from the LFS micro-data during the period 1994-2006. *Review of Economic Analysis*, *4*(2), 224-246.

Sam, S. O. (2016). *Modelling Economic Determinants of Youth Unemployment in Kenya.* University of Nairobi.

Samiullah. (2014). Relationship Between Unemployment and Human Capital . *Journal of Resources Development and Management*.

Schmid, V., & Held, L. (2004). Bayesian extrapolation of space–time trends in cancer registry data. *Biometrics*, *60*(4), 1034-1042.

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, *6*(2), 461-464.

Sclove, S. L. (2011). A Review of Statistical Model Selection Criteria: Application to Prediction in Regression, Histograms, and Finite Mixture Models. *Histograms, and Finite Mixture Models (June 30, 2011)*.

Sun, D., Tsutakawa, R. K., Kim, H., & He, Z. (2000). Spatio-temporal interaction with disease mapping. *Statistics in medicine*, *19*(15), 2015-2035.

Sunde, T., & Akanbi, O. A. (2016). Sources of unemployment in Namibia: an application of the structural VAR approach. *International Journal of Sustainable Economy*, *8*(2), 125-143.

Tu, Y. K., & Greenwood, D. C. (Eds.). (2012). *Modern methods for epidemiology*. Springer Science & Business Media.

Waller, L. A., Carlin, B. P., Xia, H., & Gelfand, A. E. (1997). Hierarchical spatio-temporal mapping of disease rates. *Journal of the American Statistical association*, *92*(438), 607-617.

Wilkinson, D. (2010, August 15). *Metropolis-Hastings MCMC algorithms*. Retrieved from Darren Wilkinson's research blog: Statistics, computing, data science, Bayes, stochastic modelling, systems biology and bioinformatics: https://darrenjw.wordpress.com/2010/08/15/metropolis-hastings-mcmc-algorithms/

Xia, H., & Carlin, B. P. (1998). Spatio-temporal models with errors in covariates: mapping Ohio lung cancer mortality. *Statistics in medicine*, *17*(18), 2025-2043.

Yildirim, I. (2012). Bayesian inference: Gibbs sampling. *Technical Note, University of Rochester.*

Zimmer, T. E. (2016). The Importance of Education for the Unemployed. *IBR*.

Appendix A: Stata Codes for data preparation and Extraction

```
/* -----------------------------------------------------------------------
	NAMIBIA
	LFS 2014, 2016 and 2018
	In this file we rename, re-label and recode variables to prepare and extract data for analysis
	Author: Fransina Amutenya
	Email:  afransina@rocketmail.com
	Stata version 14.0
	This version: December 10, 2019
 ------------------------------------------------------------------------- */
	set more off

//	2014 NLFS DATA SET
//	loading data
use "C:\Users\famutenya\Desktop\lfs data\2014 LFS.dta", clear
*1. Variants of Sex: Zero missings
	table RegionCode B04Sex [pw=indweight], row col format(%12.0fc)	//to check frequencies

*2. Mean Regional age: 236 unknown
	//	drop the 236 cases (0.6% )whose age is unknown before getting the regional average age
	drop if B05Age==99
	table RegionCode [pw=indweight], c(mean B05Age) row col format(%4.1fc)

*3. Educational level: Zero missings
	generate edu_cat = Education if B05Age>=6
	label var edu_cat "Education level"
```

```
recode edu_cat (9=1) (1=1) (2=2) (3/4=3) (5/8=4)

label define edu_cat  1 "None" 2 "Primary" 3 "Secondary" 4 "Tertiary"

label values edu_cat edu_cat
```

```
* Variants of Education Level

table RegionCode edu_cat [pw=indweight], row col format(%12.0fc) //to check fequencies

//*****************************************************************//
```

```
//	2016 NLFS DATA SET

//	Loading data

use "C:\Users\famutenya\Desktop\lfs data\2016 LFS.dta", clear

*1. Variants of Sex: Zero missings

table REGION_C SEX_C [pw=FINAL_WEIGHT], row col format(%12.0fc) //to check frequencies


*2. Mean Regional age: Zero missings

table REGION_C [pw=FINAL_WEIGHT], c(mean AGE_YEARS) row col format(%4.1fc)


*3. Educational level : Zero missings

generate edu_cat = Highest_Edu_level if AGE_YEARS>=6

label var edu_cat "Education level"

recode edu_cat (9/10=1) (1=1) (2=2) (3/4=3) (5/8=4)

label define edu_cat  1 "None" 2 "Primary" 3 "Secondary" 4 "Tertiary"

label values edu_cat edu_cat


* Variants of Education Level
```

table REGION_C edu_cat [pw=FINAL_WEIGHT], row col format(%12.0fc) //to check fequencies

//***//

// **2018 NLFS DATA SET**

// **Loading data**

use "C:\Users\famutenya\Desktop\lfs data\2018 LFS.dta", clear

*1. Variants of Sex: Zero missings

table REGION SEX [pw=Calib_Weight], row col format(%12.0fc) //to check frequencies

*2. Mean Regional age: Zero missings

table REGION [pw=Calib_Weight], c(mean AGE_YEARS) row col format(%4.1fc)

*3. Educational level : Zero missing

generate edu_cat = New_Highest_Edu_level

label var edu_cat "Education level"

recode edu_cat (10=1) (99=1) (1=1) (2=2) (3/4=3) (5/8=4)

label define edu_cat 1 "None" 2 "Primary" 3 "Secondary" 4 "Tertiary"

label values edu_cat edu_cat

* Variants of Education Level

table REGION edu_cat [pw=Calib_Weight], row col format(%12.0fc) //to check frequencies

//***//

Appendix B: R Codes for data Analysis

/*--

This version: 30 January 2020

#Type of models to be fitted

##################### 1.normal models (Regression models)##################

#1. Model with only intercept and random effects ####CAR MODEL####

#2. Parametric model with only intercept, random effects and time component trend

#3. Non-Parametric model now including fixed effects(CAR+fixed+time)with no space-interaction yet.

#4. Non-Parametric model now including fixed effects(CAR+fixed+time)with space-interaction.

#5. Non-Parametric model now including fixed effects(CAR+fixed+time)with space-interaction and covariates one by one.

#################### 2.Beta regression models####################

#1. Model with only intercept and random effects ####CAR MODEL####

#2. Parametric model with only intercept, random effects and time component trend.

#3. Non-Parametric model now including fixed effects(CAR+fixed+time)with no space-interaction yet.

#4. Non-Parametric model now including fixed effects(CAR+fixed+time)with space-interaction.

#5. Non-Parametric model now including fixed effects(CAR+fixed+time)with space-interaction and covariates oneby one.

--- */

#Clear memory

rm(list=ls())

#to install INLA run the code below:

#install.packages("INLA", repos=c(getOption("repos"), INLA="https://inla.r-inla-download.org/R/stable"), dep=TRUE)

library(parallel)

library(sp)

```
library(Matrix)
library(INLA)
#Load the package for building the map and import the shapefile
library(splines)
library(spData)
library(sf)
library(spdep)
library(INLA)
library(foreign)
library(shapefiles)
akimaPermit()
library(BayesX)
library(lattice)
library(maps)
library(mapproj)
library(RColorBrewer)
library(latticeExtra)
library(sp)
library(spData)
gpclibPermit()
library(maptools)
library(Matrix)
library(spdep)
library(carData)
library(car)
# Load shapefile
####14Regions
shapeR = rgdal::readOGR("C:\\cina\\Regions_boundaries_14\\Regions.shp")
```

```
#transform the shapefiles into adjancy matrix
neigR <- poly2nb(shapeR)

#make it compatible to INLA
nb2INLA("NamRegions14.graph",neigR)

NamRegions14.adj<-paste(getwd(),"/NamRegions14.graph",sep="")

#import the graph in R-inla
R<-inla.read.graph(filename="NamRegions14.graph")
image(inla.graph2matrix(R),xlab="",ylab="")

# plot neighbourood (Regions'map)
plot(shapeR)

## read the data from C DRIVE
data<-read.csv(file="C:\\cina\\UnemploymentRate.csv")

dis<-data.frame(data)
data<-data.frame(dis)
head(data,10)
```

#1.Gaussian Models

```
#-----------------------------------------------------------------------------------------------------#
```

#M1

```
#model with only intercept and random effects   ####CAR MODEL####

formula0 <- UR ~ 1+f(ID2,model="bym",graph=R)
```

```
Gaussian.model0<-inla(formula0,family="gaussian",
                    data=data,E=Exp_UR,
                    control.family=list(link="identity"),
                    #control.predictor=list(link=1, compute=TRUE),
                    control.compute=list(dic=TRUE, cpo=TRUE, waic=TRUE))
summary(Gaussian.model0)
```

#M2

```
#parametric model with time component
formula12 <- UR ~ 1+f(ID2,model="bym",graph=R)+f(ID1,YEAR.ID1,model="rw1")

Gaussian.model12<-inla(formula12,family="gaussian",
                     data=data,E=Exp_UR,
                     control.family=list(link="identity"),
                     #control.predictor=list(link=1, compute=TRUE),
                     control.compute=list(dic=TRUE, cpo=TRUE, waic=TRUE))
summary(Gaussian.model12)
```

#M3

```
#Non-Parametric model (spatial random effects,time,with no space-time interaction)
formula13 <- UR ~
1+f(ID2,model="bym",graph=R)+f(ID1,YEAR.ID1,model="rw1")+f(YEAR.ID2,model="iid")

Gaussian.model13<-inla(formula13,family="gaussian",
                     data=data,E=Exp_UR,
                     control.family=list(link="identity"),
                     #control.predictor=list(link=1, compute=TRUE),
                     control.compute=list(dic=TRUE, cpo=TRUE, waic=TRUE))
```

summary(Gaussian.model13)

#M4

#None parametric (spatial random effects,time,with space-time interaction)

formula122 <- UR ~
1+f(ID2,model="bym",graph=R)+f(YEAR.ID1,model="rw1")+f(YEAR.ID2,model="iid")+f(REGION.YEAR.ID,model="iid")

Gaussian.model122<-inla(formula122,family="gaussian",

data=data,E=Exp_UR,

control.family=list(link="identity"),

#control.predictor=list(link=1, compute=TRUE),

control.compute=list(dic=TRUE, cpo=TRUE, waic=TRUE))

summary(Gaussian.model122)

#M5

#Type I interaction with covariates, primary up

#Non-Parametric model (spatial random effects,time,with space-time interaction, covariates)

formula123 <- UR ~
1+primary_up+f(ID2,model="bym",graph=R)+f(YEAR.ID1,model="rw1")+f(YEAR.ID2,model="iid")+f(REGION.YEAR.ID,model="iid")

Gaussian.model123<-inla(formula123,family="gaussian",

data=data,E=Exp_UR,

control.family=list(link="identity"),

#control.predictor=list(link=1, compute=TRUE),

control.compute=list(dic=TRUE, cpo=TRUE, waic=TRUE))

summary(Gaussian.model123)

#M6

#Type I interaction with covariates, primary up, pop density

#Non-Parametric model (spatial random effects,time,with space-time interaction, covariates)

```
formula124 <- UR ~ 
1+primary_up+pop_density+f(ID2,model="bym",graph=R)+f(YEAR.ID1,model="rw1")+f(YEAR.ID2,model="iid")+f(REGION.YEAR.ID,model="iid")

Gaussian.model124<-inla(formula124,family="gaussian",
                   data=data,E=Exp_UR,
                   control.family=list(link="identity"),
                   #control.predictor=list(link=1, compute=TRUE),
                   control.compute=list(dic=TRUE, cpo=TRUE, waic=TRUE))
summary(Gaussian.model124)
```

#M7

#Type I interaction with covariates, primary up, pop density, age

#Non-Parametric model (spatial random effects,time,with space-time interaction, covariates)

```
formula125 <- UR ~ 
1+primary_up+pop_density+mean_age+f(ID2,model="bym",graph=R)+f(YEAR.ID1,model="rw1")+f(YEAR.ID2,model="iid")+f(REGION.YEAR.ID,model="iid")

Gaussian.model125<-inla(formula125,family="gaussian",
                   data=data,E=Exp_UR,
                   control.family=list(link="identity"),
                   #control.predictor=list(link=1, compute=TRUE),
                   control.compute=list(dic=TRUE, cpo=TRUE, waic=TRUE))
summary(Gaussian.model125)
```

#M8

```
#Type I interaction with covariates, primary up, pop density, age, females

#Non-Parametric model (spatial random effects,time,with space-time interaction, covariates)


formula126 <- UR ~
1+primary_up+pop_density+mean_age+prop_female+f(ID2,model="bym",graph=R)+f(YEAR.ID1,model="rw1")+f(YEAR.ID2,model="iid")+f(REGION.YEAR.ID,model="iid")


Gaussian.model126<-inla(formula126,family="gaussian",
                data=data,E=Exp_UR,
                control.family=list(link="identity"),
                #control.predictor=list(link=1, compute=TRUE),
                control.compute=list(dic=TRUE, cpo=TRUE, waic=TRUE))
summary(Gaussian.model126)
```

#2.Beta regression models

```
#-----------------------------------------------------------------------------------------------------#
```

#M1

```
#beta model with random effects only
formula2 <- UR ~ 1+ f(ID2,model="bym",graph=R)


Beta.car.model<-inla(formula2,family="beta",
              data=data,E=Exp_UR,
              control.family=list(link="logit"),
              control.predictor=list(link=1, compute=TRUE),
              control.compute=list(dic=TRUE, cpo=TRUE, waic=TRUE))


summary(Beta.car.model)
#pred<-predict(Beta.car.model,newdata=data)
#pred2<-Beta.car.model$summary.fix
```

#pred2

#M2

#Parametric model(CAR)

formula3 <- UR ~ -1+f(ID2,model="bym",graph=R)+f(ID1,YEAR.ID1,model="rw1")

bym.parametric.Betamodel <- inla(formula3,family="beta",

data=data,E=Exp_UR,

control.family=list(link="logit"),

control.predictor=list(link=1, compute=TRUE),

control.compute=list(dic=TRUE, cpo=TRUE, waic=TRUE))

summary(bym.parametric.Betamodel)

#M3

#Non-parametric model(CAR)With no space-time interaction

formula4 <- UR ~ -
1+f(ID2,model="bym",graph=R)+f(ID1,YEAR.ID1,model="rw1")+f(YEAR.ID2,model="iid")

BYM.nonparametric.model1 <- inla(formula4,family="beta",

data=data,E=Exp_UR,

control.family=list(link="logit"),

control.predictor=list(link=1, compute=TRUE),

control.compute=list(dic=TRUE, cpo=TRUE, waic=TRUE))

summary(BYM.nonparametric.model1)

#M4

#Type I interaction (Non-parametric with space-time interaction)

#It is easy to expand this model to allow for an interaction between space

#and time, which would explain diferences in the time trend of high unemployment rate for diferent areas

```
formula5 <- UR ~
1+f(ID2,model="bym",graph=R)+f(YEAR.ID1,model="rw1")+f(YEAR.ID2,model="iid")+f(RE
GION.YEAR.ID,model="iid")


BYM.IID.nonparametric.modelInterI <-inla(formula5,family="beta",
                    data=data,
                    control.family=list(link="logit"),
                    control.predictor=list(link=1, compute=TRUE),
                    control.compute=list(dic=TRUE, cpo=TRUE, waic=TRUE))
summary(BYM.IID.nonparametric.modelInterI)


#inla.doc("iid")
#inla.doc("RW1")
#inla.doc("random")
```

#M5

```
#Type I interaction with space-time interaction, primary_up
formula6 <- UR ~
1+primary_up+f(ID2,model="bym",graph=R)+f(YEAR.ID1,model="rw1")+f(YEAR.ID2,model
="iid")+f(REGION.YEAR.ID,model="iid")


BYM.IID.nonparametric.modelcovInterI1 <-inla(formula6,family="beta",
                    data=data,
                    control.family=list(link="logit"),
                    control.predictor=list(link=1, compute=TRUE),
                    control.compute=list(dic=TRUE, cpo=TRUE, waic=TRUE))
summary(BYM.IID.nonparametric.modelcovInterI1)
```

#M6

#Type I interaction with space-time interaction,primary_up, density

formula7 <- UR ~
1+primary_up+pop_density+f(ID2,model="bym",graph=R)+f(YEAR.ID1,model="rw1")+f(YEAR.ID2,model="iid")+f(REGION.YEAR.ID,model="iid")

BYM.IID.nonparametric.modelcovInterI2 <-inla(formula7,family="beta",

data=data,

control.family=list(link="logit"),

control.predictor=list(link=1, compute=TRUE),

control.compute=list(dic=TRUE, cpo=TRUE, waic=TRUE))

summary(BYM.IID.nonparametric.modelcovInterI2)

#M7

#Type I interaction with space-time interaction,Primary_up, density,age

formula8 <- UR ~
1+primary_up+pop_density+mean_age+f(ID2,model="bym",graph=R)+f(YEAR.ID1,model="rw1")+f(YEAR.ID2,model="iid")+f(REGION.YEAR.ID,model="iid")

BYM.IID.nonparametric.modelcovInterI3 <-inla(formula8,family="beta",

data=data,

control.family=list(link="logit"),

control.predictor=list(link=1, compute=TRUE),

control.compute=list(dic=TRUE, cpo=TRUE, waic=TRUE))

summary(BYM.IID.nonparametric.modelcovInterI3)

#M8

#Type I interaction with space-time interaction,Primary_up, density,age,Females

formula9 <- UR ~
1+primary_up+pop_density+mean_age+prop_female+f(ID2,model="bym",graph=R)+f(YEAR.ID1,model="rw1")+f(YEAR.ID2,model="iid")+f(REGION.YEAR.ID,model="iid")

```
BYM.IID.nonparametric.modelcovInterI4 <-inla(formula9,family="beta",
                              data=data,
                              control.family=list(link="logit"),
                              control.predictor=list(link=1, compute=TRUE),
                              control.compute=list(dic=TRUE, cpo=TRUE, waic=TRUE))
summary(BYM.IID.nonparametric.modelcovInterI4)

#################################################################
# Draw maps of regional UR (as fig 4.2)// OBSERVED VALUES
#Clear Memory
rm(list=ls())

library(sp)
#Regions
shape = rgdal::readOGR("C:\\cina\\Regions_boundaries_14\\Regions.shp")
shape@data

#Create areas IDs used here
shape$ID<-1:nrow(shape@data)
data<-read.csv("C:\\cina\\URincidences.csv")

### Put the data into shape@data
old.shape = shape
shape<-data.frame(shape@data,data,shape$ID)

old.shape$Prev1<-data$UR1
old.shape$Prev2<-data$UR2
```

```
old.shape$Prev3<-data$UR3


spplot(old.shape, c("Prev1","Prev2","Prev3")
    ,names.attr=c("2014","2016","2018"),col.regions = rainbow(99, start=.1), layout=c(3,1))



###########################################################################
#Figure 4.3 Interaction term
#Clear Memory
rm(list=ls())

library(sp)
#Regions
shape = rgdal::readOGR("C:\\cina\\Regions_boundaries_14\\Regions.shp")
shape@data

#Create areas IDs used here
shape$ID<-1:nrow(shape@data)
data<-read.csv("C:\\cina\\interaction.csv")

### Put the data into shape@data
old.shape = shape
shape<-data.frame(shape@data,data,shape$ID)



old.shape$prev1<-data$prev1
old.shape$prev2<-data$prev2
old.shape$prev3<-data$prev3
```

```
spplot(old.shape, c("prev1","prev2", "prev3")
      ,names.attr=c("2014","2016", "2018"),col.regions = rainbow(99, start=.1), layout=c(3,1))

##########################################################################
# Draw maps of regional fitted probabilities of UR (as fig 4.4)

#Clear Memory
rm(list=ls())

library(sp)
#Regions
shape = rgdal::readOGR("C:\\cina\\Regions_boundaries_14\\Regions.shp")
shape@data

#Create areas IDs used here
shape$ID<-1:nrow(shape@data)
data<-read.csv("C:\\cina\\M5_Beta_fitted.csv")

### Put the data into shape@data
old.shape = shape
shape<-data.frame(shape@data,data,shape$ID)

#Probability maps
#Figure 4.5 Beta model 5
old.shape$Prev1<-data$beta1
old.shape$Prev2<-data$beta2
old.shape$Prev3<-data$beta3
```

```
spplot(old.shape, c("Prev1","Prev2","Prev3")
    ,names.attr=c("2014","2016","2018"),col.regions = rainbow(99, start=.1), layout=c(3,1))

###############################################################################
#FIGURE 4.5 TIME SERIES PLOT

#BETA DISTRIBUTION
X<-c("2014","2016","2018")
x<-seq(1:3)
global<-read.csv(file="C:\\cina\\m5_year_id2.csv")
plot(X, global$mean,type="l", main="",xlab="Time period (Years)",ylab="Posterior mean of the time trend", ylim=c(-0.03,0.03))
lines(X,global$LB,lty=2)
lines(X, global$UB,lty=2)

BYM.IID.nonparametric.modelcovInterI1$marginals.random$REGION.YEAR.ID
```

```
#end of source code.
```

Appendix C: Correlation matrix between independent variables

	mean_age	prop_female	prop_none	prop_primary	prop_secondary	pop_density
mean_age	1	-0.4961	-0.54351	-0.63385	0.813711	-0.22194
prop_female	-0.4961	1	-0.1128	0.56344	-0.34713	0.706274
prop_none	-0.54351	-0.1128	1	0.173554	-0.79182	-0.25517
prop_primary	-0.63385	0.56344	0.173554	1	-0.69397	0.112147
prop_secondary	0.813711	-0.34713	-0.79182	-0.69397	1	0.008224
pop_density	-0.22194	0.706274	-0.25517	0.112147	0.008224	1
UR	-0.50821	0.175793	0.59751	0.105329	-0.4928	0.073509