

Investigation of Regional Geomagnetic Field Modelling in Indonesia

By

Yosi Setiawan

A thesis
submitted to Victoria University of Wellington
in fulfilment of the requirements for the degree of
Master of Science in Physics

Victoria University of Wellington
2020

Abstract

This thesis deals with the application of the Spherical Cap Harmonic Analysis (SCHA) modelling technique to obtain geomagnetic field models for Indonesia, which have better resolution and accuracy than the International Geomagnetic Reference Field (IGRF). B-splines basis function and autoregressive forecasting are applied to improve estimates of secular variation and its forecast over the Indonesian region. The modelling technique is applied to geomagnetic observation data compiled from 68 geomagnetic repeat stations in Indonesia covering the period 1985 - 2015 from BMKG (*Badan Meteorologi Klimatologi dan Geofisika* / Agency for Meteorology, Climatology, and Geophysics) Indonesia, definitive data from five BMKG geomagnetic observatories and 13 INTERMAGNET (*The International Real-Time Magnetic Observatory Network*) observatories. Synthetic cartesian X, Y, and Z components at sea level at 17 fixed locations, calculated from IGRF-13, are also used. The area covered by the models in this thesis is the Indonesian region with a spherical cap half-angle of 30° and with the coordinate of the spherical cap pole at 122°E and 3°S . From statistical analysis and comparison with the IGRF, the SCHA model with index $k = 7$ is considered as the best SCHA model, both in resolution and accuracy. Compared with the root mean square deviation (RMSD) of the IGRF model, the RMSD of the SCHA model with index $k = 7$ is lower by 28 nT, 11 nT, and 34 nT for X, Y, and Z components, respectively. A model from interpolation of the SCHA with index $k = 7$ using the B-splines basis function for the year 1985.5 – 2015.5 shows that the SCHA model gives better results than the IGRF. The forecasting calculation for the year 2015.5 – 2020.5 suggests that the autoregressive order 3 of the SCHA with index $k = 7$ gives better results than the forecasting of the IGRF model, especially in the X, Z, and F components. However, in the Y component, the IGRF is still better than the SCHA model. The RMSD of the forecasted SCHA model is 154.92 nT, 200.87 nT, 104.39 nT, and 135.81 nT for X, Y, Z, and F components, respectively, while the RMSD of the IGRF model is 172.62 nT, 95.52 nT, 117.55 nT, and 162.38 nT for X, Y, Z, and F components. Thus, the forecasted SCHA model is suitable for data reduction of geomagnetic surveys in the Indonesian region but not preferable for navigation.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisor, Malcolm Ingham, for always guiding me to perform this research. I thank him for encouragement, patience, and guidance in all the time spent researching and writing this thesis.

I owe my most sincere gratitude to Gillian Turner for providing me with the opportunity to perform this research and providing me with invaluable resources during this research.

I am grateful to the Wellington University International staff who supported me relentlessly to finish this research.

Also, I would like to thank the New Zealand Ministry of Foreign Affairs and Trade for the financial support they provided for this study. This research was conducted under a New Zealand ASEAN Scholarship programme.

I thank the Head of the Agency for Meteorology, Climatology, and Geophysics, Deputy Head for Geophysics, and Head of Deli Serdang Geophysical Station, who issued a study permit for my Master's programme. I also thank the Division for Geopotential and Time Code for providing me with the data used in this research.

I express gratitude to my family: my parents, my parents-in-law, and my siblings for their emotional support and encouragement throughout my life in general.

Finally, I cannot finish without saying that my love and gratitude goes to my beloved wife, Imma Redha Nugraheni, for her support and endless love, which gave me strength during my studies and the courage and determination to pursue my goals even further.

Contents

Abstract	iii
Acknowledgements	v
Contents	vii
List of Figures	ix
List of Tables	xiii
Chapter 1 Introduction	1
1.1 Background of the Study	1
1.2 Thesis Aims.....	5
1.3 Thesis Structure	5
Chapter 2 Literature Review	7
2.1 The Origin of Geomagnetic Field.....	7
2.2 Geomagnetic Field Modelling.....	9
2.2.1 Spherical Harmonic Analysis (SHA)	9
2.2.2 Spherical Cap Harmonic Analysis (SCHA)	13
2.3 The IGRF Model.....	18
2.4 Secular Variation.....	20
2.4.1 Origin of Secular Variation	21
2.4.2 Secular Variation Prediction	22
Chapter 3 Data and Methods	25
3.1 Data Sets and Data Processing	25
3.1.1 Geomagnetic Observatories Data.....	26
3.1.2 Repeat Stations Data	27
3.1.3 The IGRF Model.....	28
3.1.4 Outlier Detection and Removal	28
3.2 Modelling Procedure	30
3.2.1 Calculation of the Associated Legendre Function and the SCHA Degree	30
3.2.2 Coordinate Rotation.....	32
3.2.3 Removal of a Global Reference Field.....	34
3.2.4 Calculation of the Gauss Coefficients	35
3.3 Spatial Model Selection.....	36
3.3.1 Model Statistics and Validation.....	36
3.3.2 Model Truncation and Maximum Resolution	37

3.4	Temporal Modelling and Secular Variation Forecasting	38
3.4.1	Time-dependent Model Based on Cubic B-splines	39
3.4.2	Autoregressive Forecasting	40
Chapter 4	Results and Discussions	43
4.1	Spatial Modelling Analysis	43
4.1.1	RMSD of the Model.....	43
4.1.2	Stability of the Model.....	45
4.1.2.1	IGRF Model	46
4.1.2.2	SCHA with Spatial Truncation Index $k = 7$	52
4.1.2.3	SCHA with Spatial Truncation Index $k = 8$	58
4.1.2.4	SCHA with Spatial Truncation Index $k = 9$	63
4.1.3	Validation of the Model.....	69
4.1.3.1	Bandar Lampung.....	69
4.1.3.2	Tual	72
4.1.4	Difference between the Regional and Global Model	75
4.1.4.1	Difference between SCHA Index $k = 7$ and the IGRF Model.....	76
4.1.4.2	Difference between SCHA Index $k = 8$ and the IGRF Model.....	81
4.1.4.3	Difference between SCHA Index $k = 9$ and the IGRF Model.....	86
4.1.5	Selection of the Truncation Level.....	91
4.1.6	Misfit Analysis between Observed and Predicted Data.....	92
4.2	Temporal Modelling Analysis	97
4.2.1	Temporal Modelling using B-splines.....	97
4.2.2	Secular Variation Forecasting using Autoregressive Forecasting	102
Chapter 5	Summary	113
5.1	Summary.....	113
5.2	Recommendations for Further Research.....	114
References		117

List of Figures

Figure 1. Selected cap region.....	5
Figure 2. Geomagnetic field components, from Hulot et al. (2015).....	8
Figure 3. Schematic representation of the spherical cap	14
Figure 4. An approximate grand spectrum of the geomagnetic field, reproduced from Constable and Constable (2004)	20
Figure 5. Data distribution used in this research	26
Figure 6. Y Component of the magnetic field at Denpasar and Timika repeat stations (black line with green dots), with identified outliers (red dot)	29
Figure 7. Spherical cap structure, from Fiori (2020).....	32
Figure 8. Transformation of the geographic coordinate system to the spherical cap coordinate system.....	33
Figure 9. RMSD of the IGRF and SCHA model with different index k from 1985.5 – 2015.5	44
Figure 10. X component contour map from the IGRF model from 1985.5 - 2015.5.	48
Figure 11. Y component contour map from the IGRF model from 1985.5 – 2015.5.	49
Figure 12. Z component contour map from the IGRF model from 1985.5 – 2015.5.	50
Figure 13. F component contour map from the IGRF model from 1985.5 – 2015.5.	51
Figure 14. X component contour from the SCHA using index $k=7$ from 1985.5 – 2015.5	54
Figure 15. Y component contour from the SCHA using index $k=7$ from 1985.5 – 2015.5	55
Figure 16. Z component contour from the SCHA using index $k=7$ from 1985.5 – 2015.5	56
Figure 17. F component contour from the SCHA using index $k=7$ from 1985.5 – 2015.5	57
Figure 18. X component contour from the SCHA using index $k=8$ from 1985.5 – 2015.5	59
Figure 19. Y component contour from the SCHA using index $k=8$ from 1985.5 – 2015.5	60

Figure 20. Z component contour from the SCHA using index $k=8$ from 1985.5 – 2015.5.....	61
Figure 21. F component contour from the SCHA using index $k=8$ from 1985.5 – 2015.5.....	62
Figure 22. X component contour from the SCHA using index $k=9$ from 1985.5 – 2015.5.....	65
Figure 23. Y component contour from the SCHA using index $k=9$ from 1985.5 – 2015.5.....	66
Figure 24. Z component contour from the SCHA using index $k=9$ from 1985.5 – 2015.5.....	67
Figure 25. F component contour from the SCHA using index $k=9$ from 1985.5 – 2015.5.....	68
Figure 26. Differences between models and the observed field at Bandar Lampung from 1985.5 – 2015.5.....	70
Figure 27. The RMSD between models and the observed field at Bandar Lampung from 1985.5 – 2015.5.....	71
Figure 28. Differences between models and the observed field at Tual from 1985.5 – 2015.5.....	73
Figure 29. The RMSD between models and the observed field at Tual from 1985.5 – 2015.5.....	74
Figure 30. Difference contour at X component between the SCHA using index $k=7$ and IGRF from 1985.5-2015.5.....	77
Figure 31. Difference contour at Y component between the SCHA using index $k=7$ and IGRF from 1985.5-2015.5.....	78
Figure 32. Difference contour at Z component between the SCHA using index $k=7$ and IGRF from 1985.5-2015.5.....	79
Figure 33. Difference contour at F component between the SCHA using index $k=7$ and IGRF from 1985.5-2015.5.....	80
Figure 34. Difference contour at X component between the SCHA using index $k=8$ and IGRF from 1985.5-2015.5.....	82
Figure 35. Difference contour at Y component between the SCHA using index $k=8$ and IGRF from 1985.5-2015.5.....	83
Figure 36. Difference contour at Z component between the SCHA using index $k=8$ and IGRF from 1985.5-2015.5.....	84
Figure 37. Difference contour at F component between the SCHA using index $k=8$ and IGRF from 1985.5-2015.5.....	85

Figure 38. Difference contour at X component between the SCHA using index $k=9$ and IGRF from 1985.5-2015.5	87
Figure 39. Difference contour at Y component between the SCHA using index $k=9$ and IGRF from 1985.5-2015.5	88
Figure 40. Difference contour at Z component between the SCHA using index $k=9$ and IGRF from 1985.5-2015.5	89
Figure 41. Difference contour at F component between the SCHA using index $k=9$ and IGRF from 1985.5-2015.5	90
Figure 42. Misfit map at X component between the SCHA using index $k=7$ and observed data from 1985.5-2015.5	94
Figure 43. Misfit map at Y component between the SCHA using index $k=7$ and observed data from 1985.5-2015.5	95
Figure 44. Misfit map at Z component between the SCHA using index $k=7$ and observed data from 1985.5-2015.5	96
Figure 45. Annual mean values at Tuntungan, Bandar Lampung and Tual	98
Figure 46. Annual rate of change at Tuntungan, Bandar Lampung and Tual	101
Figure 47. Partial autocorrelation plot at Tuntungan, Bandar Lampung, and Tual	103
Figure 48. Annual mean values at Tuntungan, Bandar Lampung, and Tual together with the values predicted by the SCHA and the IGRF	105
Figure 49. Forecasted X component contour from 2016.5 – 2020.5	107
Figure 50. Forecasted Y component contour from 2016.5 – 2020.5	108
Figure 51. Forecasted Z component contour from 2016.5 – 2020.5	109
Figure 52. Forecasted F component contour from 2016.5 – 2020.5	110
Figure 53. 95% and 80% confidence interval from the forecasted SCHA model at Bandar Lampung and Tual	111

List of Tables

Table 1. The number of repeat stations surveyed by BMKG	27
Table 2. Number of observation data points before and after outlier rejection using the MAD method including two validation locations.....	29
Table 3. $n_k(m)$	31
Table 4. RMSD from the IGRF model in 1985.5 - 2015.5.....	46
Table 5. RMSD from the SCHA model with index $k = 7$	52
Table 6. RMSD from the SCHA model with index $k = 8$	58
Table 7. RMSD from the SCHA model with index $k = 9$	63

Chapter 1

Introduction

1.1 Background of the Study

The magnetic field is one of the oldest observable properties of the Earth. The legend of the Greek shepherd Magnes whose iron-studded boots and his staff kept sticking to the path, is thought to date back to around 900 BC. It was the Roman scholar and writer Pliny the Elder who recounted this story almost a millennium later in his multi-volume encyclopedia, *Naturalis Historia*. However, the earliest perception of the nature of magnetism is usually attributed to Thales (c. 624 – 546 BC), a Greek philosopher who provided an animistic explanation for magnetic force (Turner, 2011). Nevertheless, it was the Chinese, from about the second century AD, who discovered the fact that magnets have the property to align themselves in the north-south direction and utilised this property to develop compasses. Later, William Gilbert identified the magnetic field as a property of the Earth itself in his book *De Magnete* in 1600, and he concluded that the Earth is a spherical magnet (Kono, 2015).

The discovery of the geomagnetic components took place at different times. The existence of the declination was recognised in 1088 in China and was used for navigation. However, Europeans only realised it in the middle of the fifteenth century AD. The inclination was recognised in 1544 (Kono, 2015), and the intensity itself was first measured in 1791 (Jackson & Finlay, 2015).

Knowledge of the Earth's magnetic field developed slowly. However, nowadays, it has developed rapidly, and more details about the sources of the observed geomagnetic field and its variations are known. Further applications of geomagnetic data apart from navigation and understanding the generation of the geomagnetic field have been recognised. With current understandings, it is known that every observed geomagnetic field observation is the sum of several contributions, i.e. internal sources and external sources. The largest part, 90%, of the internal sources is generated by the dynamo process in the fluid core and change slowly by what is known as *secular variation*. This part of the field is also known as the core or main field and has features with scales of several thousands of kilometres. The rest of the internal field comes from the remanent magnetisation of igneous rocks and sediments obtained at the time of its formation. This field is also known as the *lithospheric* or *crustal field*, and it is mostly static. On the other hand, the external field is driven by the Sun's activity and is often varying rapidly in time (Korte & Mandea, 2019).

For centuries, with the directional preference of the needle of a compass, the Earth's magnetic field has been used by navigators to indicate directions. Currently, although the magnetic orientation remains essential, the geomagnetic field is used to explore the dynamics of the Earth's interior and its space environment, and geomagnetic data are used for surveying, geophysical exploration, mineral exploration, risk mitigation, and other practical applications such as studies of large-scale crustal structures (Love, 2008). To produce all of this information, a global network of magnetic observatories is needed to provide accurate records of the geomagnetic field at fixed locations and over long periods of time. These records are vital to predict geomagnetic field variations. However, it is impractical to accurately predict the field for the very distant future at any point because the Earth's magnetic field is continuously changing. Thus, a mathematical representation of the Earth's main magnetic field and how it is changing is needed.

Since the early days of global exploration, the efforts to represent the changing Earth's magnetic field at the surface was performed by creating magnetic charts by hand. However, the analytical breakthrough in modelling the geomagnetic field came in 1838 when Carl Friedrich Gauss applied the spherical harmonic analysis (SHA) to the Earth's magnetic field. Developed initially to describe Earth's gravitational field by Pierre Simon Laplace, Adrien-Marie Legendre, and Siméon-Denis Poisson (Turner, 2011), SHA has become the standard method of analysis of the geomagnetic field since Gauss time (Kono, 2015). The SHA method provides an efficient method of modelling the geomagnetic field and gives a simple mathematical description of different wavelength characteristics. The separation of internal and external contributions to the field and the upward and downward continuation can be performed correspondingly. This means the SHA technique can calculate the geomagnetic field in three dimensions.

SHA has become the de-facto standard for analysing the global field because of its many advantages (Barton, 1988). By adopting this method, the International Association of Geomagnetism and Aeronomy (IAGA) produced the International Geomagnetic Reference Field (IGRF) for the year 1965, which was the first generation of the IGRF (Kono, 2015). However, the sensitivity of the SHA basis functions at a global scale aligns poorly with the local nature of the geological sources. As a result, it is not suitable for modelling unevenly distributed data or for crustal field modelling (Schott & Thébault, 2011). The global distribution of geomagnetic observatories is highly irregular. Consequently, the model in regions with few data points is strongly influenced by observations in areas with large numbers of

geomagnetic observatories. To address this problem, several methods have been developed to model the geomagnetic field regionally with much denser data sets.

Principally, regional geomagnetic models are developed to provide compass information for navigation, details of secular variation for reducing magnetic survey data to a common epoch with better accuracy and spatial resolution, and information about magnetic anomalies with intermediate wavelength (Barton, 1988). Several approaches to regional modelling of potential fields have been undertaken. The simplest analytical method uses a polynomial expression in latitude and longitude, as has been done by De Santis et al. (2003) with his Italian Geomagnetic Reference Field (ITGRF). However, this technique sacrifices the physical constraint that the field should be derived from a scalar potential satisfying Laplace's equation. Moreover, this method cannot be applied if there are variations in altitude in the data set (Torta et al., 2006). Another technique employed for smaller regions on the globe is the Rectangular Harmonic Analysis (RHA) which was developed by Alldredge (1981). The RHA involves solving Laplace's equation in a Cartesian coordinate system. However, this method is only efficient in modelling short-wavelength features (De Santis et al., 1990). Thus, the RHA is only appropriate for modelling the geomagnetic field on the local scale. The most promising method is probably the Spherical Cap Harmonic Analysis (SCHA) introduced by Haines (1985) to provide a reference field for Canada (Haines, 1985b). The SCHA is developed by solving Laplace's equation over a spherical cap. The solution is developed in terms of two sets of spherical harmonic functions with integer-order but non-integer degree.

SCHA has been widely adopted in various regional models, including reference field models, secular variation, crustal field, external field, and even in applications outside of geomagnetism (Schott & Thébault, 2011). However, the SCHA model still has some shortcomings. The SCHA model at the edge of the cap deteriorates if the data distribution is not uniform (De Santis, 1991). SCHA is also insufficiently well behaved with data recorded at different altitudes; the more significant the altitude difference, the worse the behaviour of the model. This led to the introduction of the Revised Spherical Cap Harmonic Analysis (R-SCHA) by Thébault et al. (2006) to remedy the drawbacks of SCHA. Nevertheless, to compromise the boundary problem in the SCHA, the size of the cap can be enlarged in order to increase the spectral coverage in low harmonics, which is probably more significant (Torta et al., 1992). Then, the gaps in data distribution are filled with IGRF values as in Torta et al. (1992) and Qamili et al. (2010). Many regional reference fields have been calculated using SCHA methods, such as the model for Canada (Haines, 1985b), Scandinavia (Nevanlinna et

al., 1988), Italy (De Santis et al., 1990), Spain (Torta et al., 1993), Albania and South-East Italy (Qamili et al., 2010), and several other models.

In Indonesia, the regional geomagnetic model has been calculated using collocated cokriging and kriging with external drift (Syirojudin et al., 2018). However, this geomagnetic field model does not satisfy the conditions imposed by potential theory, where the field is supposed to be the gradient of a potential. Specifically, the potential theory requires the field to be the gradient of a function that satisfies Laplace’s differential equation, i.e., the curl and divergence of the field be zero (Haines, 1985a). The global model, the IGRF, only incorporated three geomagnetic observatories from Indonesia to model the IGRF-12, i.e., KPG (Kupang), PLR (Pelabuhan Ratu), and TND (Tondano) (Thébault et al., 2015) and only two geomagnetic observatories to calculate the IGRF-11 model, i.e., TND (Tondano) and TUN (Tuntungan) (Finlay et al., 2010). For the IGRF-10, only TND (Tondano), TUN (Tuntungan), and TNG (Tangerang) were included (Macmillan & Maus, 2005). Thus, the IGRF model is not very accurate in Indonesia, and there is a significant discrepancy between the global model and the actual observational values which have been reported by Syirojudin and Bijaksana (2013). On the other hand, the need for accurate geomagnetic field models is increasing as they are essential for navigation and reduction of magnetic surveys. Regional geomagnetic field modelling using SCHA can be used to give a more accurate geomagnetic reference field that satisfies potential theory and utilises denser data in Indonesia.

In SCHA modelling, the spherical cap corresponds to an area defined around a selected point by a predetermined angle θ_0 subtended at the centre of the Earth. The spherical cap half-angle size has to be large enough because some errors will arise when attempting to map wavelengths that are larger than the spherical cap (Fiori, 2020). The area modelled in this thesis is the Indonesian region with a spherical cap half-angle 30° , and the coordinate of the spherical cap pole is at 122°E and 3°S . This spherical cap covers the area, as shown in Figure 1. This thesis is concerned with developing a mathematical model of the geomagnetic field-specific to this area.

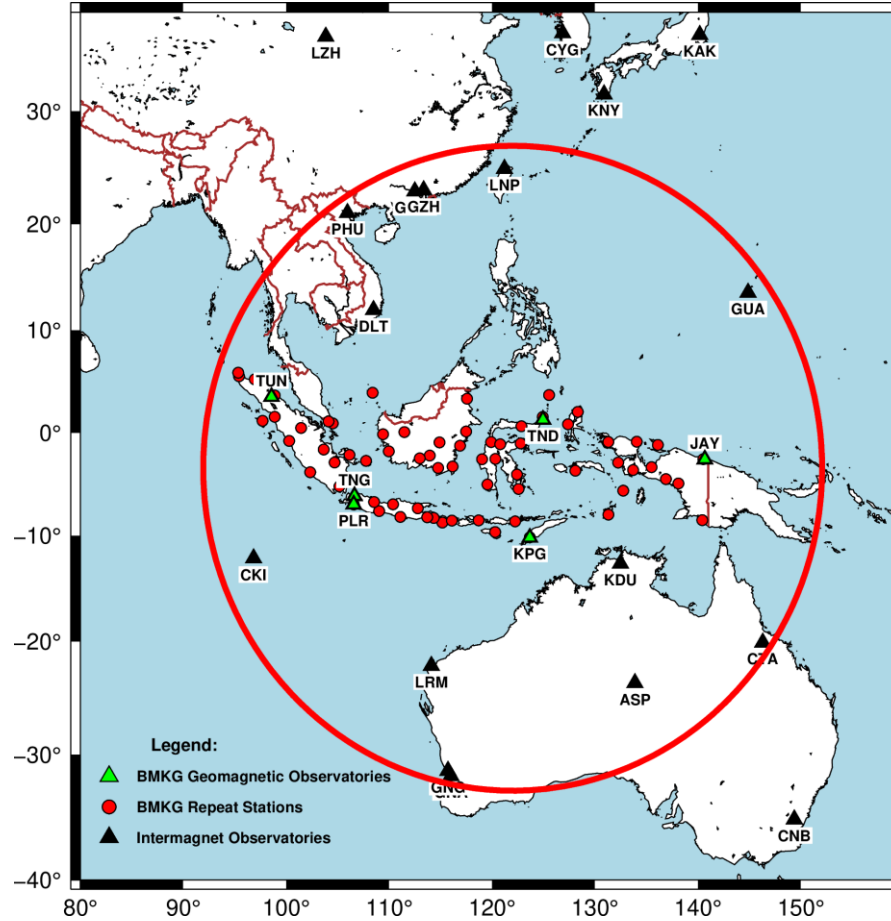


Figure 1. Selected cap region. The spherical cap radius $\theta_0 = 30^\circ$, centered at geographic coordinates 122°E and 3°S .

1.2 Thesis Aims

There are two main aims in this thesis. The first aim is to make available a geomagnetic field model with a better resolution and better accuracy over the Indonesian region than provided by the IGRF. The second is to improve the secular variation forecast in geomagnetic field models used for navigation and reduction of magnetic surveys carried out over the Indonesian region to a common epoch.

1.3 Thesis Structure

This section outlines the thesis structure by briefly describing the contents of each chapter. This thesis is structured as follows:

- **Chapter 1:** An introduction to the motivation and objectives of this study. The history of geomagnetism is introduced first and followed by an overview of efforts to model the geomagnetic field and its secular variation, globally

and regionally. The recent study of geomagnetic modelling in Indonesia is then outlined. Lastly, the region of the research area is described.

- **Chapter 2:** The review of related literature and studies in this study. This starts with an explanation of the origin of geomagnetic field theory and continues with a review of modelling techniques in geomagnetism, both globally using SHA and regionally using SCHA. A review of the IGRF model is described in detail because it is used to compare the results of this study. Finally, the concept of secular variation and its origin is explained.
- **Chapter 3:** The details about the data collected and the methodology employed in this study. The dataset used in this study and the data processing technique are described in this chapter, followed by a discussion of the procedure used in the SCHA modelling and in selecting the best SCHA model spatially. Finally, the temporal modelling using cubic B-splines and secular variation forecasting using autoregressive forecasting are described.
- **Chapter 4:** Results calculated using the methods outlined in Chapter 3. The analysis of the results is divided into the spatial and temporal analysis. In the spatial modelling analysis, the quality of the SCHA model is investigated, followed by an explanation of how the best model was selected and an analysis of the misfit of the selected model. The results of the B-splines calculation and secular variation forecast are described in the temporal modelling analysis section.
- **Chapter 5:** Concluding remarks, including key findings, limitations on this study, and recommendations for future work.

Chapter 2

Literature Review

2.1 The Origin of Geomagnetic Field

After the publication of *De Magnete* by Gilbert in 1600, the geomagnetic field was commonly accepted as being due to the permanent magnetisation of the Earth as a whole. However, in the late nineteenth century, Pierre Curie discovered that a ferromagnetic material loses its magnetic properties at high temperatures. As the temperature below the Earth's crust exceeds the temperature at which ferromagnetism is lost, alternative theories in the origin of the geomagnetic field needed to be considered (McKnight, 1990).

In 1919, Joseph Larmor suggested a self-sustaining dynamo in the Earth's liquid core might generate the Earth's magnetic field. However, it was Walter M. Elsasser and Edward Bullard, who are well acquainted with the magnetohydrodynamic (MHD) theory for the main geomagnetic field and showed how the Earth's magnetic field generated by the motion in the liquid core (Turner, 2011). In this theory, the motion of the mixture of molten iron and nickel in the Earth's outer core due to convection currents generate electric currents and these result in the magnetic field. The convection currents are generated by the heat escaping from the core (Olson, 2015). It was Glatzmaier and Roberts (1995) who produced the first successful simulation of dynamo action, demonstrating that electric currents in the Earth's outer core, generated by the convection currents and the planet's rotation, are the source of the geomagnetic field. They simulated the magnetic field of the Earth and showed that although the geomagnetic field at the Earth's surface is smooth with growing and decaying features, the field in the outer core is relatively unstable (Turner, 2011).

At any location on the Earth's surface, the geomagnetic field can be represented as a vector that has both a size (magnitude) and a direction, as has been described by Hulot et al. (2015) and can be seen in Figure 2. The relations between the components are (Jankowski & Sucksdorff, 1996)

$$\begin{aligned}
F &= \sqrt{X^2 + Y^2 + Z^2} = \sqrt{H^2 + Z^2} \\
H &= \sqrt{X^2 + Y^2} \\
X &= H \cos D \\
Y &= H \sin D = X \tan D \\
Z &= F \sin I = H \tan I \\
D &= \tan^{-1}(Y/X) \\
I &= \tan^{-1}(Z/H)
\end{aligned} \tag{1}$$

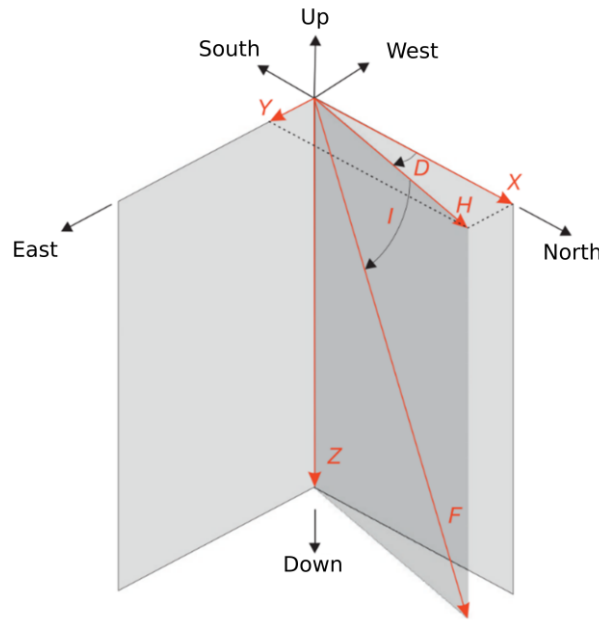


Figure 2. Geomagnetic field components, from Hulot et al. (2015). The magnetic elements X (north), Y (east), and Z (vertical) are the components of the field vector B in an orthogonal coordinate system. The angle between geographic north and the compass needle horizontal direction is denoted as declination (D), and the angle between the field vector and the horizontal direction is denoted as an inclination (I). The vector's horizontal component along the local magnetic meridian is denoted as H , and the magnetic field vector of total intensity is denoted as F .

Commonly, the geomagnetic field is displayed as an *isomagnetic chart* or a contour map of a particular magnetic element with similar values. For the declination component, this chart is termed *isogonic*, *isoclinic* for inclination, *isodynamic* for the equal intensity of a particular component, and *isoporic* for the secular variation (Merrill et al., 1996). Furthermore, Maxwell's equations can be used to represent the geomagnetic field quantitatively in a mathematical model, as will be discussed in the next chapter.

2.2 Geomagnetic Field Modelling

Geomagnetic field charts were hand-drawn from the 17th century, and due to their importance for navigation, these were usually of declination (Barton, 1988). It was Edmund Halley who published the first isogonic map of the Atlantic in his *General Chart of the Variation of the Compass* (1701). This publication is the first magnetic chart, and it shows the importance of the study of geomagnetism (Kono, 2015).

Since then, a large amount of magnetic vector and scalar data have been measured. This comprises data at different ranges of altitude, from the ground to satellite level. Furthermore, several techniques have been proposed to model the geomagnetic measurements, both globally and regionally. Spherical harmonic expansions remain the fundamental tool for modelling the geomagnetic field at a global scale, although other techniques such as wavelets (Holschneider et al., 2003), multi-scale (Maier & Mayer, 2003), and Slepian functions (Simons & Dahlen, 2006) are available.

However, the spherical harmonic basis functions are not well suited for modelling the crustal field nor for use with unevenly distributed data. To solve this problem, several techniques have been proposed to model the geomagnetic measurements regionally. These include polynomial fitting, Rectangular Harmonic Analysis (RHA), Cylindrical Harmonic Analysis (CHA), and Spherical Cap Harmonic Analysis (SCHA). The polynomial fitting might be the most straightforward technique and has been implemented in several regional models, such as the Italian Geomagnetic Reference Field (De Santis et al., 2003) and the New Zealand regional model (McKnight, 1996). Nevertheless, the SCHA technique is probably the most popular regional modelling method because it is the closest relative to SHA (Kono, 2015). In this chapter, the geomagnetic field modelling techniques which will be discussed are the SHA for modelling at a global scale and SCHA for modelling at a regional scale.

2.2.1 Spherical Harmonic Analysis (SHA)

It was Gauss who was applying the spherical harmonic analysis (SHA) to calculate the mathematical model of the Earth's magnetic field for the first time. There are two assumptions. The first of these is that there are no electric currents between the surface of the Earth and the ionosphere in the region where the measurements of geomagnetic fields are generally made. The second assumption is that there are no sources of the magnetic field in this region. For the first assumption, the amount of electric current flowing across the boundary between the Earth and its atmosphere is relatively insignificant, and the electric field changes are negligible. Thus, applying

Maxwell's equation at the surface of the Earth, the curl of magnetic field B equals zero, or

$$\nabla \times B = i \left(\frac{\delta B_z}{\delta y} - \frac{\delta B_y}{\delta z} \right) + j \left(\frac{\delta B_x}{\delta z} - \frac{\delta B_z}{\delta x} \right) + k \left(\frac{\delta B_y}{\delta x} - \frac{\delta B_x}{\delta y} \right) = 0 \quad (2)$$

where i, j , and k represent the three orthogonal directions, and δ indicates the partial derivatives. This equation requires that the field can be obtained from the negative gradient of a scalar potential (U)

$$B = - \left[i \frac{\delta U}{\delta x} + j \frac{\delta U}{\delta y} + k \frac{\delta U}{\delta z} \right] = -\nabla U \quad (3)$$

The second assumption is that there is no source of the magnetic field. Thus, by the other Maxwell's equation, the divergence of the field equals zero

$$\nabla \cdot B = \left[\frac{\delta B_x}{\delta x} + \frac{\delta B_y}{\delta y} + \frac{\delta B_z}{\delta z} \right] = 0 \quad (4)$$

Putting (3) and (4) together, thus

$$\nabla \cdot \nabla U = \nabla^2 U = 0 \quad (5)$$

The equation (5) is read as the Laplacian of scalar U is zero.

The solution of the potential function U for the Earth's main field can be found by solving Laplace's equation in spherical coordinates (r, θ, φ) . As devised by Gauss in 1838, the solution is

$$U = a \sum_{n=1}^{\infty} \left[\left(\frac{r}{a} \right)^n S_n^e + \left(\frac{a}{r} \right)^{n+1} S_n^i \right] \quad (6)$$

Where r is the distance between the observational point and the centre of the Earth, a is the radius of the Earth, and S are the functions called *Legendre polynomials*. The i and e superscript represent the internal and external part of the field, respectively.

From the measurements at the surface of the Earth, essentially all the contribution comes from the internal part of the potential function expansion. Therefore, the solution of the spherical harmonic functions becomes

$$U(r, \theta, \varphi) = a \sum_{n=1}^{\infty} \sum_{m=0}^n \left(\frac{a}{r}\right)^{n+1} (g_n^m \cos m\varphi + h_n^m \sin m\varphi) P_n^m(\cos \theta) \quad (7)$$

where r is the radial distance, θ is the colatitude, φ is the longitude, a is the reference radius of the Earth ($a=6371.2$ km), n is the degree and m is the order of spherical harmonics, and both are integral, $P_n^m(\cos \theta)$ are Schmidt-normalised associated Legendre polynomials, and g_n^m and h_n^m are the Gauss coefficients, which denote the amplitudes of the respective harmonics or amplitude coefficients for the function being fit to the data.

The Schmidt-normalised associated Legendre polynomials are

$$P_n^m(\cos \theta) = \sin^m \theta \frac{\partial^m}{\partial (\cos \theta)^m} (P_n(\cos \theta)) \quad (8)$$

where $P_n(\cos \theta)$ are determined using Rodrigues' formula as (Merrill et al., 1996)

$$P_n(\cos \theta) = \frac{1}{2^n n!} \frac{\partial^n}{\partial (\cos \theta)^n} (\cos^2 \theta - 1)^n$$

and the normalising terms as (Blakely, 1995; Merrill et al., 1996)

$$P_n^m = \begin{cases} \left[\frac{2(n-m)!}{(n+m)!} \right]^{1/2} P_{n,m} & m > 0 \\ P_{n,m} & m = 0 \end{cases}$$

The $P_n^m(\cos \theta)$ are functions of colatitude θ only. These polynomials require that n be greater than or equal to m in size. There are m sine and cosine waves fitted around each latitude circle, which are called sectoral harmonics, and there are zonal harmonics of Legendre polynomial waves around each great circle of longitude. The number of zonal harmonics is $n - m + 1$ if $m > 0$ or just n if $m = 0$ along a great circle of longitude. The shortest wavelength representation along a latitude line is found by dividing 360° by the largest value of m , while along a longitude line, it is found by dividing 360° by the largest value of n (Campbell, 2003). The $\cos m\varphi$ and $\sin m\varphi$ are Fourier harmonic type sinusoidal oscillations around a latitude circle. The product of the $P_n^m(\cos \theta)$ with the $\cos m\varphi$ and $\sin m\varphi$ terms form spherical harmonics of degree n and order m . Then, these harmonics are weighted by their corresponding Gauss coefficients.

The most general solution for U is an infinite summation of such functions over n and m . However, it is impossible to determine an infinite set of coefficients. Therefore, in practice, the series is truncated at some point. The level of complexity of the model is determined by the level of truncation (Malin, 1983).

To determine the number of Gauss coefficients, the number of measurements on the Earth's surface must be known. The total number of g_n^m to degree N is

$$\left[\frac{2 + (N + 1)}{2} \right] N = \frac{N^2 + 3N}{2} \quad (9)$$

while the total number of h_n^m to degree N is

$$\left[\frac{1 + N}{2} \right] N \quad (10)$$

Thus, summing (9) and (10), to determine the potential U up to degree N , requires a total number of measurements (Merrill et al., 1996)

$$(N + 1)^2 - 1 \quad (11)$$

In principle, a higher degree, n , and order, m , and Gauss coefficients can be calculated if more observation data is known. However, in practice, n and m are generally limited to low values. The Gauss coefficients are generally determined using a least-squares best-fit model from the observation data.

As U is the scalar magnetic potential, the magnetic field components of B can be found from the gradient of the potential in spherical coordinates

$$X(\theta, \varphi) = -B_\theta = \frac{1}{r} \frac{\delta U}{\delta \theta} \quad (12)$$

$$Y(\theta, \varphi) = +B_\varphi = \frac{-1}{r \sin \theta} \frac{\delta U}{\delta \varphi} \quad (13)$$

$$Z(\theta, \varphi) = -B_r = \frac{\delta U}{\delta r} \quad (14)$$

where X, Y, Z are positive for the northward, eastward, and downward components, respectively (Campbell, 2003).

Generally, the SHA method is used to determine the magnetic potential on a global scale where the data are spread over the entire globe. However, it is not well suited if the spatial density of the data set is varied. The basis functions of the SHA are global, and its degree must represent the shortest spatial wavelength in the data set. However, if the appropriate spatial sampling rate is only available in a limited region, then the coefficients multiplying these degree terms will not be well constrained (Purucker & Whaler, 2015). Thus, the SHA is not well suited to represent the magnetic potential over a restricted area on the surface of the Earth. In these circumstances, to solve Laplace's equation in spherical coordinates over only a limited area of the Earth, a regional modelling technique such as the Spherical Cap Harmonic Analysis (SCHA) is preferred.

2.2.2 Spherical Cap Harmonic Analysis (SCHA)

Regional geomagnetic modelling is principally used when measurements are restricted to a localised region or when high spatial frequency fields with much fewer parameters are required (Torta, 2019). Spherical Cap Harmonic Analysis (SCHA) was developed by Haines (1985a) to model the field over a small portion of the globe. Among regional modelling methods, SCHA is the closest relative to SHA (Schott & Thébault, 2011). Moreover, according to Feng et al. (2016), the SCHA method has two distinct advantages, the SCHA model satisfies the geomagnetic potential theory, and it can be used to describe the three-dimensional structure of the geomagnetic field. In other words, it is possible to use a downward or upward continuation of the SCHA model to calculate the geomagnetic field at levels other than merely the Earth's surface.

The basis functions of SHA and SCHA consist of associated Legendre functions in colatitude, trigonometric functions in longitude, and powers of radial distance (Haines, 1988). The difference is the associated Legendre functions of SHA have integral degrees, while for SCHA, they have non-integral degrees. SCHA's domain of interest is shown in Figure 3, as has been described by Haines (1985a) and is defined by the half-angle of the spherical cap θ_0 and the radial distances between $r = a$ and $r = b$.

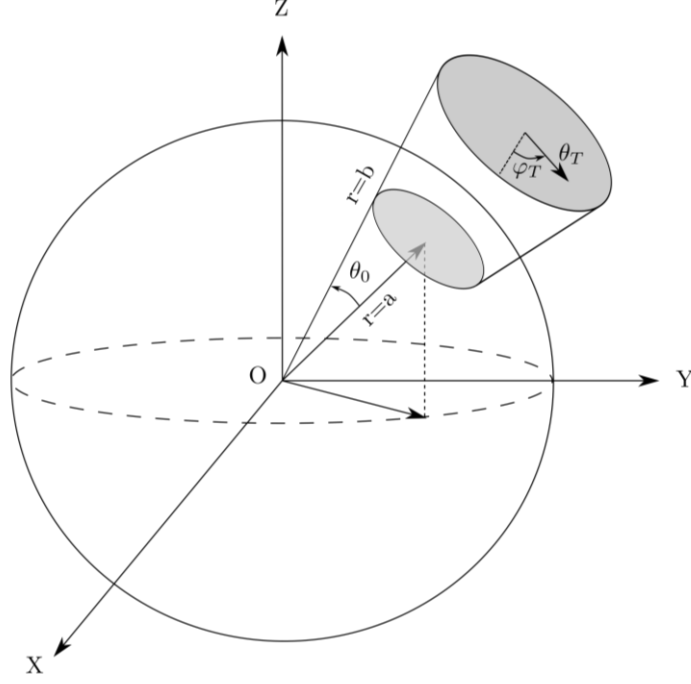


Figure 3. Schematic representation of the spherical cap. The domain of study is bounded by the terrestrial surface $r = a$ and the upper surface $r = b$. X, Y, and Z are the orthogonal geographic reference system. The spherical cap half-angle denoted as θ_0 , while the θ_T and φ_T are the colatitude and longitude of a point relative to the cap centre.

While the spherical harmonic functions of the SHA are orthogonal over the entire globe, they are orthogonal over the cap in SCHA. In SCHA, the boundary conditions must be applied to the edge of the cap in order to solve Laplace's equation for the scalar magnetic potential. The boundary conditions require the continuity of the potential in longitude and regularity at the spherical cap pole (Torta, 2019). In the SCHA model, the scalar magnetic potential U at θ_0 or at the edge of the spherical cap and its derivative in accordance with θ must be arbitrary

$$U(r, \theta_0, \varphi_T) = f(r, \varphi_T) \quad (15)$$

$$\frac{\partial U(r, \theta_0, \varphi_T)}{\partial \theta_T} = g(r, \varphi_T) \quad (16)$$

where f and g are arbitrary functions. The colatitude and longitude (θ_T, φ_T) , in these boundary conditions and the SCHA calculations are in the spherical cap reference frame. As a result, it is essential to transform the site locations and the magnetic field components into this spherical cap reference frame.

An SCHA model has much the same form as that for SHA, but n no longer has to be integral. SCHA for the internal field is based upon the following expression

$$U(r, \theta, \varphi) = a \sum_{m=0}^{\infty} \sum_{k=m}^{\infty} \left(\frac{a}{r}\right)^{n_k(m)+1} (g_k^m \cos(m\varphi_T) + h_k^m \sin(m\varphi_T)) P_{nk(m)}^m(\cos\theta_T) \quad (17)$$

This is similar to the SHA form; the difference is that now the associated Legendre polynomials are real and of integral order (m) but not necessarily of integral degree ($n_k(m)$). The Gauss coefficients no longer describe the best fit to the geocentric dipole, quadrupole, and so on, as in SHA, although they still form the basis of the model field. The coefficients are determined using the same methodology as SHA. The spatial truncation index (k) and order (m) in SCHA describe the sectioning over the spherical cap. This division defines the resolution of the features that can be mapped using SCHA (Fiori, 2020).

The Schmidt normalised associated Legendre polynomials in the solution above are

$$P_n^m(\cos\theta) = \sum_{k=0}^{\infty} A_k(m, n) \left\{ \frac{1 - \cos\theta}{2} \right\}^k \quad (18)$$

and its derivatives when $m = 0$ are

$$\frac{dP_n^m(\cos\theta)}{d\theta} = \frac{\sin\theta}{2} \sum_{k=1}^{\infty} k A_k(m, n) \sin^{2(k-1)}\left(\frac{\theta}{2}\right) \quad (19)$$

and when $m > 0$ are

$$\frac{dP_n^m(\cos\theta)}{d\theta} = \frac{\sin\theta}{2} \sum_{k=1}^{\infty} k A_k(m, n) \sin^{2(k-1)}\left(\frac{\theta}{2}\right) + \cos\theta \left[\frac{m}{\sin\theta} P_n^m(\cos\theta) \right] \quad (20)$$

with

$$A_0(m, n) = K(m, n) \sin^m\theta_T$$

$$A_k(m, n) = \frac{(k+m-1)(k+m) - n(n+1)}{k(k+m)} A_{k-1}(m, n)$$

In this case, $K(m, n)$ is a normalising factor and obtained by

$$K_n^m = 1 \text{ if } m=0$$

$$K_n^m = \frac{2^{-m}}{\sqrt{m\pi}} \left(\frac{n+m}{n-m} \right)^{\frac{n+1}{4} + \frac{1}{4}} P^{\frac{m}{2}} \exp(e_1 + e_2 + \dots) \text{ if } m>0$$

Within this P and e_1, e_2 can be obtained from

$$\begin{aligned} P &= \left(\frac{n}{m} \right)^2 - 1 \\ e_1 &= -\frac{1}{12m} \left(1 + \frac{1}{P} \right) \\ e_2 &= \frac{1}{360m^3} \left(1 + \frac{3}{P^2} + \frac{4}{P^3} \right) \end{aligned}$$

For any combination of (m, k) , the Associated Legendre Polynomials are dependent on degree, $n_k(m)$ and colatitude within the cap, θ_T . Also, the boundary conditions (15) and (16) are equivalent to

$$P_{n_k(m)}^m(\cos \theta_0) = 0, \quad \text{for } k - m = \text{odd} \quad (21)$$

$$\frac{\partial P_{n_k(m)}^m(\cos \theta_0)}{d\theta_T} = 0, \quad \text{for } k - m = \text{even} \quad (22)$$

A numerical solution for all $n_k(m)$ that satisfies one of equations (21) or (22) must be carried out in determining the non-integral degrees of the spherical harmonic functions in SCHA. Any $n_k(m)$ can satisfy only one of the boundary conditions. Thus, the solution comprises two infinite sets of spherical harmonic functions which are mutually orthogonal within themselves but not orthogonal to the other. Nevertheless, these two series can still be differentiated term by term to find the different magnetic field components as follows

$$\begin{aligned}
B_x &= \frac{1}{r} \frac{\delta U(r, \theta, \varphi)}{\delta \theta} \\
&= \sum_{m=0}^{\infty} \sum_{k=m}^{\infty} \left(\frac{a}{r}\right)^{n_k(m)+2} \cdot \frac{\partial P_{n_k(m)}^m(\cos \theta_T)}{\partial \theta_T} \{g_k^m \cos m\varphi_T + h_k^m \sin m\varphi_T\} \\
B_y &= -\frac{1}{r \sin \theta} \frac{\delta U(r, \theta, \varphi)}{\delta \varphi} \\
&= \sum_{m=0}^{\infty} \sum_{k=m}^{\infty} m \left(\frac{a}{r}\right)^{n_k(m)+2} \cdot \frac{P_{n_k(m)}^m(\cos \theta_T)}{\sin \theta_T} \{g_k^m \sin m\varphi_T - h_k^m \cos m\varphi_T\} \quad (23) \\
B_z &= \frac{\delta U(r, \theta, \varphi)}{\delta r} \\
&= -\sum_{m=0}^{\infty} \sum_{k=m}^{\infty} (n_k(m) + 1) \left(\frac{a}{r}\right)^{n_k(m)+2} \cdot P_{n_k(m)}^m(\cos \theta_T) \{g_k^m \cos m\varphi_T \\
&\quad + h_k^m \sin m\varphi_T\}
\end{aligned}$$

The SCHA technique must be used with special care, especially when the cap is small. The result of the SCHA technique deteriorates as the cap becomes smaller. This is because a high degree expansion is needed in order to include large wavelengths over a small-cap and can cause a problem if not enough data are available (Verbanac, 2007). As a rule, the cap region must be at least coincident with the spatial content of the features of the field being represented (Torta, 2019). This problem has been solved by Torta et al. (1992) by increasing the cap half-angle to extend the wavelength content of the basis functions artificially. Nevertheless, beyond $\theta = 20^\circ$, the SCHA was relatively insensitive to the chosen cap size (Düzgüt & Malin, 2000).

The SCHA technique has been widely used in numerous problems of physics in general and of geophysics in particular, such as regional secular variation (Korte & Haak, 2000; Torta et al., 1992), the crustal magnetic anomaly field (De Santis et al., 1989), ionospheric plasma flow (Fiori et al., 2014), and even modelling sea level data (Hwang & Chen, 1997). For the main-field, several reference models have been developed in different locations using SCHA, e.g., Canada (Haines, 1985b), Scandinavia (Nevanlinna et al., 1988), Italy (De Santis et al., 1990), Spain (Torta et al., 1993), Antarctica (Torta et al., 2006), the Pacific Ocean (Toh et al., 2007), and

Albania and South-East Italy (Qamili et al., 2010). Moreover, the SCHA technique has been developed into several modelling techniques, e.g., TOSCA (Translated Origin Spherical Cap Harmonic Analysis) by De Santis (1991), R-SCHA, the revised version of the SCHA, which is suitable to model the multilevel magnetic data (Thébault et al., 2006) and R-SCHA2D, an alternative version of the SCHA and R-SCHA which is suitable for modelling geomagnetic data at a quasi-constant altitude (Thébault, 2008).

In this research, the results of the SCHA modelling are compared with the currently available models in the research area. Some regional models are already available, such as the reference maps by Syirojudin et al. (2018). However, these are not based on potential theory, where the curl and divergence of the field be zero. In the Indonesian region, the most widely used reference model is the IGRF (International Geomagnetic Reference Field).

2.3 The IGRF Model

The International Geomagnetic Reference Field (IGRF) is a series of mathematical models used to describe the large-scale part of Earth's internal magnetic field. This model is valid between epochs 1900 A.D. until the present and also gives the secular variation prediction for the next five years. The IGRF is used to study the dynamics of the Earth's core field, space weather, local magnetic anomalies at the Earth's crust, and as a source of orientation information (Thébault et al., 2015). Nevertheless, the main purpose of the IGRF is to predict the internal main field at the epoch of validity and the next five years to come (Hulot et al., 2015).

The internal main field has the deepest origin inside the Earth and represents the part of the Earth's field that must be removed from the observation data to obtain the crustal contribution. This crustal contribution is also known as the *residual crustal field* or *surface magnetic anomalies* and is essential to study the thermal states, tectonic and geological structures, and the geodynamic of the Earth's crust. The quantitative definition of the surface magnetic anomalies, at given point P , is as follows (Lanza & Meloni, 2006)

$$B_a(P) = B(P, t) - B_m(P, t) - B_e(P, t) \quad (24)$$

where B_a denotes the crustal magnetic anomalies, B denotes the measured field at given point P at the time t , B_m denotes the main field at that point and time, and B_e denotes the external field at that point and time.

The external field is obtained from a magnetic observatory or an independent simultaneously recording instrument that operates at a fixed location near the survey area and during the survey duration. Meanwhile, the main field is obtained from the internal main field model, such as the IGRF or the other geomagnetic reference field models. This correction or reduction of the survey data will give better results if the internal main field model used is more accurate.

The IGRF has been endorsed by the International Association of Geomagnetism and Aeronomy (IAGA) since 1965 and updated at five years intervals. The latest version of the IGRF model is the IGRF-13. The IGRF model is formulated using the SHA method up to degree 10 for the epochs up to and including 1995.0 and up to degree 13 for the epoch 2000.0 onwards. For the predictive secular variation on all generations of the IGRF, the model is up to degree 8 (Hulot et al., 2015). However, the secular variation between two epochs is assumed to vary linearly with time.

There are six geomagnetic observatories operated by Badan Meteorologi Klimatologi dan Geofisika (BMKG) in Indonesia. They are TUN (Tuntungan), TNG (Tangerang), PLR (Pelabuhan Ratu), KPG (Kupang), TND (Tondano) and JAY (Jayapura). However, not all geomagnetic observatories are used in calculating the IGRF model. There are only three geomagnetic observatories from Indonesia, which are used to model the IGRF-12, i.e., KPG, PLR, and TND (Thébault et al., 2015). For the IGRF-11, according to Finlay et al. (2010), there are only two geomagnetic observatories from Indonesia, which are used to calculate the model, i.e., TND and TUN. For the IGRF-10, it is only TND, TUN, and TNG (Macmillan & Maus, 2005). As yet, there is no publication available on the IGRF-13.

Although widely used, the IGRF has many limitations. There are two types of errors that limit the accuracy of the IGRF. The first error is the error of omission, and it is caused by the field that the IGRF is not attempting to model. This error is dominated by the crustal field, and its root mean square (RMS) values are estimated to be within 200 – 300 nT. The second error is caused by the difference between the IGRF and the part of the field that it is attempting to model, and this error is called the error of commission. The RMS value of this error is thought to be within 50 – 300 nT of the true value, but when satellite data are available, the RMS is thought to be within 5 - 10 nT. The true value, in this case, is the internal field up to a maximum spherical degree at the Earth's surface. This error varies with time and can be estimated by comparing different generations of the IGRF on the same date for both (Macmillan & Finlay, 2011).

As the geomagnetic field generated in the Earth's outer core continuously changes, all of the geomagnetic reference field models must be revised regularly, whether it is a global reference field model such as IGRF or a regional reference field model. This change is called the secular variation and occurs slowly on timescales of years to centuries (Jackson & Finlay, 2015).

2.4 Secular Variation

The geomagnetic field varies on an extensive range of scales in space and time domains. On the time domain, the geomagnetic field is changing on a range of timescales from milliseconds to millions of years. The milliseconds to days variations are associated with the effect of the Sun on the main field, e.g., geomagnetic storms. On the timescales from years to millions of years, this variation is called *secular variation* (SV) (Merrill et al., 1996). The approximate spectrum of the geomagnetic field, as has been described by Constable and Constable (2004), is shown in Figure 4. The secular variation of the main geomagnetic field has been extensively studied, and it is an essential constraint on geodynamo models (Demetrescu & Dobric, 2005). The source of the secular variation is believed to originate within the Earth's core.

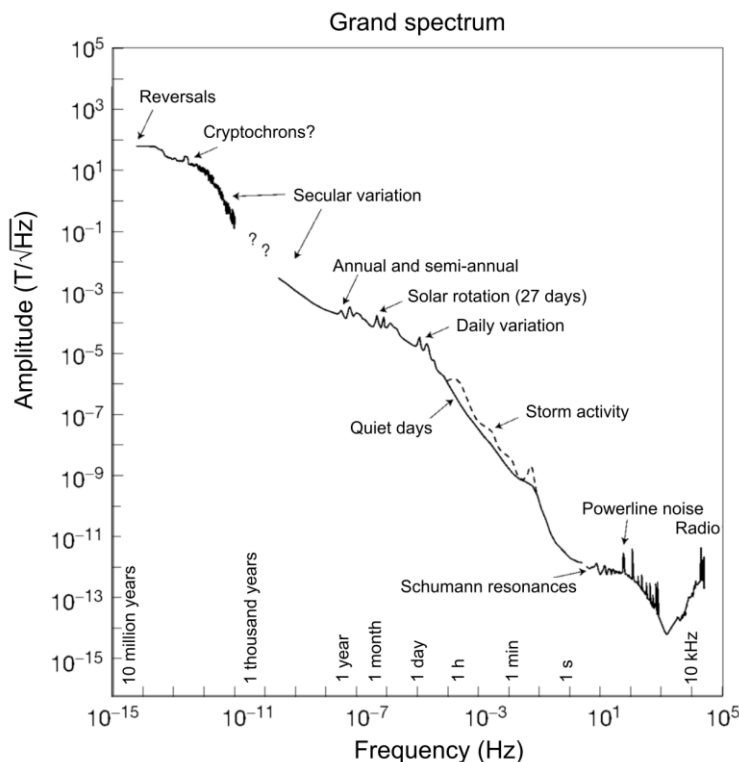


Figure 4. An approximate grand spectrum of the geomagnetic field, reproduced from Constable and Constable (2004). At periods longer than a few months, the internal variations correlated with the Earth's fluid core motions dominate the geomagnetic field. The external magnetic field has an energy between several hours to several second bands. The human-made sources dominate at the highest frequencies.

Several methods have been developed to calculate the secular variation within the time range of an IGRF model, i.e., interpolation of the model. One of the most popular methods is by utilising B-splines basis functions. Also, there are several methods developed to forecast the secular variation, i.e., extrapolation of the model. However, the secular variation is not constant in time, and it can change rapidly on a time-scale of only 1 or 2 years because of rapid and strong acceleration. This limits forecasting secular variation to a relatively short period (Thébault et al., 2015). Nevertheless, many extensive efforts have been made in forecasting secular variation. In this chapter, the discussion of the secular variation will start by describing its origin and then the efforts to predict its values.

2.4.1 Origin of Secular Variation

The first studies of secular variation were initiated by the observation of a change in declination over time in London. The first measurement of declination was accurately conducted in 1580. However, in 1635, Henry Gellibrand repeated the measurement and found that the declination had changed by about 7° . This marks the founding of secular variation (Kono, 2015). The first theory that tries to explain the secular variation was hypothesised by Halley in 1692. He hypothesised that the change in declination could be explained by westward drift caused by a fluid layer deep in the Earth's interior (Bloxham & Jackson, 1991). Today, it is known that secular variation results from a number of factors, including a change in the magnitude of the main dipole part of the field, a shift in the alignment of the dipole axis, and changes in the nondipole parts of the main field (Campbell, 2003). However, another theory has been proposed by Ryskin (2009), who suggested that the geomagnetic secular variation originates from ocean flow.

As indicated above, secular variation can be separated into two parts, i.e. the dipole part and the non-dipole part. The dipole part of the main field is moving slowly by nearly $0.1^\circ/\text{year}$ on a motion called *precessional drift*. Nowadays, the dipole north pole (the geomagnetic pole) is moving about 18 km northward and 5 km westward each year (Campbell, 2003). Moreover, since the time of Gauss's analysis in 1835, the intensity of the dipole field has also decreased at the rate of about 5% per century (Merrill et al., 1996). For the non-dipole part, there are two kinds of variation. The first part is a drifting part that consists mainly of low harmonics (degree < 3) (Creer, 1988). This part seems to be drifting westward at a rate of about 0.2° or 0.3° each year and is faster than the dipole part movement; thus, this movement is called *westward drift*. The other part is the standing part which remains stationary but

exhibits intensity fluctuations. These secular variation models are commonly referred to as the drifting and standing model for secular variation (Merrill et al., 1996).

The westward motion of the geomagnetic field at the Earth's surface may be the most noticeable aspect of geomagnetic secular variation. The westward motion of the field can be seen by following the motion of the agonic lines where declination is equal to zero (Jackson & Finlay, 2015). The change in westward drift is highly correlated with the change in the Earth's rotation speed (Campbell, 2003). Apart from the slow westward drift, a relatively sharp change in the time variation of the geomagnetic field occurred over a few months in 1969, which was later called a *geomagnetic jerk* (Campbell, 2003).

A geomagnetic jerk is a feature in the time evolution of the main field expressed as a sudden change in the slope of the secular variation (first time derivative), or a step-like change in the secular acceleration (second time derivative) of the field, or a Dirac distribution of the third time derivative of the field (Demetrescu & Dobric, 2005). Since the beginning of the twentieth century, several geomagnetic jerks have been identified, e.g., in 1901, 1913, 1925, 1969, 1978, 1991, 1999, 2003, 2011, and 2014. Since 1999, recent jerks have been identified to occur every 3 – 4 years. This suggests that some unknown oscillatory phenomena cause the geomagnetic jerks within the core (Chulliat & Maus, 2014). Geomagnetic jerks are not seen in all components at all observatories (Mandea et al., 2010). Moreover, they do not coincide over the globe; some jerks may be identified in different epochs from one region to another (Brown et al., 2013). Even in a limited region, e.g., China, jerks have some offset times (Feng et al., 2018). The most apparent change of slope is in the Atlantic-African region, where the secular variation is also the largest (Chulliat et al., 2010). Recent findings by Torta et al. (2015) show that the rate of change of the secular variation in Europe after the jerk in 2014 has increased, and its value is becoming closer to that of the Atlantic-African region. All of these recent findings show that the secular variation is something unpredictable, as has been discussed by Malin (1985). Thus, the prediction of the secular variation from the IGRF and related models might be more impoverished than currently expected in the upcoming years.

2.4.2 Secular Variation Prediction

The challenge in secular variation prediction is primarily caused by the core processes controlling the secular variation, which are not sufficiently well understood. As unpredictable abrupt geomagnetic secular variation changes cause the prediction of the field to become complicated, this subject is crucial. The prediction of secular variation is essential for a geophysicist who needs to produce navigational charts that

will be valid for several years after their date of publication. Another usage of secular variation prediction is for reducing data from magnetic surveys conducted in the future after the geomagnetic reference field model is published. The secular variation prediction can be separated into global and local schemes (Malin, 1985).

The global schemes vary from the physical models, such as the westward drift hypothesis, drifting and standing parts separation of the field, changing dipoles, and current-loop models (Malin, 1985). As the observed secular variation of the internal geomagnetic field is a consequence of the motions in the liquid metal outer core, to model the mechanism underlying this variation, the mathematical framework of magnetohydrodynamics (MHD) must be employed. According to Beggan and Whaler (2010), several secular variation predictions which incorporate physical approximations have been developed since 2005. Some models, such as those in Sun et al. (2007) and Fournier et al. (2007), outline frameworks that assimilate the observed magnetic field data into physical MHD models. Another method is assimilating historical field data into numerical dynamo models to investigate the improvements which can be made to the forecasted field models (Kuang et al., 2009). Another model by Beggan and Whaler (2010) presents a forecast based upon a steady core flow model generated from satellite magnetic data. This physically constrained secular variation forecasting model is useful in producing more accurate predictions.

The local schemes are mostly concerned with extrapolation of the data without it being constrained physically. Currently, the secular variation of the global reference field model (IGRF) uses a combination of polynomial extrapolation of satellite data and linear prediction filters applied to the observatory data (Beggan & Whaler, 2010). Therefore, the secular variation prediction of the IGRF might not be very accurate as the secular variation is not linear. Kotzé (2019) reported that the linear prediction of the IGRF model was potentially not accurate in South Africa, where unpredictable abrupt geomagnetic secular variation changes characterise the region.

Nevertheless, regional geomagnetic modelling, incorporating physical constraints in the secular variation prediction, is far from simple. Thus, most regional geomagnetic models do not appeal to any physical modelling constraints, for example, using polynomial extrapolation from previous measurements such as that done by Qamili et al. (2010). Some improvements have been made by calculating the secular variation prediction using an autoregressive forecasting technique such as that of De Santis et al. (2003) on the Italian Geomagnetic Reference Field, and Metodiev and Trifonova (2017) on the Bulgarian Geomagnetic Reference Field. Pavón-Carrasco et al. (2013) also provide secular variation prediction using a bootstrap algorithm.

Chapter 3

Data and Methods

3.1 Data Sets and Data Processing

This chapter describes the data used in this research and the method used to remove data of insufficient quality. The data used to calculate a mathematical model of the geomagnetic field are typically made up of measurements of the three orthogonal strength components, i.e., north (X), east (Y), and vertical (Z); the total field strength and two angles, i.e., declination (D), inclination (I), and total intensity (F); or alternatively using two strength components and an angle, i.e., horizontal (H), vertical (Z), and declination (D). The relationship between these geomagnetic components can be seen in Figure 2.

The data used in this study are the X, Y, and Z components compiled from 68 geomagnetic repeat stations in Indonesia, covering the period 1985 - 2015 from BMKG (Badan Meteorologi Klimatologi dan Geofisika). However, two locations are not included in the calculation and will be used later to validate the model. These validation observatories are Bandar Lampung (105.175°E, 5.240°S) and Tual (132.735°E, 5.662°S). The repeat station data are combined with definitive data from five of BMKG's geomagnetic observatories from 1985 - 2019. To complete the dataset, the definitive data from a similar period from 13 INTERMAGNET observatories are used. In spherical cap harmonic modelling it is essential to have data around the edges of the chosen spherical cap. Because of the irregularities in the data distribution in the research area, the synthetic cartesian X, Y, and Z components at sea level at 17 fixed points are calculated from IGRF-13 at five-year intervals from 1985 – 2015. The IGRF-13 is also used to remove the main field before applying the techniques to calculate the regional field. Figure 5 shows the data locations used in this research. For checking and filtering to remove data of insufficient quality, the Median Absolute Deviation (MAD) (Hampel, 1974) from the median is used.

In this research, the coordinate of the spherical cap pole is at 122°E and 3°S, and the spherical cap half-angle is 30°. This covers the Indonesian region, as shown in Figure 5. According to Torta et al. (2006), the use of a large spherical cap is necessary to obtain the main field and its secular variation; otherwise, the results are unrealistic. Research by Qamili et al. (2010) used a half-angle of 16°, although their data mostly concentrated in the central 4° half-angle. Torta et al. (1992) and Torta and De Santis

(1996) explain that the RMS residual of the SCHA model decreases as the cap half-angle increases. However, Düzgit and Malin (2000) obtained different results and reported that the quality of the fit of the SCHA model is relatively insensitive to the cap radius. Nevertheless, the cap half-angle chosen in this research is the most optimal as it covers the Indonesian region, and there are several geomagnetic observatories located near the boundary of the cap.

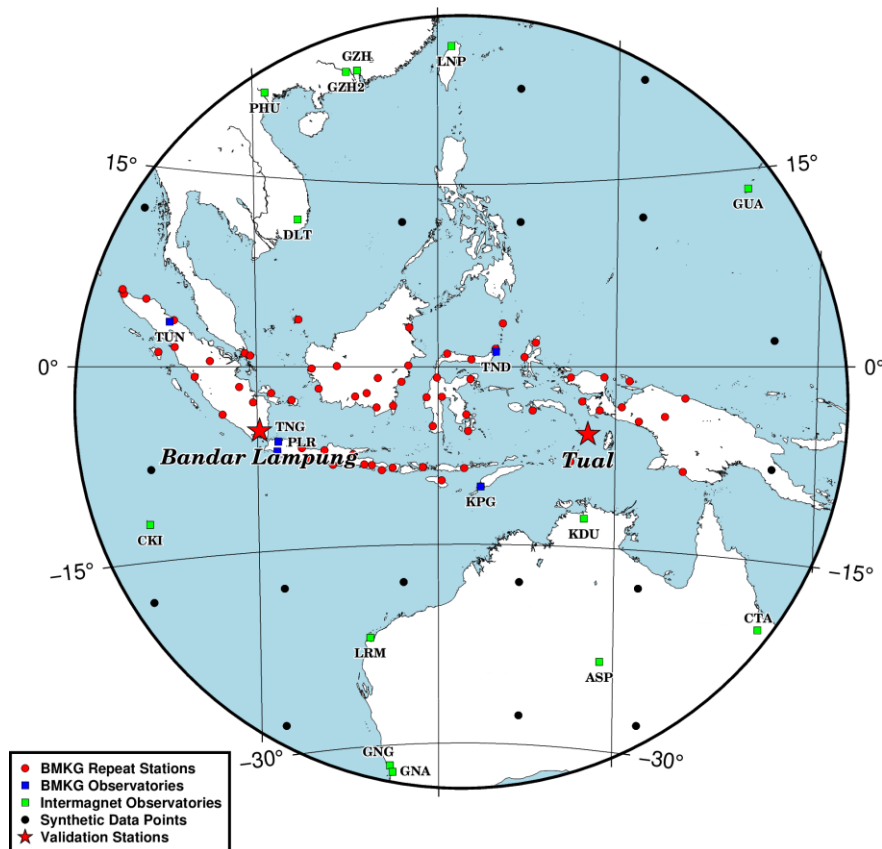


Figure 5. Data distribution used in this research. The data are compiled from 68 geomagnetic repeat stations (red circle), 5 BMKG's geomagnetic observatories (blue square), 13 INTERMAGNET observatories (green square), and 17 fixed synthetic data points (black circle) derived from IGRF-13. Two validation stations are shown as red stars.

3.1.1 Geomagnetic Observatories Data

At present, there are six geomagnetic observatories in Indonesia operated by BMKG: Tangerang (TNG), Tuntungan (TUN), Tondano (TND), Pelabuhan Ratu (PLR), Kupang (KPG), and Jayapura (JAY). However, the data continuity in these observatories is not very good, and there are many gaps in the annual means data. Furthermore, there is no available definitive data at JAY due to it being a new observatory, and the data being generated is still unstable. As a result, the JAY observatory is not used in this research.

There are 13 INTERMAGNET observatories used in this research: Dalat (DLT), Kakadu (KDU), Learmonth (LRM), Cocos Keeling Islands (CKI), Phu Thuy (PHU), Gingen (GNG), Alice Springs (ASP), Guam (GUA), Gwangara (GNA), Luning (LNP), Charters Towers (CTA), Guangzhou (GZH), and Zhaoqing (GZH2). The definitive annual means from INTERMAGNET observatories are obtained from World Data Centres for Geomagnetism, Edinburgh (<http://www.wdc.bgs.ac.uk>).

The data used extend from 1985 – 2019. The data from 1985 – 2015 are used to calculate the models, while the data from 2016 and forward are used to validate the secular variation forecasting. In the internal field modelling, any external geomagnetic activity must be removed as it can affect the accuracy of the model calculation. The data used here are the definitive annual means, and it is assumed that the annual means calculations eliminate the external geomagnetic field.

3.1.2 Repeat Stations Data

Geomagnetic repeat station surveys are carried out in many countries to map the geomagnetic field components and their secular variation regionally. In Indonesia, the geomagnetic repeat station surveys have been carried out by BMKG since 1960. Limited surveys were conducted in 1960 - 1962 and 1970 - 1974. Since 1985, BMKG has conducted a regular survey every five years (Syirojudin & Bijaksana, 2013).

The repeat station data used in this research are the data from 1985 – 2015 because the data before this year are irregularly measured. However, the number of repeat stations varies in every epoch. Nevertheless, the use of repeat stations will improve the regional geomagnetic model because it covers the research area more densely than the geomagnetic observatories used by the IGRF model. Table 1 shows the number of repeat stations in each measurement year.

Table 1. The number of repeat stations surveyed by BMKG

Year	Number of Stations
1985	51
1990	48
1995	49
2000	50
2005	53
2010	53
2015	68

3.1.3 The IGRF Model

As a model comparison, the global geomagnetic field model is needed. The IGRF-13 (<http://www.geomag.bgs.ac.uk/research/modelling/IGRF.html>) is preferred because this model is endorsed by the International Association of Geomagnetism and Aeronomy (IAGA) and is used widely. To obtain the IGRF model in such a large and dense grid, the Geomag version 7.0 is used. This is software from the National Centers for Environmental Information (NCEI), an agency operated by the National Oceanic and Atmospheric Administration (NOAA). This software is already integrated with the IGRF-13 coefficients. The IGRF model is also used to remove the global reference field before performing the SCHA technique.

Because of the low coverage of the research area with geomagnetic data, especially at the borders of the cap, the synthetic cartesian X, Y, and Z components at sea level at 17 fixed points are calculated from IGRF-13 and included in the calculation of the SCHA model. The synthetic data points are calculated at five-year intervals from 1985.5 – 2015.5. The locations of these points can be seen in Figure 5 as the black dots. Qamili et al. (2010) used a similar method when developing the regional geomagnetic model for Albania and South-East Italy and obtained satisfying results.

The data point locations are chosen such that they fill the empty area inside the cap border but are not too close to actual observatories. This is because the purpose of the addition of synthetic data points is to improve the data distribution without significantly influencing the model at the observatory sites (Verbanac, 2007). The addition of synthetic data points can minimise the edge effects and maximise the reliable coverage of the regional model.

3.1.4 Outlier Detection and Removal

In the process of data acquisition, there is a possibility that outliers exist. These data points that deviate significantly from other observation data may be due to artificial magnetic disturbances, inaccuracy in operations, or natural short-term external geomagnetic activity. The first two cases should be identified and removed from the observational data as they do not relate to the physical processes of the geomagnetic field. Moreover, in the internal field modelling, all of these outliers, including those caused by external geomagnetic activity, must be removed as they can affect the accuracy of the model calculation.

There are various statistical methods of identifying and rejecting outliers in time series data. One method that has been applied in the geomagnetic field context is the Median Absolute Deviation (MAD) from the median, defined as

$$MAD = \text{median}(|x_i - \text{median}(x)|) \quad (25)$$

for a set of observations x_i (Hampel, 1974). The rejection criterion of a value x_i is

$$\frac{|x_i - \text{median}(x)|}{MAD} > \alpha \quad (26)$$

where α is a user-specified threshold. Although this threshold determination is inevitably subjective, Leys et al. (2013) strongly recommended using 2.5 as the threshold when using the MAD method for outlier detection. The MAD method has been applied by Cox et al. (2018) for detecting outliers in geomagnetic observatory data when they were developing MagPySV software. A sample of the observation data with outlier detection is shown in Figure 6.

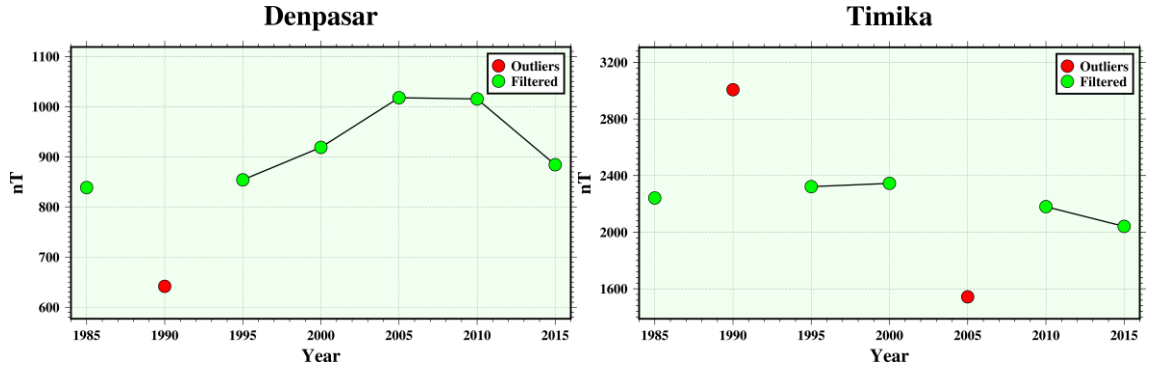


Figure 6. Y Component of the magnetic field at Denpasar and Timika repeat stations (black line with green dots), with identified outliers (red dot). A similar outlier detection method is used on all components in all data.

After outlier detection using the MAD method, the amount of observational data is reduced. The most significant decrease is in 2005 when there are 16 outliers rejected. The complete information on data rejection for each geomagnetic component is shown in Table 2. Data from the validation observatories are also removed in each epoch to be used later to validate the final model.

Table 2. Number of observation data points before and after outlier rejection using the MAD method including two validation locations

Year	Before Outlier Rejection	After Outlier Rejection	Outlier
1985	57	53	4
1990	55	47	8
1995	56	52	4
2000	58	57	1
2005	61	45	16
2010	63	60	3
2015	80	69	11

3.2 Modelling Procedure

Overall, the procedure for regional geomagnetic modelling using the SCHA method is started by calculating the associated Legendre functions. After that, the non-integral degrees ($n_k(m)$) of the SCHA can be determined by numerically solving the boundary conditions. Because the observational data are in geographic coordinates, the system must be rotated to the new spherical cap coordinate system. Then, to increase the rate of convergence in the calculation and to reduce the error beyond the spherical cap boundary, the global reference field is subtracted from the observational data. Lastly, the Gauss coefficients can be determined using simple matrix inversion. More detailed explanations of each step are given below.

3.2.1 Calculation of the Associated Legendre Function and the SCHA Degree

The first step in SCHA modelling is calculating the associated Legendre functions and the SCHA non-integer degree. The associated Legendre functions $P_n^m(\cos\theta)$ can be solved using equations (18), (19), and (20). Simultaneously, the SCHA non-integer degree $n_k(m)$ is solved numerically that separately satisfy equation (21) and (22). Four decimal places are used to speed up the calculations, as has been done by Haines (1985). A slight error in n only affects the orthogonality of the relevant basis functions of the solutions. The decimal place required in the $n_k(m)$ depends on the half angle θ_0 of the spherical cap. Fewer decimal places are required to obtain large $n_k(m)$ at high values of the spherical truncation index k in order to maintain the orthogonality of the functions to an acceptable accuracy (Haines, 1988). The values of $n_k(m)$ up to $k = 9$ for $\theta_0 = 30^\circ$ are given in Table 3 below. The calculations of the $P_n^m(\cos\theta)$ and $n_k(m)$ are performed using *Matlab* software.

Those values for which $k - m$ is odd (the second, fourth, sixth, ... values) are the values of n for which the function itself is zero at $\theta = 30^\circ$ or satisfying equation (21). Correspondingly, those values for which $k - m$ is even (the first, third, fifth, ... values) are the values of n for which the derivative of the function is zero at $\theta = 30^\circ$ or satisfying equation (22). The $P_n^m(\cos\theta)$ and $n_k(m)$ are the functions of the spherical cap size. Therefore, these values will be similar for all regions as long as the spherical cap size is the same.

Table 3. $n_k(m)$

m										
k	0	1	2	3	4	5	6	7	8	9
0	0.0000									
1	4.0837	3.1196								
2	6.8354	6.8354	5.4928							
3	10.0386	9.7121	9.3733	7.7524						
4	12.9083	12.9083	12.372	11.8074	9.9589					
5	16.0248	15.8215	15.6154	14.9180	14.1778	12.1334				
6	18.9364	18.9364	18.5830	18.2219	17.3910	16.5041	14.2861			
7	22.0183	21.8702	21.7210	21.2461	20.7588	19.8121	18.7979	16.4228		
8	24.9514	24.9514	24.6858	24.4170	23.8390	23.2437	22.1940	21.0663	18.5472	
9	28.0145	27.8979	27.7807	27.4166	27.0470	26.3784	25.6881	24.5451	23.3144	20.6618

3.2.2 Coordinate Rotation

A conical section with a radius θ_0 constructs the spherical cap from a sphere and in consideration of the outer slab of the section which extends from the radius r_a to r_b as has been described by Fiori (2020) and can be seen in Figure 7. A spherical cap coordinate system is determined by the centre of the outer shell of the sphere of the spherical cap with θ representing colatitude and φ representing longitude in the spherical cap coordinate system. The colatitude spans from 0° to θ_0 , while the longitude spans the full 0° to 360° . In this research $r_a = r_b$ and the spherical cap reduces to a thin spherical cap shell. The geomagnetic measurements data positions and components on the Earth's surface are usually given in a geographic coordinate system. Therefore, it is necessary to transform both coordinate locations and the vector-pointing directions to the spherical cap coordinate system to simplify calculations.

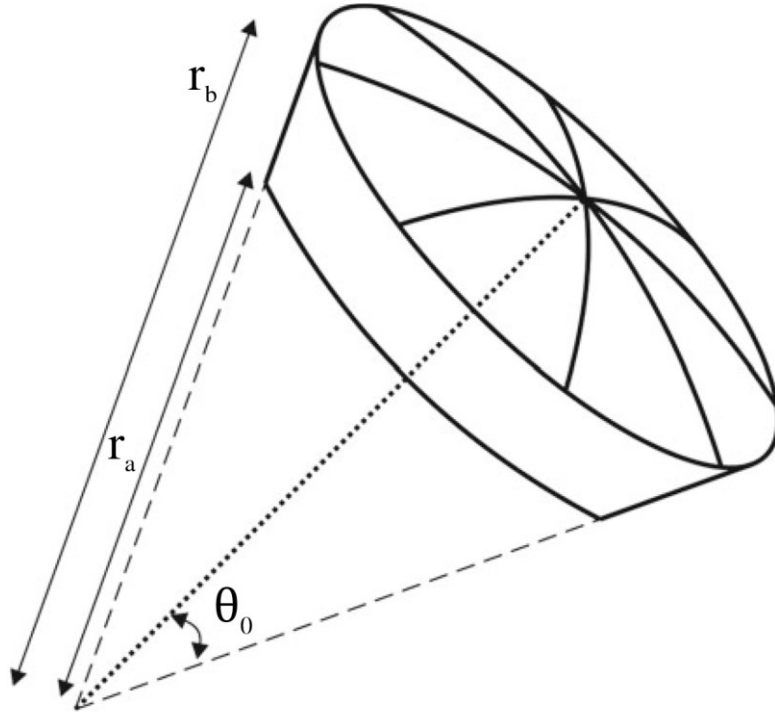


Figure 7. Spherical cap structure, from Fiori (2020). The domain of study is bounded by the terrestrial surface r_a and the upper surface r_b with spherical cap half-angle θ_0 .

From the spherical cap coordinate system, the transformed longitude and colatitude for any data site can be found by considering a triangle on the Earth's surface with the centre of the spherical cap, the North geographic pole, and the data points as shown in Figure 8.

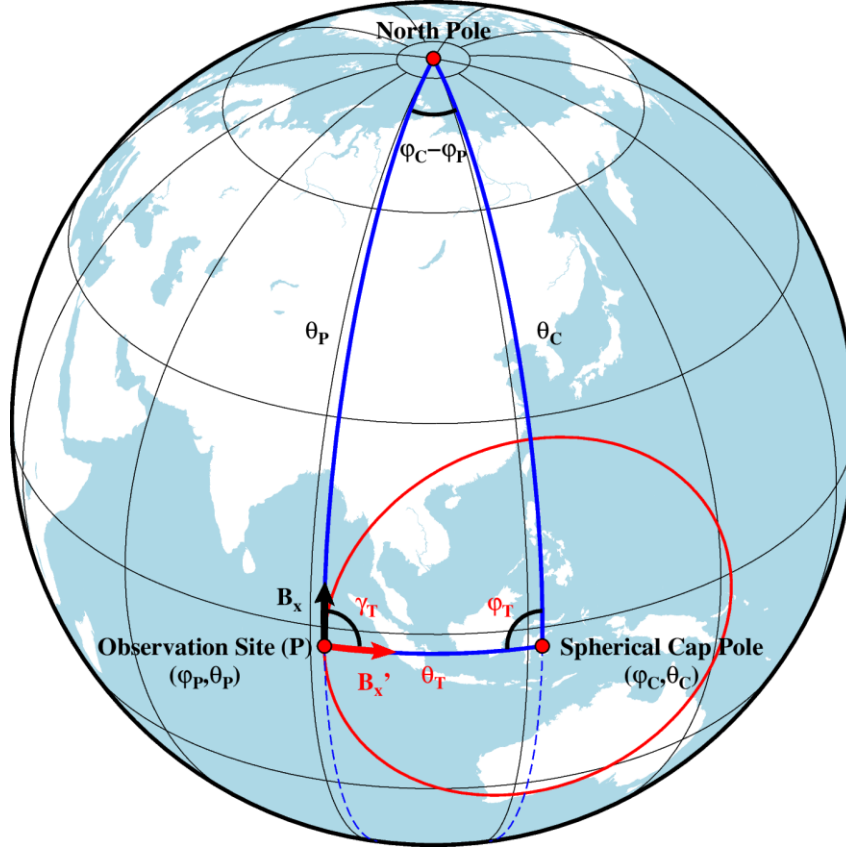


Figure 8. Transformation of the geographic coordinate system to the spherical cap coordinate system. (φ_C, θ_C) are the spherical cap pole coordinates. (φ_P, θ_P) are the coordinates of the observation site in the geographic coordinate system, while (φ_T, θ_T) are the coordinates of the observation site in the spherical cap coordinate system. γ_T is the rotation angle needed to transform the geomagnetic component B_X in geographic reference frame to B_X' in spherical cap reference frame.

Consider an arbitrary point P in the geographic coordinate system which has coordinate (φ_P, θ_P) and the spherical cap pole with coordinate (φ_C, θ_C) . The meridian of these points (θ_C) and (θ_P) pass through the centre of the cap and data sites, and simply become the sides of the spherical triangle with longitude difference between these two sides becoming the angle $(\varphi_C - \varphi_P)$. Using the spherical version of the cosine rule, the transformed colatitude (θ_T) of the data point, i.e., relative to the centre of the cap, can be determined as follows (Ingham, 2009)

$$\theta_T = \cos^{-1}(\cos \theta_C \cos \theta_P + \sin \theta_C \sin \theta_P \cos(\varphi_C - \varphi_P)) \quad (27)$$

while the transformed longitude (φ_T) can be determined using the spherical version of the cosine rule as follows

$$\varphi_T = \cos^{-1} \left(\frac{\cos \theta_P - \cos \theta_T \cos \theta_C}{\sin \theta_T \sin \theta_C} \right) \quad (28)$$

Similarly, the geomagnetic data at point P also need to be transformed into the cap coordinate system. The rotation angle γ needs to be determined for every data point using the spherical version of the cosine rule

$$\gamma = \cos^{-1} \left(\frac{\cos \theta_C - \cos \theta_T \cos \theta_P}{\sin \theta_T \sin \theta_P} \right) \quad (29)$$

Finally, the transformed geomagnetic data for the X and Y components can be determined as

$$B'_X = B_X \cos \gamma + B_Y \sin \gamma \quad (30)$$

$$B'_Y = -B_X \sin \gamma + B_Y \cos \gamma \quad (31)$$

The downward component (B_Z) is unaffected by the coordinate transformation.

3.2.3 Removal of a Global Reference Field

Haines (1988), in his paper, explained that subtracting a global reference field from the observational data before applying the SCHA is advantageous, both statistically and practically. From the statistical point of view, this procedure can minimise the leakage of power from long wavelengths into the band being analysed. In other words, removing the global reference field is similar to pre-whitening a spectral density function. From a practical point of view, the advantage is that the rate of convergence is increased. Moreover, the final model, with the global reference field added back, will extrapolate more accurately beyond the spherical cap boundary.

However, in this research, the global reference field (IGRF-13) is used to fill the empty area inside the cap border and is considered as synthetic data points. Thus, removing the global reference field will result in zero values for the synthetic data point locations in every epoch. To solve this problem, as has been done by Qamili et al. (2010), the IGRF values from the centre of the time series are considered as the global reference field. The data time series in this research spans from 1985.5 – 2015.5, and as a result, the IGRF 2000.5 calculated on the Earth's surface is used to remove the global reference field in all epochs.

3.2.4 Calculation of the Gauss Coefficients

The expression of B_X , B_Y , and B_Z at each measurement site can be obtained using equation (23). From this equation, the magnetic field components are determined in terms of the colatitude (θ_T), longitude (φ_T) of the sites, the Gauss coefficients (g_k^m and h_k^m), and the ratio of Earth's radius (a) and distance from the measurement site to the centre of the Earth (r). In this research, all measurements are done at the surface of the Earth; thus, $r = a$ and a/r dependence can be omitted. Thus, the calculation of Gauss coefficients now only relies upon the coordinate of the site location and the measured magnetic field values.

By applying simple matrix inversion, the Gauss coefficients is determined as follows

$$\begin{bmatrix} X_{g_0^0}(\theta_1, \varphi_1) & X_{g_1^0}(\theta_1, \varphi_1) & X_{g_1^1}(\theta_1, \varphi_1) & X_{h_1^1}(\theta_1, \varphi_1) & \cdots \\ Y_{g_0^0}(\theta_1, \varphi_1) & Y_{g_1^0}(\theta_1, \varphi_1) & Y_{g_1^1}(\theta_1, \varphi_1) & Y_{h_1^1}(\theta_1, \varphi_1) & \cdots \\ Z_{g_0^0}(\theta_1, \varphi_1) & Z_{g_1^0}(\theta_1, \varphi_1) & Z_{g_1^1}(\theta_1, \varphi_1) & Z_{h_1^1}(\theta_1, \varphi_1) & \cdots \\ X_{g_0^0}(\theta_2, \varphi_2) & X_{g_1^0}(\theta_2, \varphi_2) & X_{g_1^1}(\theta_2, \varphi_2) & X_{h_1^1}(\theta_2, \varphi_2) & \cdots \\ Y_{g_0^0}(\theta_2, \varphi_2) & Y_{g_1^0}(\theta_2, \varphi_2) & Y_{g_1^1}(\theta_2, \varphi_2) & Y_{h_1^1}(\theta_2, \varphi_2) & \cdots \\ Z_{g_0^0}(\theta_2, \varphi_2) & Z_{g_1^0}(\theta_2, \varphi_2) & Z_{g_1^1}(\theta_2, \varphi_2) & Z_{h_1^1}(\theta_2, \varphi_2) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} g_0^0 \\ g_1^0 \\ g_1^1 \\ h_1^1 \\ \vdots \end{bmatrix} = \begin{bmatrix} X(\theta_1, \varphi_1) \\ Y(\theta_1, \varphi_1) \\ Z(\theta_1, \varphi_1) \\ X(\theta_2, \varphi_2) \\ Y(\theta_2, \varphi_2) \\ Z(\theta_2, \varphi_2) \\ \vdots \end{bmatrix} \quad (32)$$

Where $X_{g_k^m}(\theta_T, \varphi_T)$ and $X_{h_k^m}(\theta_T, \varphi_T)$ are the spatial terms of degree k , order m , at a measurement site (θ_T, φ_T) of Gauss coefficient g and h on X component. $X(\theta_T, \varphi_T)$, $Y(\theta_T, \varphi_T)$, and $Z(\theta_T, \varphi_T)$ are the measured B_X , B_Y , and B_Z at site (θ_T, φ_T) respectively. The modified *Matlab* programme from Ingham (2009) is used to transform the measurement data into the spherical cap coordinate system. It calculates the Gauss coefficients using least square best fit and transforms back the coordinates and the field into a geographic coordinate system. After the Gauss coefficients are obtained, the model of the magnetic fields can be calculated at any locations inside the cap border by bringing the Gauss coefficients back into equation (32).

When the solution over a portion of the spherical cap is poorly constrained, the coefficient rejection procedure described in Haines and Fiori (2013) can be used to minimise the anomalies. This procedure will eliminate non-significant fitting coefficients. However, Korte and Holme (2003) criticise the coefficient rejection practice as it does not have physical justification in SCHA. Nevertheless, the simultaneous solution of all the coefficients from all the data in a single analysis will result in a more objective estimation.

3.3 Spatial Model Selection

Appropriate selection of the non-integer degree is necessary to ensure optimal performance of the SCHA algorithm. In order to choose the best spherical cap harmonics degree, several models with spatial truncation index $k = 1 - 10$ have been calculated. The selection of the appropriate value of k and therefore $(n_k(m))$ is necessary to ensure an accurate representation of the mapped field based on the available data. If the k index is too low, then the model will be overly smooth. Conversely, if the k index is too high, the model will overfit the data (Fiori, 2020). This chapter is concerned with the statistics and validation of the SCHA model and thus precedes by explaining the procedures for selection of the best model truncation and maximum resolution calculation.

3.3.1 Model Statistics and Validation

In order to investigate the reliability and validate an SCHA model, the goodness of fit of the model needs to be analysed. In this study, the results of the SCHA calculations using spatial truncation index $k = 1 - 10$ are analysed statistically in order to select the best model. The fit of the model is not only determined by its contour plot but also can be quantified by its Root Mean Square Difference (RMSD) values. RMSD values are used to express the magnitude of the deviation between the model field and the observation field values. It can be calculated as follows

$$RMSD = \sqrt{\frac{\sum_{p=1}^P (\hat{y}_p - y_p)^2}{P}} \quad (33)$$

where \hat{y}_p are the predicted values of the SCHA model field for sites p , y_p is the observation values, and P is the number of sites.

The RMSD is calculated for the SCHA models with different spatial truncation index k . As a baseline, the RMSD of the IGRF model is used to compare the results of the RMSD of the SCHA models. This study aims to obtain a better model and better resolution than the currently available model, i.e., the IGRF. Thus, the SCHA model is considered better than the IGRF if its RMSD is lower than the RMSD of the IGRF. The RMSD of the IGRF is calculated on similar data locations as in SCHA calculations. However, in order to validate the modelling technique, two repeat stations data are excluded from the SCHA modelling, i.e., Bandar Lampung (105.175°E, 5.240°S) and Tual (132.735°E, 5.662°S). Therefore, these two validation

stations are excluded in the RMSD calculations, both for the SCHA and IGRF models.

After the RMSD calculations, the next step is checking the contour map of the SCHA models and comparing them with the contour map of the IGRF. As the RMSD calculations are only performed on the observational data locations, this procedure is needed in order to check the stability of the modelling technique over the whole spherical cap. The SCHA model is considered reliable if the contour map is smooth and does not exhibit any new structure over the cap.

The next procedure in the statistical analysis is performing the model validation. Validation of the model usually uses data that has not been used in model construction. In other words, validation of the model usually incorporates examining how far the constructed model differs from the real data. In this chapter, the validation is based on the observation data at two repeat stations, Bandar Lampung and Tual. The locations of these validation stations can be seen in Figure 5. The validations are calculated by subtracting the model value from the observational data. The smaller the differences, the more the model can be considered to be a better model.

The last step is calculating the difference between the SCHA model with different index k and the IGRF model. The difference is calculated spatially by creating a grid over the cap and creating a contour map from the difference values. The SCHA model is considered the best if the difference between the selected SCHA model and the IGRF model is not overly high or low. After the statistical analysis has been done, the best model needs to be selected, and from the maximum spatial truncation index, the maximum resolution of the SCHA model can be determined.

3.3.2 Model Truncation and Maximum Resolution

The best SCHA model can be determined following the procedures in the statistical analysis and model validation. If there are more than one of the SCHA models that fulfils the statistical requirements, these models are selected by using the spatial truncation index k , i.e., the highest index k or k_{max} . This maximum spatial truncation index k_{max} can be used to determine the maximum resolution of the respective model, as has been explained by Haines (1988)

$$k_{max} \approx \frac{\theta_0}{90^\circ} \left(\frac{360^\circ}{\omega_{min}} + 0.5 \right) - 0.5 \quad (34)$$

where ω_{min} is the minimum wavelength to be represented in the model. As a comparison, the IGRF model uses spherical harmonic degrees up to and including 13. Therefore, as has been explained in Section 2.2.1, the shortest wavelength representation of the IGRF can be obtained by dividing 360° by its degree. As 360° represents the full circumference of the Earth (40075 km), thus the minimum wavelength which can be described is approximately 3000 km. By utilising equation (34), the similar minimum wavelength of the SCHA model with $\theta_0 = 30^\circ$ is found to be given by $k_{max} = 4$. As a result, the SCHA model is considered better than the IGRF if its maximum spatial truncation index k_{max} is greater than or equal to four. The last step in the spatial modelling analysis, after the best SCHA model is determined, is to calculate the misfit of this model by differencing the value between the SCHA model and the observational data at all data point locations. Following the spatial modelling analysis, temporal modelling and secular variation forecasting are performed.

3.4 Temporal Modelling and Secular Variation Forecasting

A common practical problem when analysing geomagnetic data is how to fit a smooth function to an observational time series with sparse time intervals. In the IGRF model, the time dependence of the Gauss coefficients is assumed to be linear over the 5-year interval (Thébault et al., 2015). However, as has been discussed in Section 2.4.1, the secular variation of the geomagnetic field is somewhat unpredictable, and its variation is sometimes shorter than five years. Therefore, new time-dependent models of the geomagnetic field, which have non-linear variation, have to be developed. One of the common techniques is by utilising the B-splines function.

In this research, the SCHA model is calculated for every epoch in a 5-year interval from 1985.5 to 2015.5, and the annual interpolation is done using B-splines basis functions. The technique which takes a series of spatial models at particular epochs and then uses some form of interpolation between them is called two-step modelling (Jackson & Finlay, 2015).

Another problem in the modelling geomagnetic reference field is how to predict the field such that it is still valid several years after the model has been published. In regional geomagnetic modelling, this is mostly concerned with extrapolation of the data without taking into consideration the physical constraints. In this research, the prediction of the geomagnetic field, valid from 2015.5 to 2020.5, is determined using

the autoregressive (AR) technique. In this chapter, the interpolation technique using the B-splines basis function is described first, and then a discussion of the AR forecasting technique follows.

3.4.1 Time-dependent Model Based on Cubic B-splines

The SCHA calculations in this research are done in a 5-year interval. The resulting Gauss coefficients (g_k^m and h_k^m) are then fitted with a B-splines representation of order 4 (cubic B-splines) as the basis functions, such as used by Jackson et al. (2000) when developing *gufm1*. The procedure in this chapter is done using the *Matlab* programme.

A spline function of order k is a piecewise polynomial function of degree $k - 1$ in a variable x , and the location where the pieces of polynomial function meet are called knots. The knots can be considered as division points that subdivide the interval $[t_0, t_m]$ into knot spans. All of the B-splines basis functions are considered to have their domain on $[t_0, t_m]$. As in Verbanac (2007), for the given sequence of knots, the i -th B-splines basis function of order k is written as $B_{i,n}(t)$ and can be defined by construction of the Cox-de Boor recursion formula.

$$B_{i,1}(x) = \begin{cases} 1, & \text{if } t_i \leq x < t_{i+1} \\ 0, & \text{otherwise} \end{cases} \quad (35)$$

$$B_{i,k}(x) = \frac{x - t_i}{t_{i+k-1} - t_i} B_{i,k-1}(x) + \frac{t_{i+k} - x}{t_{i+k} - t_{i+1}} B_{i+1,k-1}(x)$$

B-splines with order 2 are called piecewise linear, order 3 are piecewise quadratic, and order 4, as has been used in this research, are piecewise cubic. In this research, to make it simple, the knots t_i are chosen to be the data points. Then, by equation (35), the cubic B-splines with equidistant knots are given as

$$B_{i,4}(x) = \begin{cases} 1/6z^3, & t_{i-2} < x < t_{i-1} \\ 1/6[1 + 3(1 + z(1 - z))z], & t_{i-1} \leq x < t_i \\ 1/6[1 + 3(1 + z(1 - z))(1 - z)], & t_i \leq x < t_{i+1} \\ 1/6(1 - z)^3, & t_{i+1} \leq x < t_{i+2} \\ 0, & \text{otherwise} \end{cases} \quad (36)$$

where the distance between the knots is given as $d = t_{i+1} - t_i$ and z is given as $z = (x - t_{i+2})/d$. As in Langel et al. (1986), the general cubic B-splines over (t_i, t_m) are given by

$$s(t) = \sum_{i=1}^m \gamma_i B_{i,4}(t) \quad (37)$$

where t_i with $i = 1, \dots, m$ are a set of strictly increasing real numbers; $s(t)$ is a polynomial of order 4 or less in each of the intervals, and γ_i are coefficients determined by the least-squares fit of $s(t)$ to the data.

In this research, each of the Gauss coefficients from the SCHA calculations is expanded in time as a combination of cubic B-splines $B_{i,4}(t)$

$$g_k^m(t) = \sum_{i=1}^m \gamma_i B_{i,4}(t) \quad (38)$$

and similarly for h_k^m . Practically, the B-splines interpolation is done using the *spline* function in the *Matlab* programme using 5-year epoch intervals and the data on the corresponding epoch as input, and then determining the annual interval as the interpolated query points. After interpolating the observation data using B-splines basis functions on an annual basis, the results are compared with the observation data and the value from the IGRF model.

Following this step, the annual rate of changes, i.e., the difference between successive years of the time series (Wardinski & Holme, 2011), can be determined as follows

$$\left. \frac{dX}{dt} \right|_t = X(T) - X(T - 1) \quad (39)$$

where T is given time, and X is the component of the geomagnetic field (Peqini et al., 2018). Similarly, the results of the annual rate of changes from the SCHA model are compared with the annual rate of changes in the observational data and the value from the IGRF model. Lastly, from the annually interpolated SCHA model, the forecasting of the secular variation can be performed using the autoregressive technique.

3.4.2 Autoregressive Forecasting

Autoregressive forecasting is a forecasting technique of the variable of interest using a linear combination of the variable from past values (Hyndman & Athanasopoulos, 2018). In other words, it is a regression of the variable against itself, as indicated by the term *autoregression*. Autoregressive forecasting is a type of extrapolation method. The autoregressive models of order p can be written as

$$y_t = c + k_1 y_{t-1} + k_2 y_{t-2} + \cdots + k_p y_{t-p} + \varepsilon_t \quad (40)$$

where c is a constant, ε_t is white noise, t is the period, and k_p are the coefficients. In this regression model, the response variable in the previous period becomes the predictor. The order p indicates the number of immediately preceding values in the series that are used to predict the value at present. The autoregressive model of order p can be written as $AR(p)$.

The order of the autoregressive model can be determined by looking at its partial autocorrelation function plot. The partial autocorrelation of an $AR(p)$ model is zero at lag $p + 1$ and greater. The partial autocorrelation at lag p is the autocorrelation between y_t and y_{t-p} which is not considered by lags 1 through $p - 1$ (NIST/SEMATECH, 2012). After the order of the $AR(p)$ is determined, the coefficients k_1, \dots, k_p can be calculated using ordinary least squares method. However, the autoregressive forecasting method can only be used if the time series is stationary, i.e., the mean, variance, and autocorrelation structure do not change over time. Thus, to make the time series stationary, the first difference of the time series needs to be calculated first.

In this research, the time series for the secular variation prediction of the model is the first difference of the annual mean values or the annual rate of change on each data location. Then the partial autocorrelation is plotted using Statsmodel (Seabold & Perktold, 2010), a library for statistical and econometric analysis in Python. The order p can be determined by looking at the partial autocorrelation plot where the maximum lag is greater than the 95% confidence interval for statistical significance. Following this step, the coefficients of the $AR(p)$ can be calculated using a similar programme, and the forecasting can be calculated but limited to the next five years to ensure accuracy. Then, the time series is transformed back into its original values as before the first differencing. Finally, the results of the $AR(p)$ forecasting method are compared with the observational values and the forecasting value from the IGRF model.

Chapter 4

Results and Discussions

4.1 Spatial Modelling Analysis

This chapter presents the analysis of the effect of the value of the truncation spatial index k on the conformity of the SCHA model and the real observation data, and the determination of the best model based on its statistical properties. The RMSD from all the calculated models is presented first and is followed by an analysis of the model stability, which is the analysis of the reliability of the modelling method. This analysis of stability is performed for several selected models only. The validation of the model by comparing its calculated values and the real observation data at two selected locations will then be discussed. Following this, the analysis of the difference between the regional and global model will be presented. Lastly, the selection of the best truncation spatial index k and the misfit between the real observed data and the predicted data from the selected model will be discussed.

4.1.1 RMSD of the Model

For the use of a magnetic reference for navigation, rather than an estimation of the core field alone, an accurate representation of the geomagnetic field at the point of observation is required (De Santis et al., 1990). The accuracy of the geomagnetic model can be determined by calculating its statistical properties, e.g., RMSD. The RMSD in this chapter is calculated using all the observational data, but the number of data locations varies between epochs. Nevertheless, in the same epoch, the RMSD of the SCHA and the IGRF model are calculated on the same data point locations. In order to ensure that the data distribution location and the modelling method give stable and reliable results, the number of data points and the RMSD of the model should not indicate any relationship, e.g., high RMSD does not correlate with a high number of data points or conversely.

Figure 9 shows the RMSD for each epoch as a function of k . From Figure 9, it is unquestionable that the RMSD of the SCHA model is significantly improved as the spatial truncation index k increases. It means the SCHA model is approaching the real observation data as index k increases. Generally, the RMSD of the SCHA model declines significantly from index $k = 1 - 3$, then declines gradually from index $k > 3$. This is consistent with Düzgüt and Malin (2000), who concluded that the SCHA model is improved rapidly when the number of coefficients is increased.

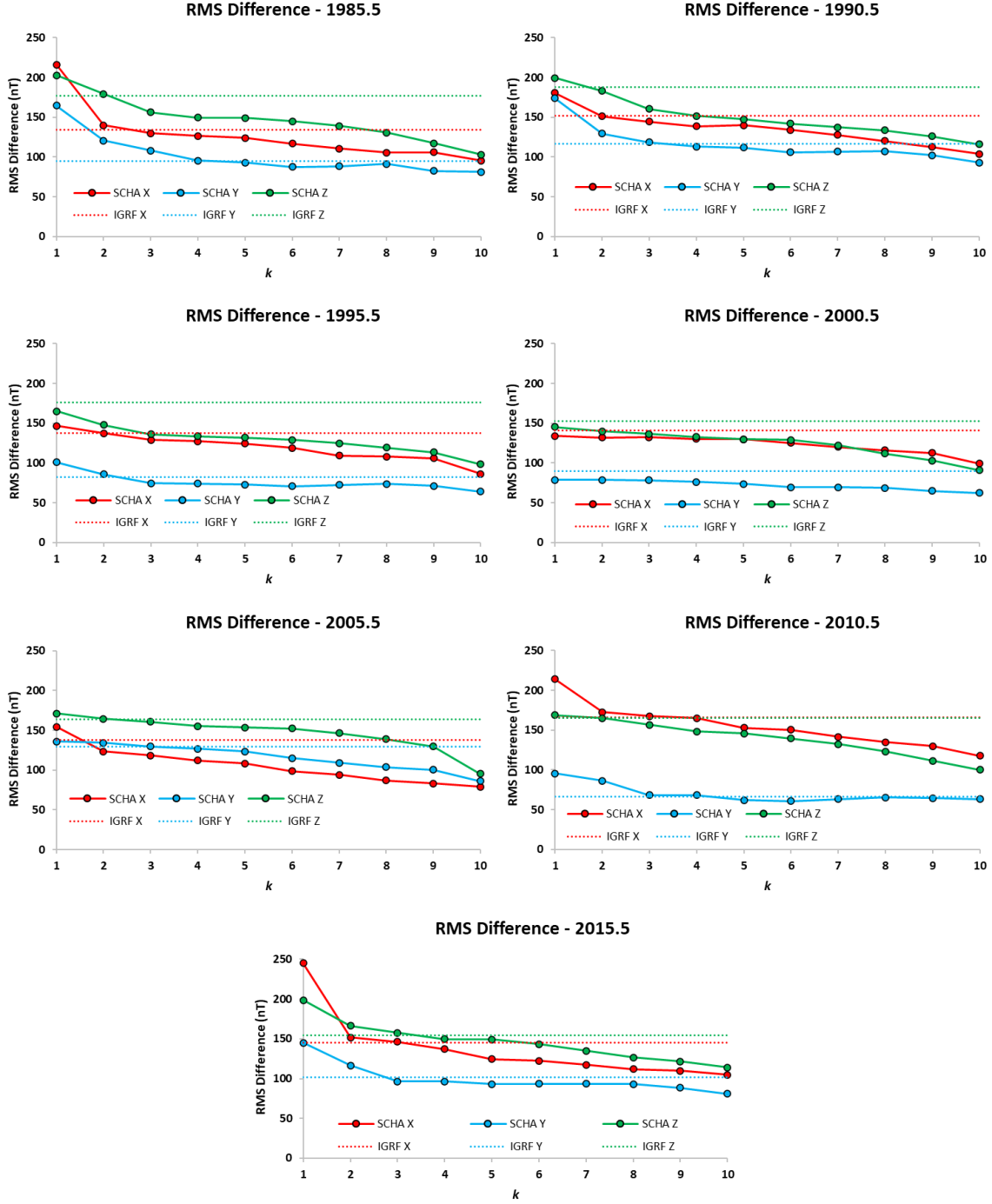


Figure 9. RMSD of the IGRF and SCHA model with different index k from 1985.5 – 2015.5. The solid lines with dots are the RMSD of the SCHA models, while the dotted lines are the RMSD of the IGRF model.

The value of index k , which starts to provide a better RMSD than the RMSD of the IGRF model varies for each component for all epochs. Nevertheless, overall, at index k equal to or greater than 4, the RMSD of the SCHA model is better than the RMSD of the IGRF model. Indeed, at index $k = 4$, the SCHA model will represent the same minimum wavelength with the IGRF, that is around 3000 km at the spherical cap

angle $\theta_0 = 30^\circ$. Therefore, at the same and better model resolution, the RMSD of the SCHA model is better than the RMSD of the IGRF model.

The higher index k and the shorter the separation of the stations, the shorter the wavelength that can be represented. An attempt to include shorter wavelength fields by extending a global spherical harmonic model to a higher degree requires a large number of coefficients and a dense distribution of the global data set. A regional modelling technique, such as SCHA, can take advantage of the local data set and utilise far fewer model coefficients by restricting the analysis to a limited region (De Santis et al., 1990).

Haines (1985), in his original paper, demonstrated that the accuracy of the least-squares solution in the SCHA method is better than the conventional spherical harmonic method, which has been used to develop the IGRF model. However, this advantage is obtained by introducing error at the boundary of the spherical cap. He explained further that the error at the boundary is the result of omitting the data outside of the spherical cap.

Nevertheless, in Figure 9, the RMSD is calculated from the observational data locations only. Since the number of observatories near the boundary is limited, it does not represent the RMSD at the boundary of the cap sufficiently. Hence, although the RMSD of the SCHA provides a good result, the contour map of the SCHA model needs to be analysed thoroughly.

4.1.2 Stability of the Model

The data distribution at each epoch varies, with the highest number of the observational data locations in 2015 and the lowest number in 2005. In order to know the reliability of the SCHA modelling method, the results must be analysed by checking the contour maps for the different epochs. As a comparison and a baseline, the contour map of the IGRF model is plotted correspondingly.

According to Walker (1989), there is a trade-off in the modelling between reliability and the maximum index. It is preferable to calculate a model using fewer coefficients rather than use a large number of coefficients to exactly fit the observations data but with zero degrees of freedom in the error estimate. He also suggests that the wavelength of the model should be greater than the separation of the stations.

The SCHA model with index $k = 4$ represents the same minimum wavelength as the IGRF, that is around 3000 km at the spherical cap angle $\theta_0 = 30^\circ$. As a result, to

obtain a better model than the global model, the index k should be greater than or equal to 4. Inspection of the contour maps shows that maps with index $k < 7$ show no significant differences. On the other hand, the contour maps at index $k = 9$ start to show a new structure that does not exist on the contour maps with a lower index. The structure becomes more erratic as the index k is increased further, especially on the boundary of the research area. Thus, by considering the minimum wavelength to represent, the SCHA model discussed in this chapter is considered to be that given by a truncation index $k = 7 - 9$.

4.1.2.1 IGRF Model

As a baseline, the IGRF model is used to compare the results of the SCHA calculation. One of the aims of this study is to obtain a better model and better resolution than the currently available model, e.g., IGRF. The IGRF model used here is the latest 13th generation accessed from the British Geological Survey (BGS). The RMSD is calculated between the IGRF model and the real observational data and is shown in Table 4. The contour map of the IGRF model is shown in Figure 10 through to Figure 13. Due to the high number of data points needed to plot the contour, the IGRF model is calculated using Geomag version 7.0 software from the National Centers for Environmental Information (NCEI) National Oceanic and Atmospheric Administration (NOAA) which is already integrated with the IGRF-13 coefficients. The IGRF grid data is plotted using Generic Mapping Tools version 6 (Wessel et al., 2019).

Table 4. RMSD from the IGRF model in 1985.5 - 2015.5

Epoch	X (nT)	Y (nT)	Z (nT)
2015.5	145.12	101.84	154.04
2010.5	166.46	66.48	165.53
2005.5	137.83	129.46	164.15
2000.5	140.61	90.15	152.53
1995.5	137.52	82.15	176.38
1990.5	151.41	116.32	187.51
1985.5	134.49	94.87	177.02

It appears that the RMSD of the IGRF model varies for each epoch. The smallest RMSD on the X component is in 1985.5, while the greatest is in 2010.5. On the Y component, the smallest RMSD is in 2010.5, and the greatest is in 2005.5. Lastly, on the Z component, the smallest RMSD is in 2000.5, and the greatest is in 1990.5. Moreover, the RMSD of the Y component is always smaller than the RMSD on the

X and Z components. For each component, the greatest and the smallest RMSD of the IGRF model never occurs in the same epoch. From 1985.5 – 2015.5, the average of the RMSD of the X, Y, and Z components is around 145 nT, 97 nT, and 168 nT, respectively.

The contour maps of the IGRF model (Figures 10-13) show that they are smooth and stable. There is very little difference for each epoch from 1985.5 – 2015.5. Nevertheless, westward drift can be seen in the contour map for the Y component.

According to Thébault et al. (2015), there are only three geomagnetic observatories from Indonesia that are used to model the IGRF-12, i.e., KPG (Kupang), PLR (Pelabuhan Ratu), and TND (Tondano). For the IGRF-11, according to Finlay et al. (2010), there are only two geomagnetic observatories from Indonesia which are used to calculate the model, i.e., TND (Tondano) and TUN (Tuntungan). For the IGRF-10, there are only TND (Tondano), TUN (Tuntungan), and TNG (Tangerang) (Macmillan & Maus, 2005). However, as yet, there is no publication available on the IGRF-13.

The maximum truncation degree of the IGRF model is 10 for the epoch up to and including 1995.0, providing a minimum wavelength of 4000 km. The same wavelength needs to be calculated by the SCHA method using spatial truncation index $k = 3$ at the spherical cap angle $\theta_0 = 30^\circ$. However, from 2000.0 onwards, the maximum truncation degree of the IGRF model is 13, and the minimum wavelength which can be represented becomes 3000 km. By using the SCHA method, the same wavelength can be obtained using $k = 4$ for the spherical cap angle $\theta_0 = 30^\circ$.

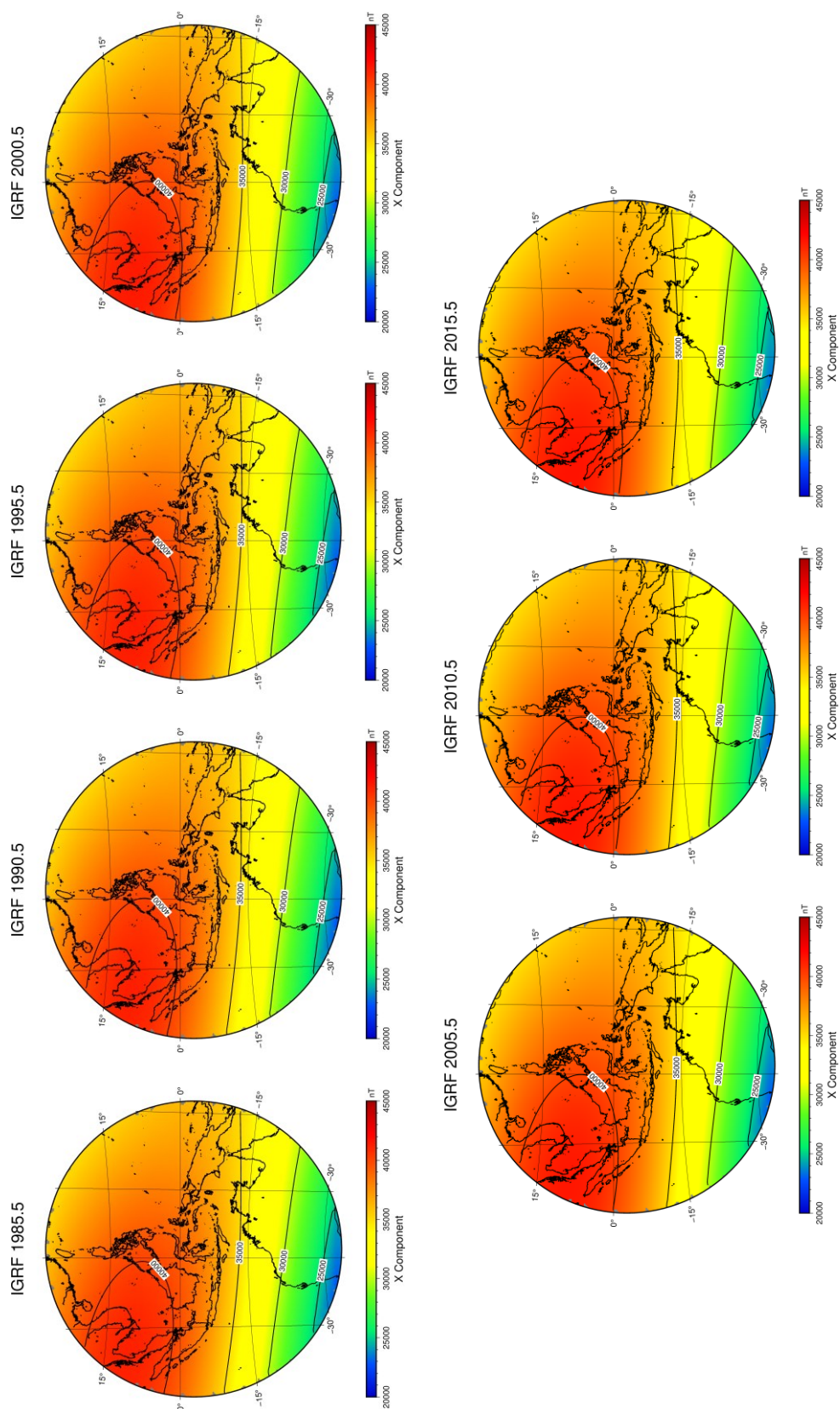


Figure 10. X component contour map from the IGRF model from 1985.5 - 2015.5.

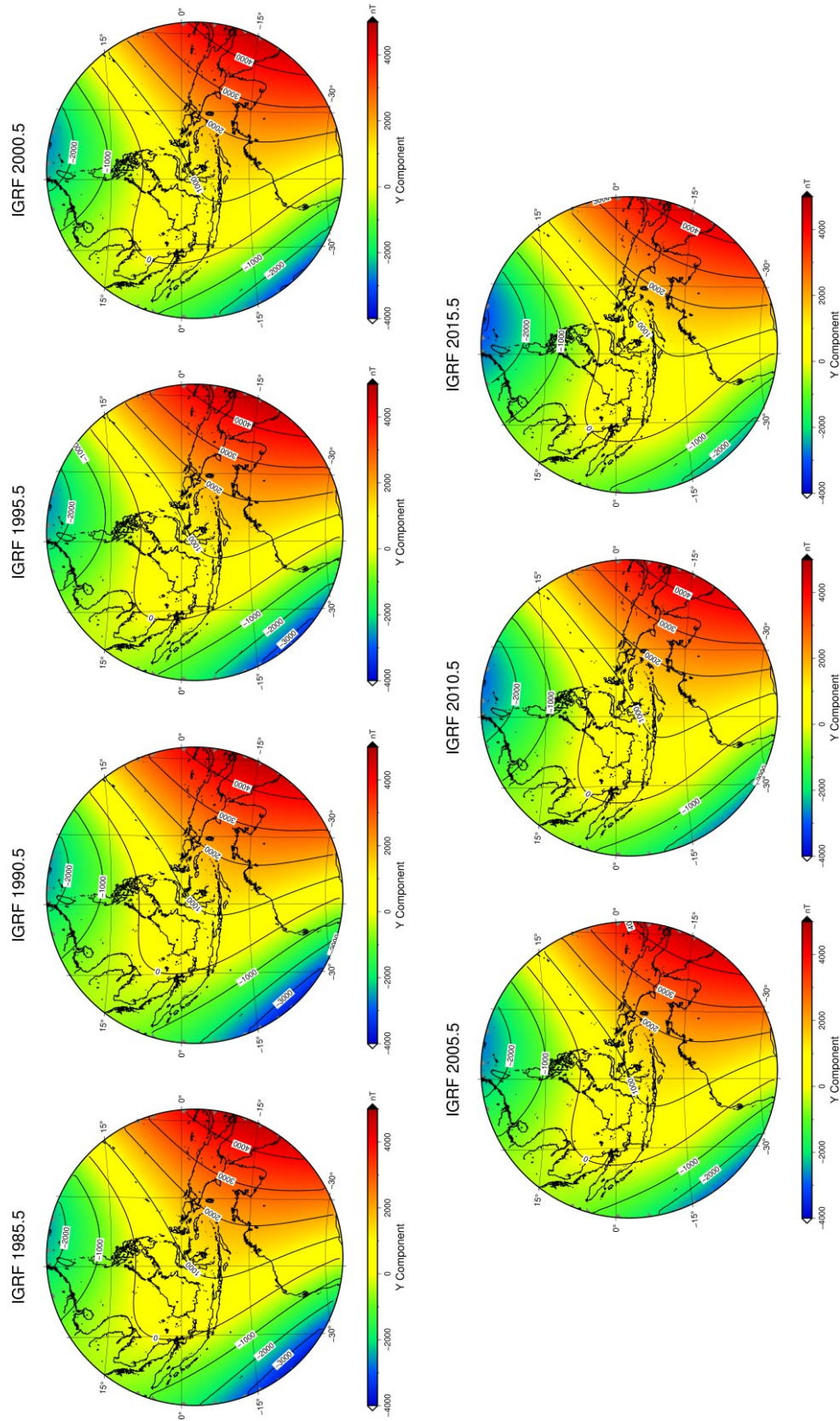


Figure 11. Y component contour map from the IGRF model from 1985.5 – 2015.5.

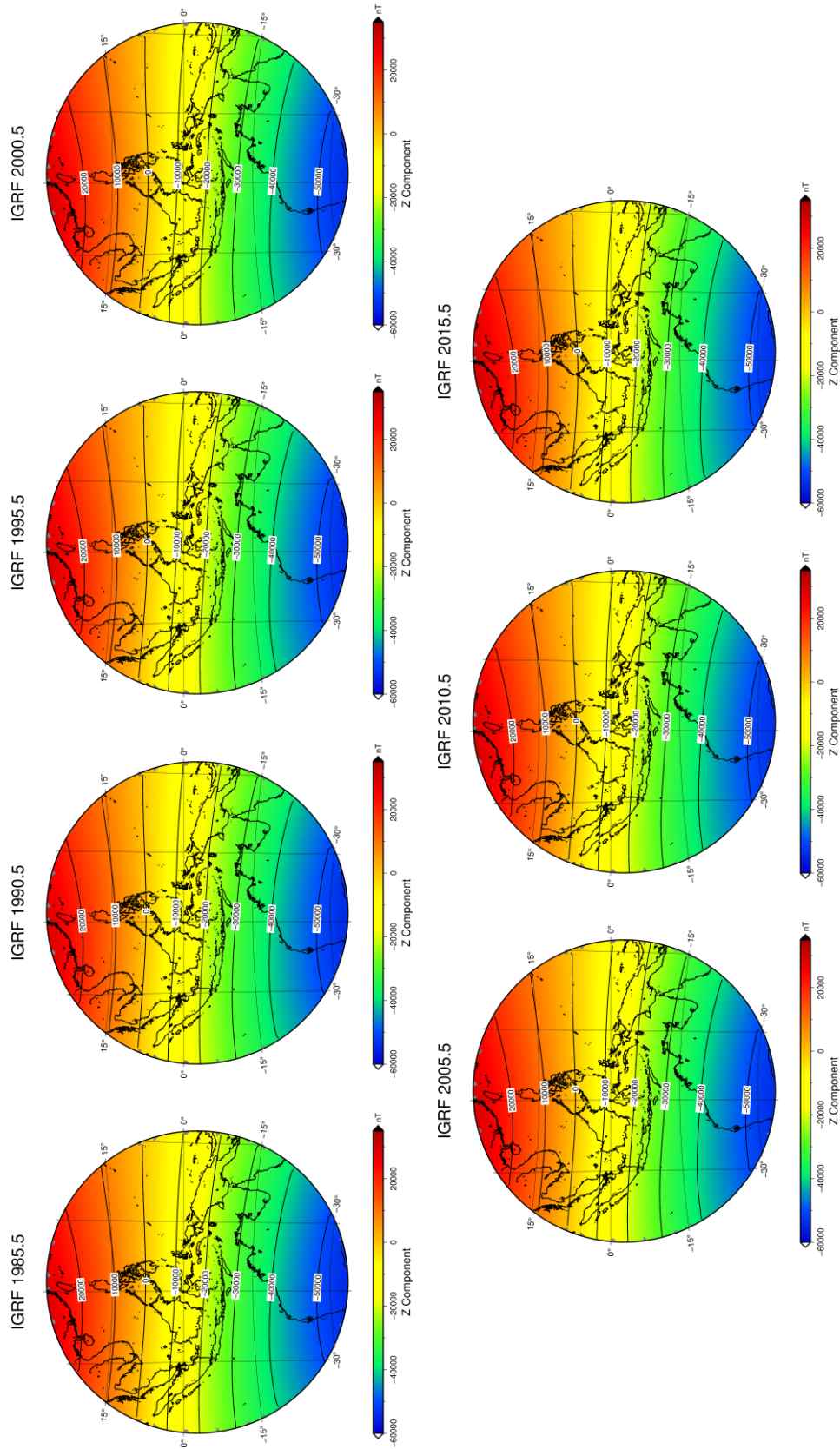


Figure 12. Z component contour map from the IGRF model from 1985.5 – 2015.5.

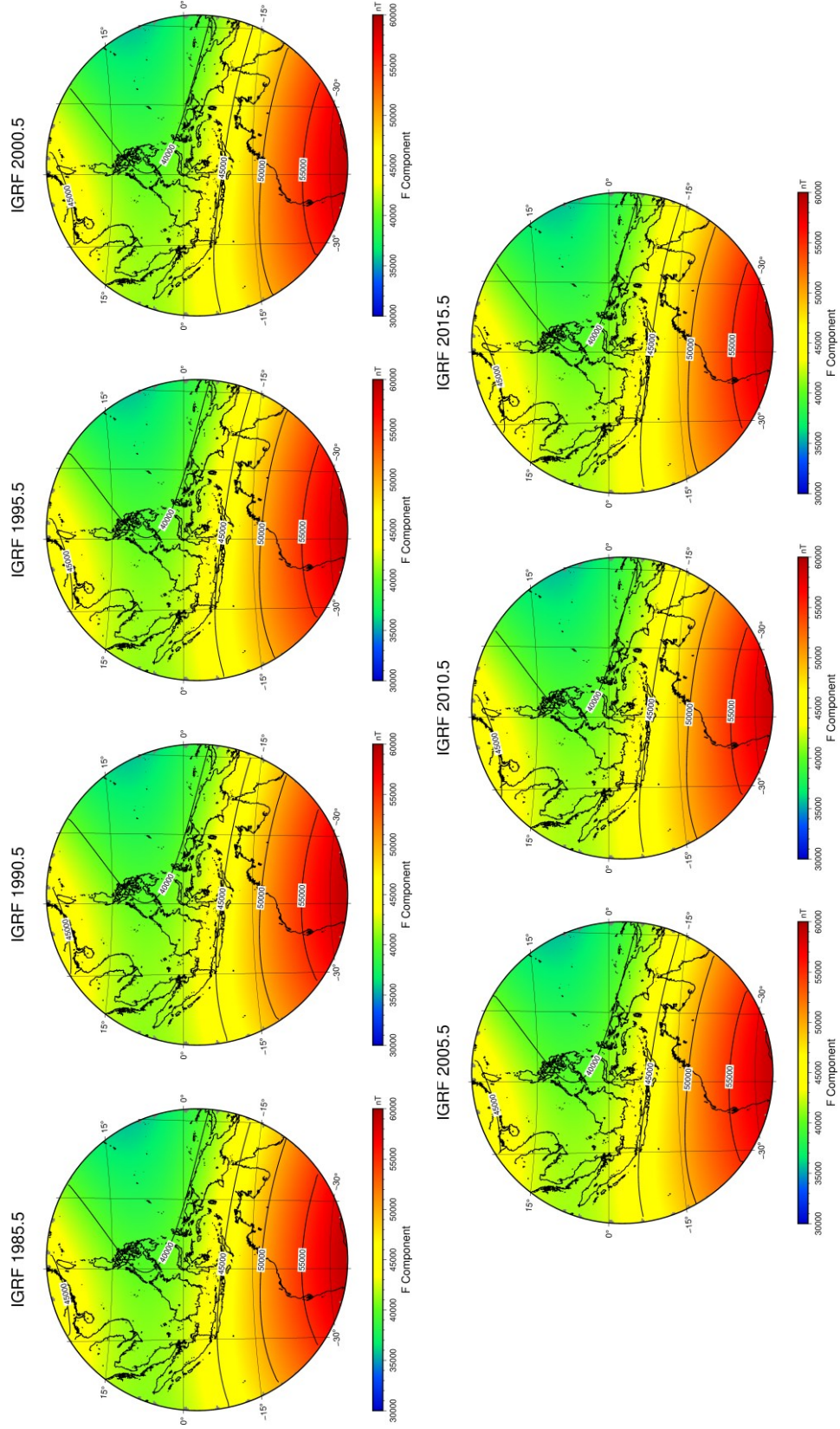


Figure 13. F component contour map from the IGRF model from 1985.5 – 2015.5.

4.1.2.2 SCHA with Spatial Truncation Index $k = 7$

The first SCHA model to be analysed is the SCHA model with spatial truncation index $k = 7$. The minimum wavelength which can be represented is approximately 1800 km, equal to that for an SHA model with a maximum truncation degree of 22. The RMSD from the SCHA model with spatial truncation index $k = 7$ is shown in Table 5, and the contour maps are shown in Figure 14 through to Figure 17.

Table 5. RMSD from the SCHA model with index $k = 7$

Epoch	X (nT)	Y (nT)	Z (nT)
2015.5	117.43	93.46	134.87
2010.5	141.55	63.02	132.19
2005.5	94.23	108.99	146.29
2000.5	120.10	69.67	122.12
1995.5	109.34	72.51	124.87
1990.5	127.66	106.95	137.56
1985.5	110.35	88.40	139.08

Compared with the RMSD from the IGRF model in Table 4, the RMSD of the SCHA model with index $k = 7$ shows a slight improvement for all components in all epochs. The RMSD of the SCHA model with index $k = 7$ is always smaller than the RMSD of the IGRF model. From 1985.5 – 2015.5, the average of the RMSD of the X, Y, and Z components is approximately 117 nT, 86 nT, and 134 nT, respectively. Thus, compared with the RMSD of the IGRF model, the RMSD of the SCHA model is reduced by 28 nT, 11 nT, and 34 nT for the X, Y, and Z components, respectively.

The position of the extrema, or maximum and the minimum RMSD values of the model is also different between the SCHA and the IGRF model. As an example, on the X component, the minimum RMSD of the SCHA model is in 2005.5, while on the IGRF model, the minimum RMSD is in 1985.5. This could probably be attributed to some crustal bias presenting itself in the Indonesian repeat stations and geomagnetic observatories, which are not used to model the IGRF. This is supported by the evidence that the extrema positions of the SCHA models lie in almost the same epoch for models with different index k .

The contour maps of the SCHA models are similar to the IGRF contour maps. The westward drift is still visible in the Y component. Westward drift is a motion of the geomagnetic field that is interpreted as an indication of core-mantle thermal coupling (Dumberry & Finlay, 2007). Therefore, the SCHA model appears to provide a reliable

result in terms of core field modelling. Note that the white triangles on the maps are the data observation locations and the blue stars are the location of the validation stations, which will be discussed later.

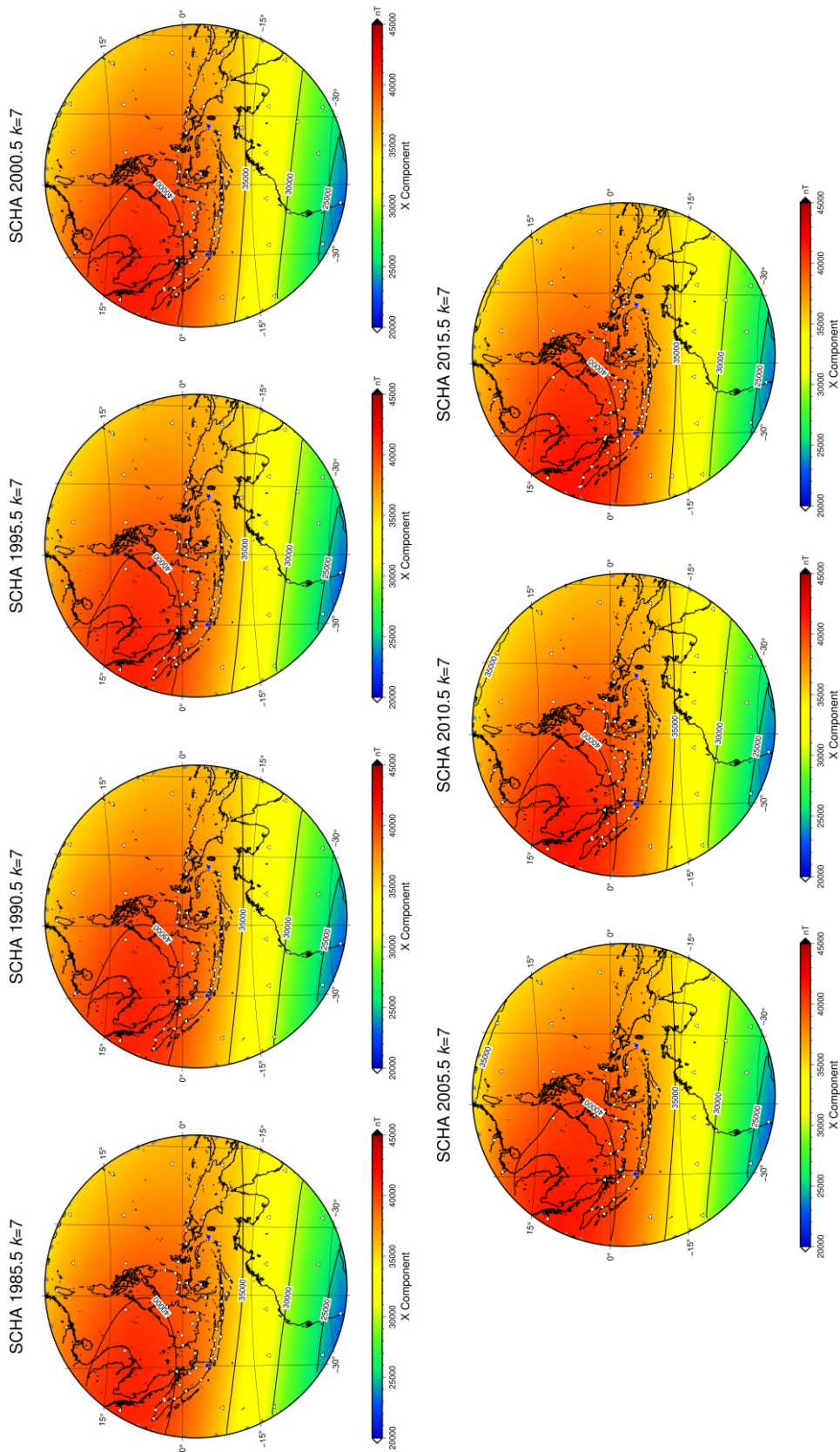
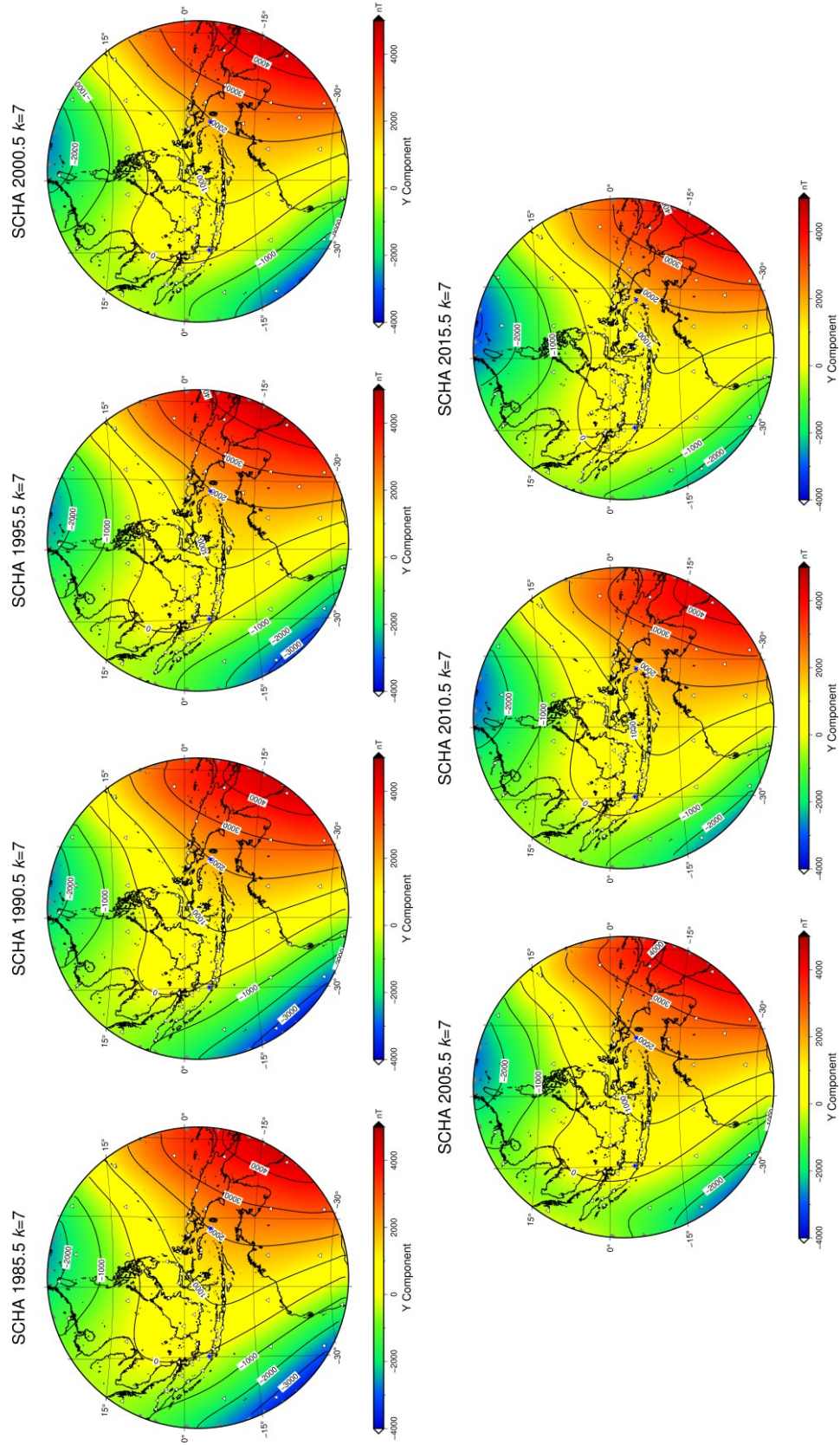


Figure 14. X component contour from the SCHa using index $k=7$ from 1985.5 – 2015.5.

Figure 15. Y component contour from the SCHa using index $k=7$ from 1985.5 – 2015.5.

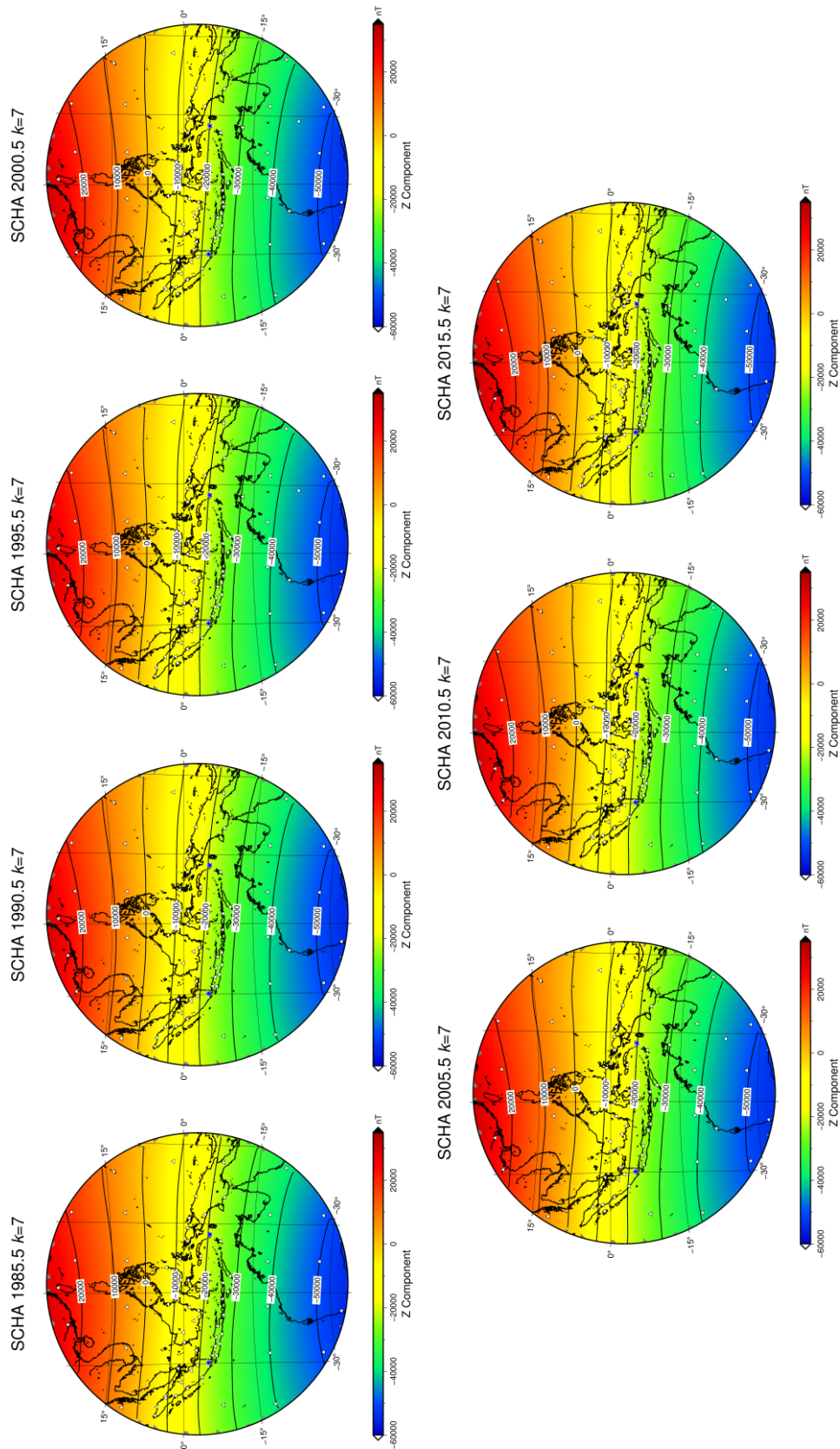
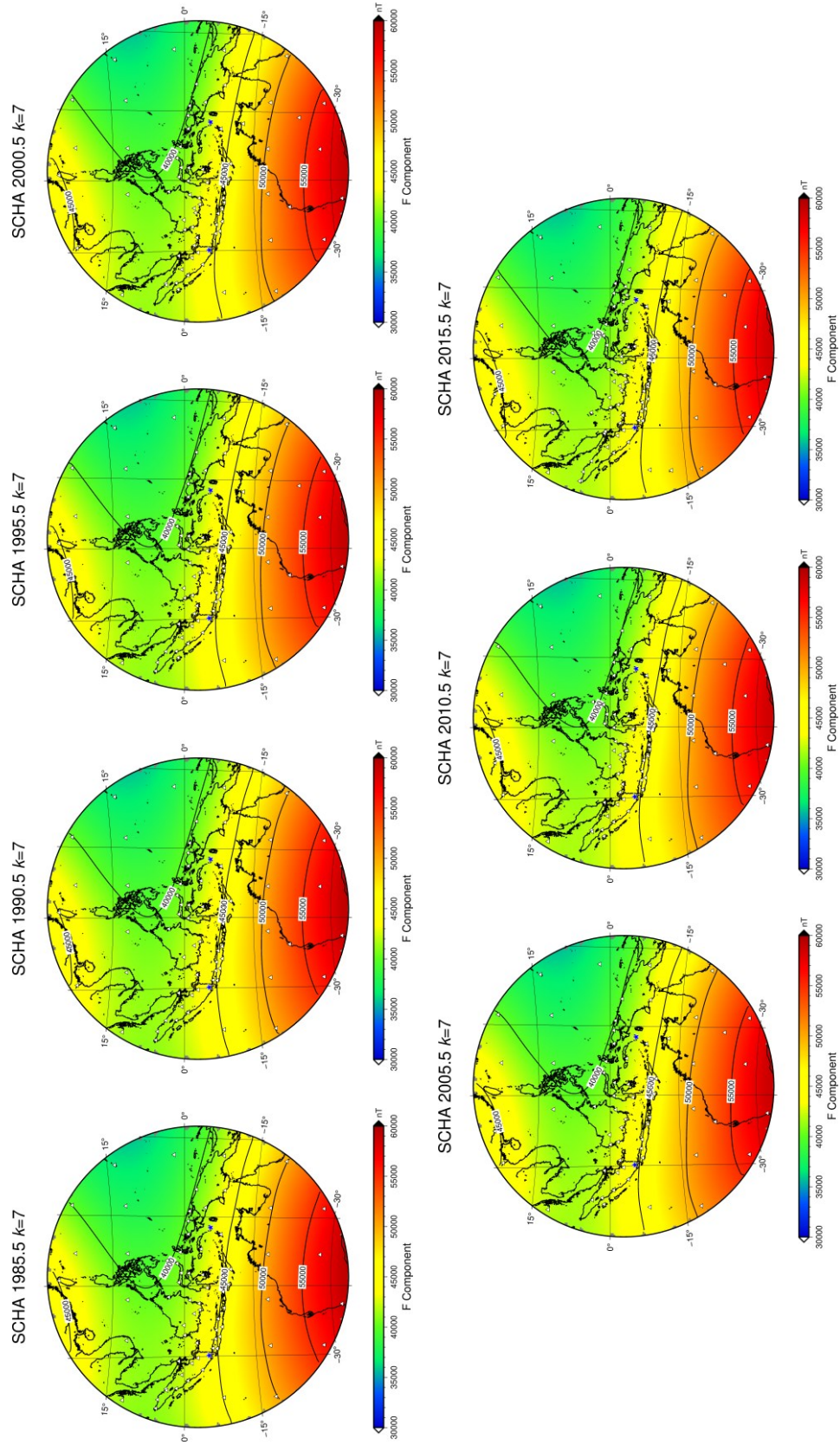


Figure 16. Z component contour from the SCHa using index $k=7$ from 1985.5 – 2015.5.

Figure 17. F component contour from the SCHAs using index $k=7$ from 1985.5 – 2015.5.

4.1.2.3 SCHA with Spatial Truncation Index $k = 8$

At index $k = 8$, the minimum wavelength which can be represented is approximately 1600 km, and it is equal to the SHA model with a maximum truncation degree of 25. The RMSD from the SCHA model with spatial truncation index $k = 8$ is shown in Table 6, and the contour maps are shown in Figure 18 through to Figure 21.

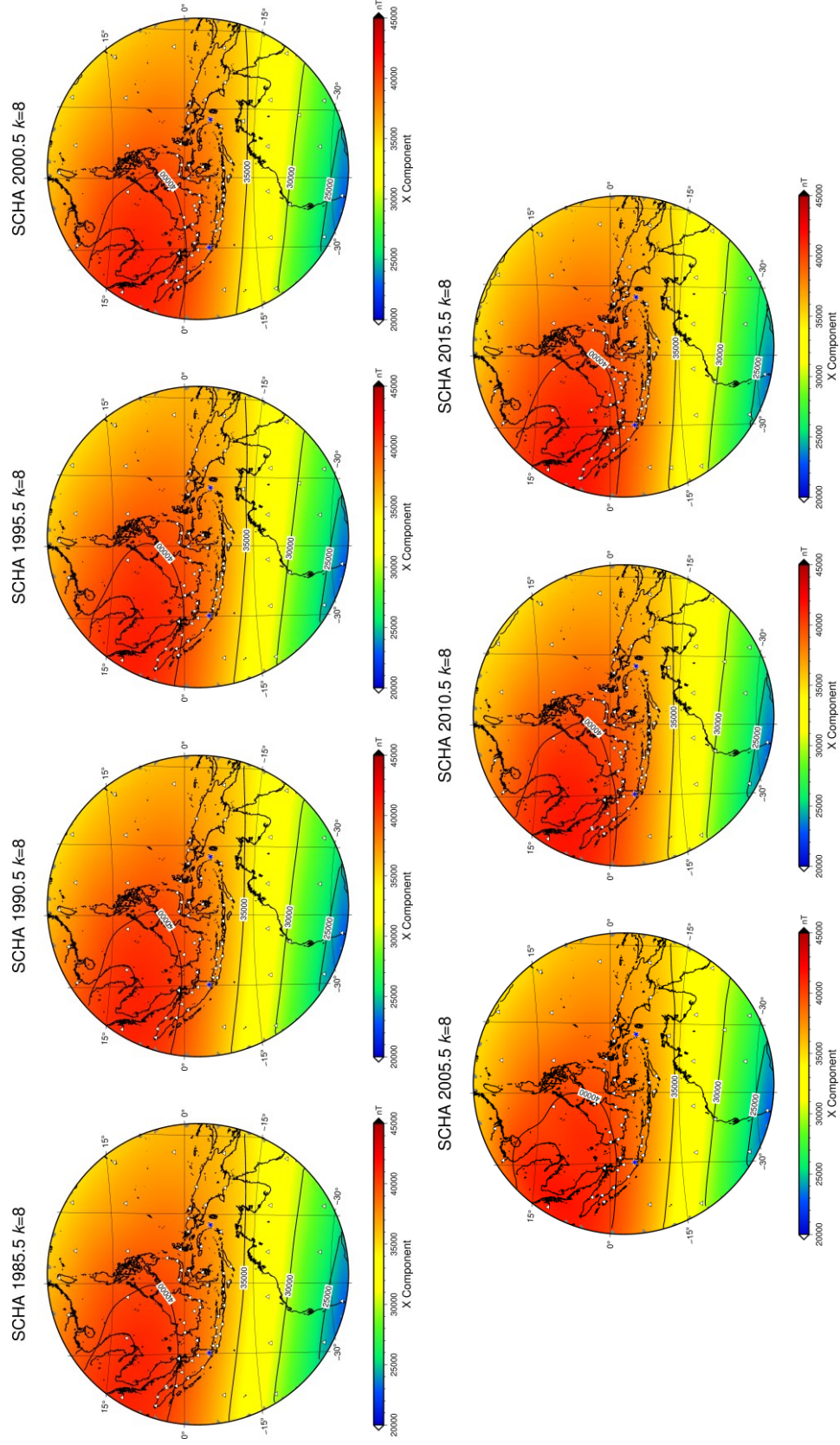
Table 6. RMSD from the SCHA model with index $k = 8$

Epoch	X (nT)	Y (nT)	Z (nT)
2015.5	111.96	93.26	126.52
2010.5	134.83	65.20	123.07
2005.5	87.19	103.89	138.87
2000.5	115.90	68.87	111.61
1995.5	108.15	73.88	119.38
1990.5	120.33	107.4	133.42
1985.5	105.61	91.25	130.48

Again, the RMSD of the SCHA shows improvement in comparison with the RMSD of the IGRF. For X, Y, and Z components, the average RMSD of the SCHA model with index $k = 8$ is 112 nT, 86 nT, and 126 nT, respectively. Indeed, compared with the RMSD of the SCHA with index $k = 7$, the SCHA with index $k = 8$ gives a better result.

As reviewed previously, the positions of the extrema of the RMSD from the SCHA with index $k = 8$ are similar to the positions with index $k = 7$, with the exception of the maximum RMSD of the Y component. While the maximum RMSD of the Y component on the SCHA with $k = 7$ is in 2005.5, on the SCHA with $k = 8$, it is in 1990.5. In 1990.5, the RMSD of the Y component is 107.4 nT, and in 2005.5, it is 103.89. Hence, the difference is not significant. These similar extrema positions on the SCHA model with different index k show that the results from the SCHA calculation are stable.

The contour map of the SCHA with index $k = 8$ is mostly similar to the contour map of the IGRF and the SCHA with index $k = 7$. An agonic line, a line on which the declination is zero, presents just to the west of Indonesia, adjacent with the boundary, for epochs 2000.5 and 2005.5. However, this agonic line does not cross Indonesia. This line is caused by the limitations of the observation data on that area. Note that the number of observations data is different in each epoch.

Figure 18. X component contour from the SCHa using index $k=8$ from 1985.5 – 2015.5.

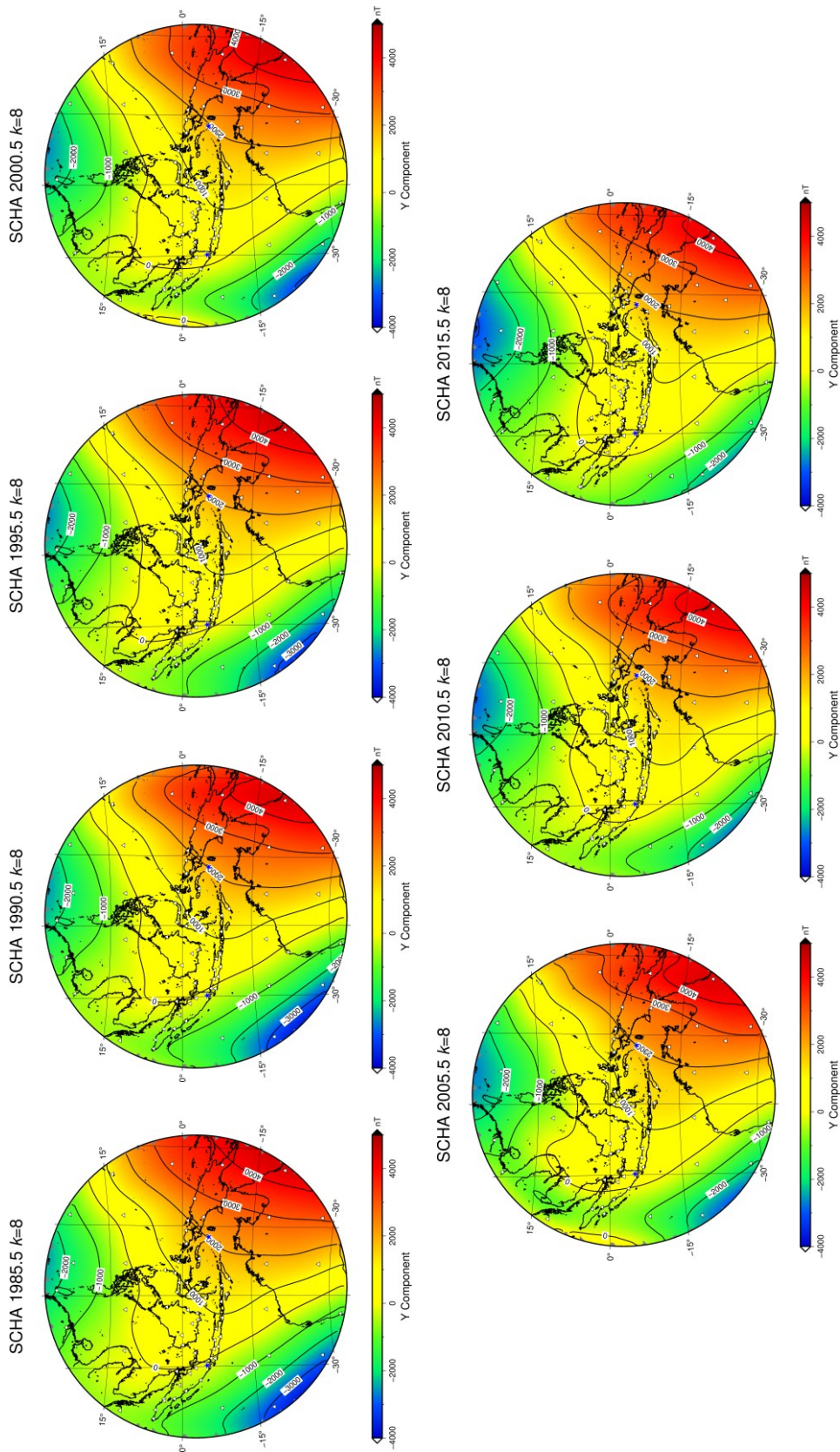


Figure 19. Y component contour from the SCHa using index $k=8$ from 1985.5 – 2015.5.

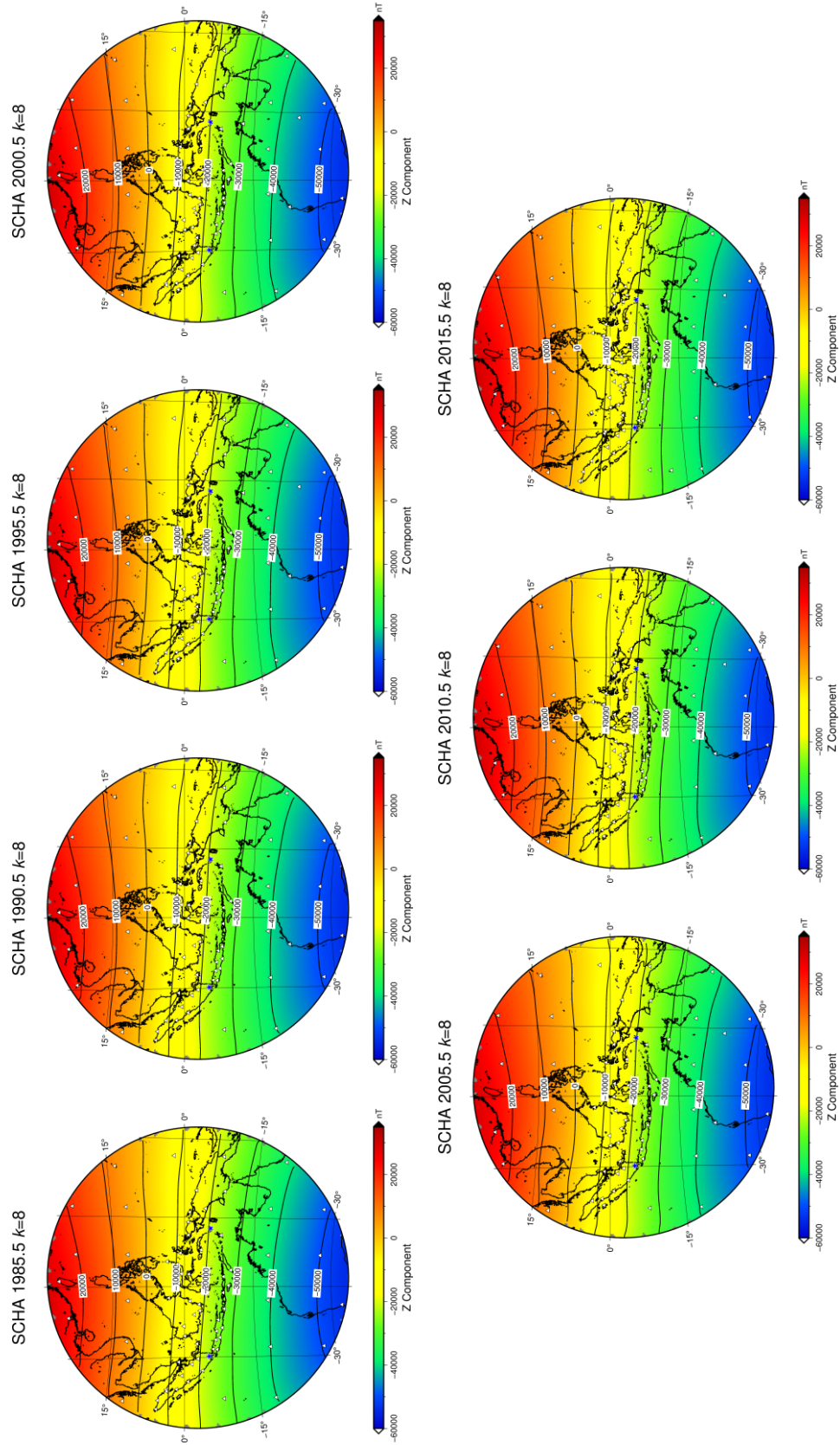


Figure 20. Z component contour from the SCHA using index $k=8$ from 1985.5 – 2015.5.

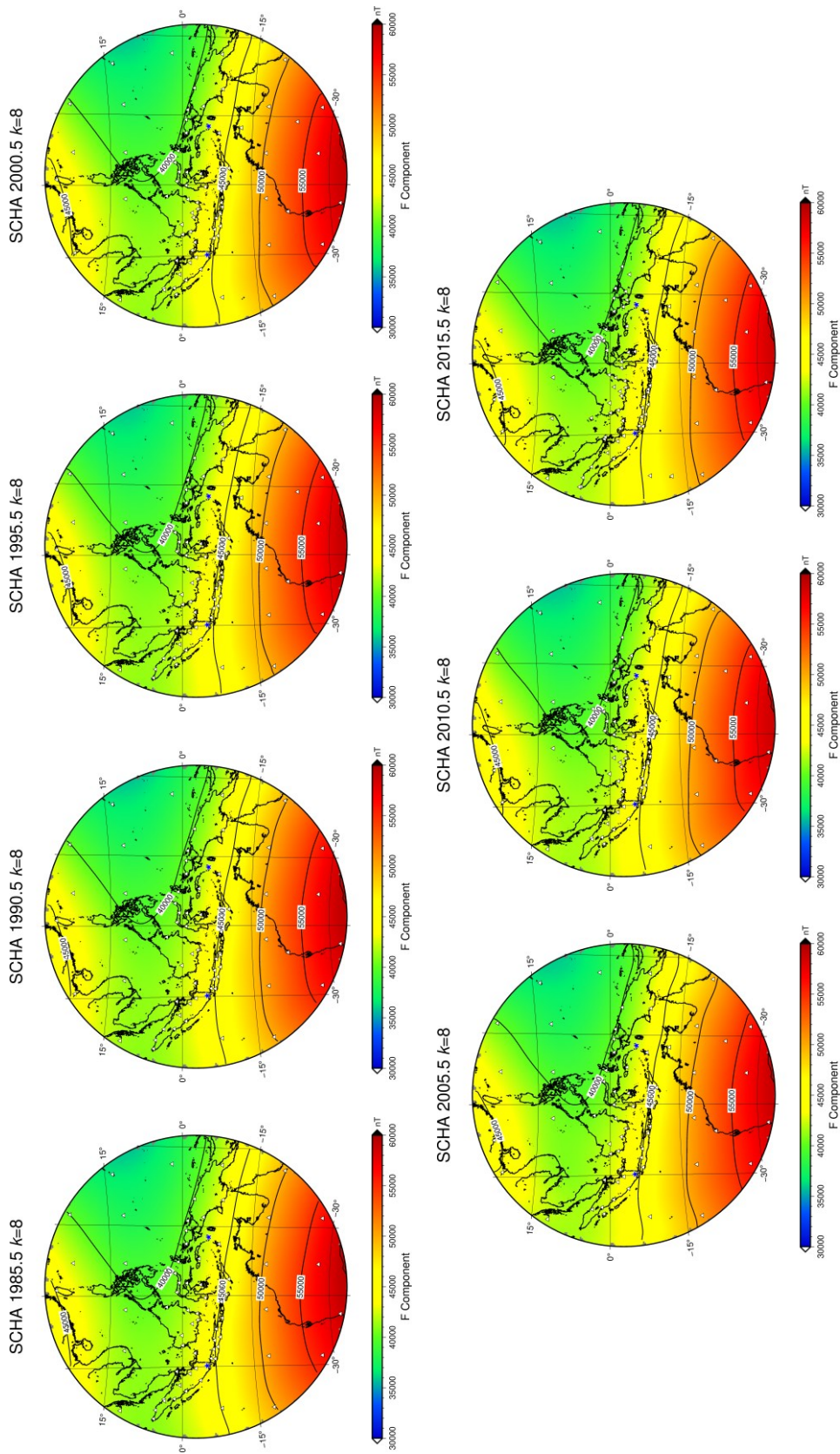


Figure 21. F component contour from the SCHa using index $k=8$ from 1985.5 – 2015.5.

4.1.2.4 SCHA with Spatial Truncation Index $k = 9$

The last SCHA model which will be discussed is that with spatial truncation index $k = 9$. SCHA models with higher index k will not be discussed, given the erratic contouring that these models show around the boundary of the spherical cap. The minimum wavelength which can be represented with $k = 9$ is approximately 1400 km, equal to the SHA model with a maximum truncation degree of 28. The RMSD from the SCHA model with spatial truncation index $k = 9$ is shown in Table 7, and the contour map is shown in Figure 22 through to Figure 25.

Table 7. RMSD from the SCHA model with index $k = 9$

Epoch	X (nT)	Y (nT)	Z (nT)
2015.5	109.78	88.46	121.63
2010.5	129.82	64.46	111.45
2005.5	83.35	100.31	129.74
2000.5	112.54	64.81	103.14
1995.5	105.76	71.38	113.55
1990.5	112.78	102.02	125.95
1985.5	105.81	82.49	117.25

The RMSD of the SCHA model is again improved compared to the RMSD of the IGRF model and the other SCHA model with lower index k . For X, Y, and Z components, the average RMSD of the SCHA model with index $k = 9$ is 109 nT, 82 nT, and 118 nT, respectively. The positions of the extrema in fit are similar to the extrema positions of the SCHA with index $k = 7$ and mostly similar to the extrema position of the SCHA with index $k = 8$. Thus, it supports the suggestion that the SCHA model is stable and adequate for regional geomagnetic modelling.

Although the RMSD of the SCHA with spatial truncation index $k = 9$ shows a promising result, the contour map needs to be inspected thoroughly. The contour map of the X, Z, and F components shows an insignificant difference, compared to the contour maps of the IGRF model and the SCHA with lower index k . However, the contour map of the Y component shows a new feature that does not exist in the other models. The agonic line, which emerges to the west of Indonesia in the SCHA with index $k = 8$, develops into a new feature in the SCHA with index $k = 9$ in the western part of Indonesia in 2005.5. In 2010.5, this feature is also visible in the same area, despite the absence of the agonic line in the SCHA with index $k = 8$ for this

epoch. Furthermore, instead of developing into a new feature, the agonic line in 2000.5 with index $k = 8$ remains unchanged in the SCHA with index $k = 9$. Thus, the RMSD of the model cannot be the sole basis in deciding the best model. It is also necessary to inspect the contour maps of the model for unrealistic features. In the SCHA with index $k = 9$, although the RMSD shows an improvement, the contour map suggests some deterioration in this regard.

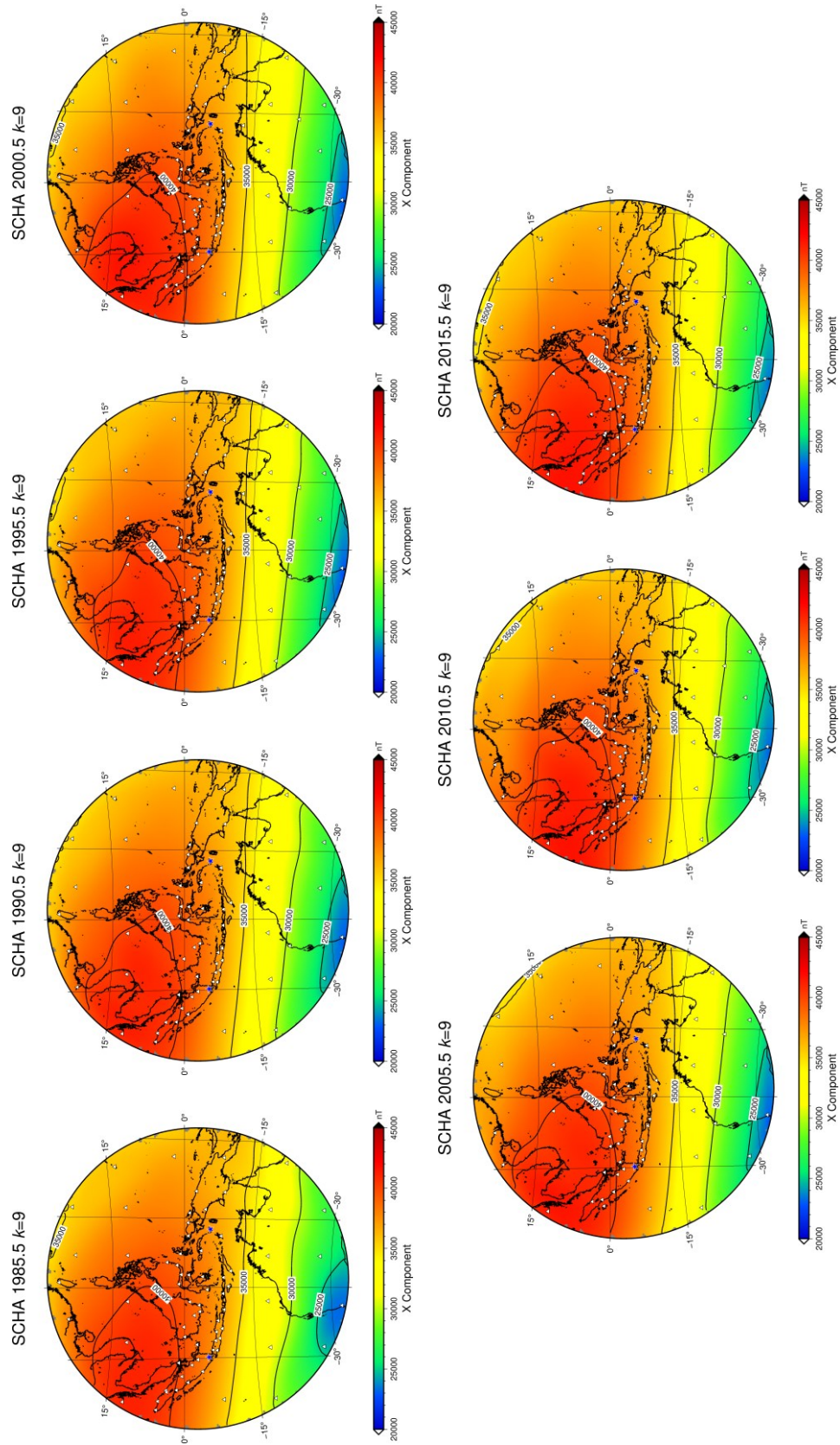


Figure 22. X component contour from the SCHa using index $k=9$ from 1985.5 – 2015.5.

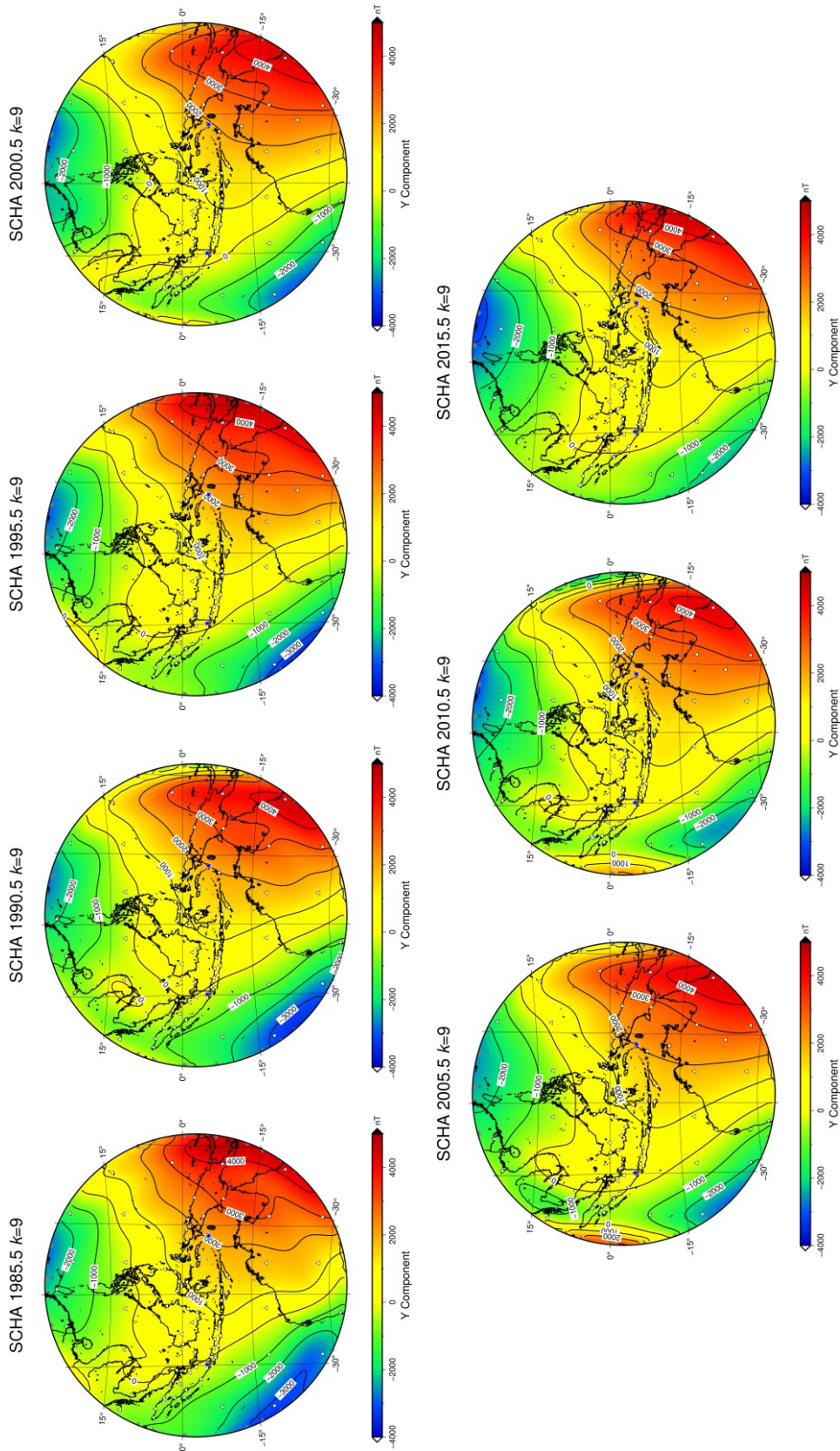
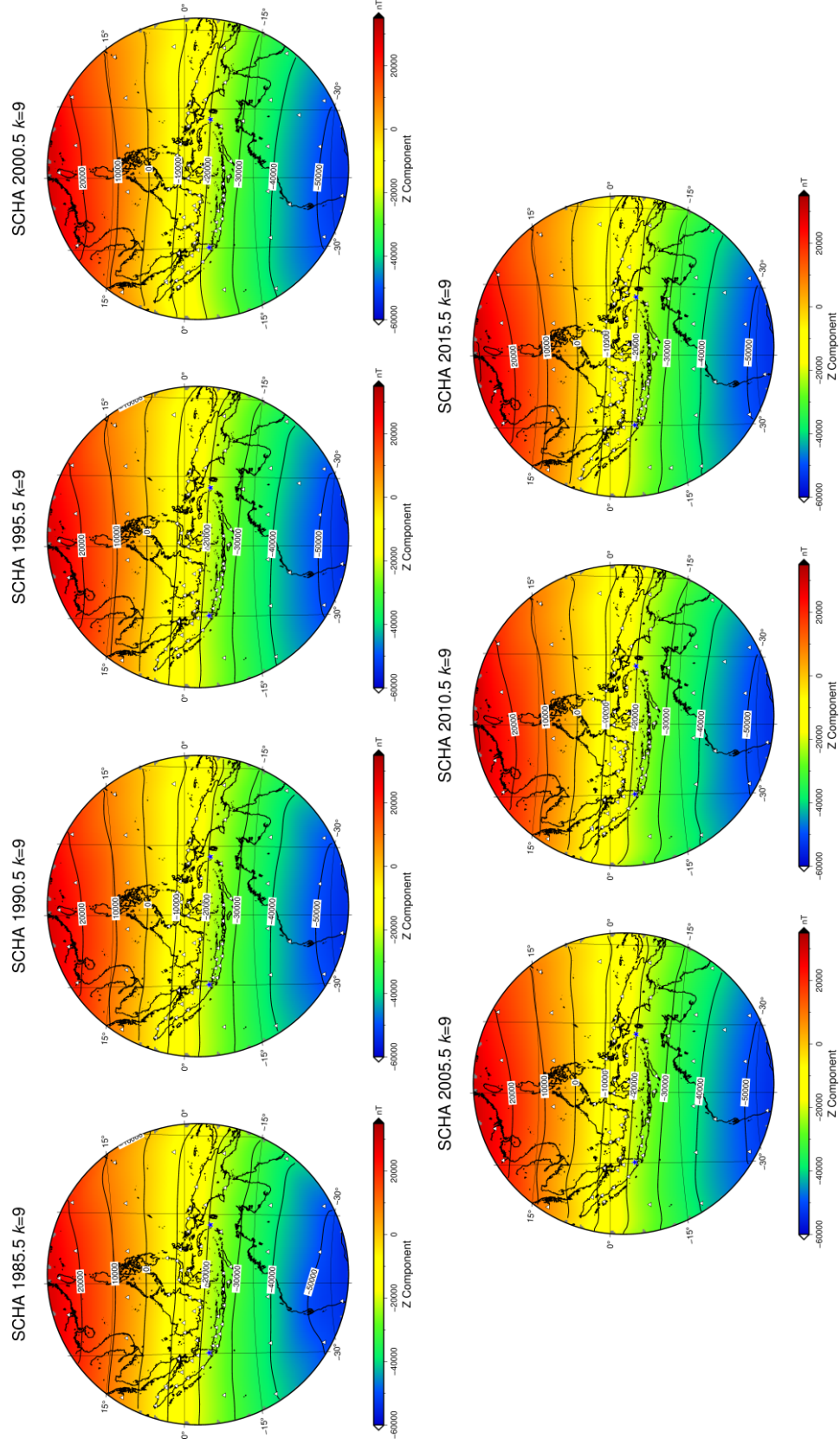


Figure 23. Y component contour from the SCHa using index $k=9$ from 1985.5 – 2015.5.

Figure 24. Z component contour from the SCHA using index $k=9$ from 1985.5 – 2015.5.

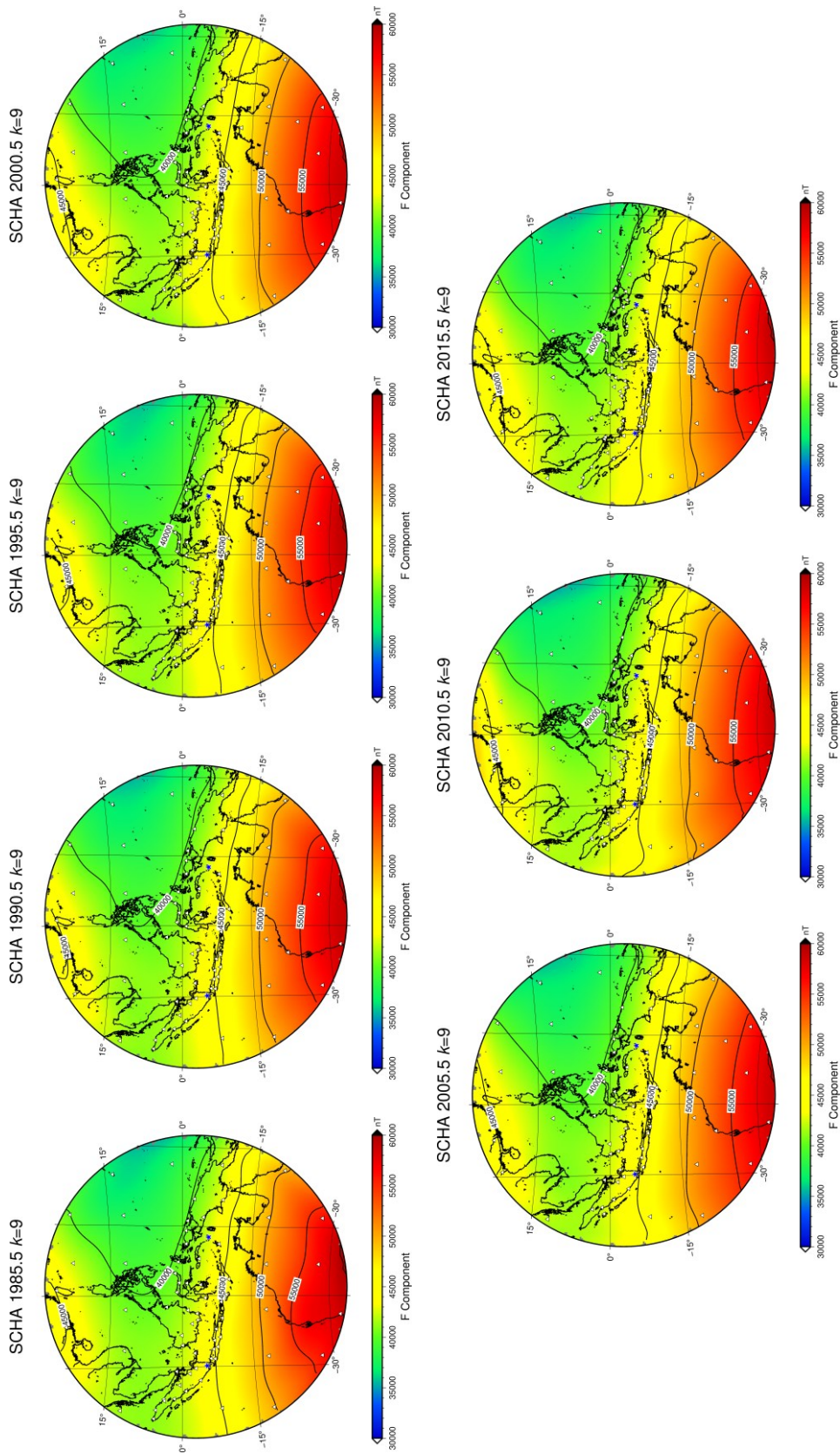


Figure 25. F component contour from the SCHa using index $k=9$ from 1985.5 – 2015.5.

4.1.3 Validation of the Model

In order to confirm that the outputs of the model are acceptable, model validation is necessary. There are two types of model validation. The first type is usually conducted by analysing the goodness of fit of the model, and it involves data that was used in the construction of the model. The second type usually uses data that was not used in model construction.

In this chapter, the validation of the model incorporates examining how far the constructed model differs from the real data in locations that are not used to develop the model. The models which are validated are the IGRF model and the SCHA with index $k = 1 - 10$. The validations are based on the observation data at two repeat stations, Bandar Lampung and Tual. These locations are selected to make a balanced comparison as Bandar Lampung is located on the western side of Indonesia, and Tual is located on the eastern side of Indonesia. The validations are calculated by subtracting the model value with the real observation data. The model can be considered as a better model if the difference is small.

4.1.3.1 Bandar Lampung

Located at the western side of Indonesia, Bandar Lampung is surrounded by several repeat stations and geomagnetic observatories. The nearest geomagnetic observatories are Tangerang (TNG) and Pelabuhanratu (PLR), located at 170 km and 200 km from Bandar Lampung, respectively. The difference between the IGRF and SCHA models with the real observed data on each component is shown in Figure 26.

For each component, there is a pattern as the index k increases. However, the pattern is not consistent for each year. As an example, in 1985.5 and 1990.5, the difference in the X component decreases as index k increases. However, in 2005.5 and 2010.5, this value increases as the index k increases. Another trend is seen in 1995.5 when the difference in the X component decreases for index $k = 1 - 6$ then increases afterwards. This inconsistent pattern also applies to the other components.

Furthermore, the minimum difference value of the SCHA model does not consistently occur at the same index k for each year for each component. Interestingly, the minimum difference value of the SCHA model is not always smaller than the difference value of the IGRF model. It means that the SCHA model for every index k is not always better than the IGRF model.

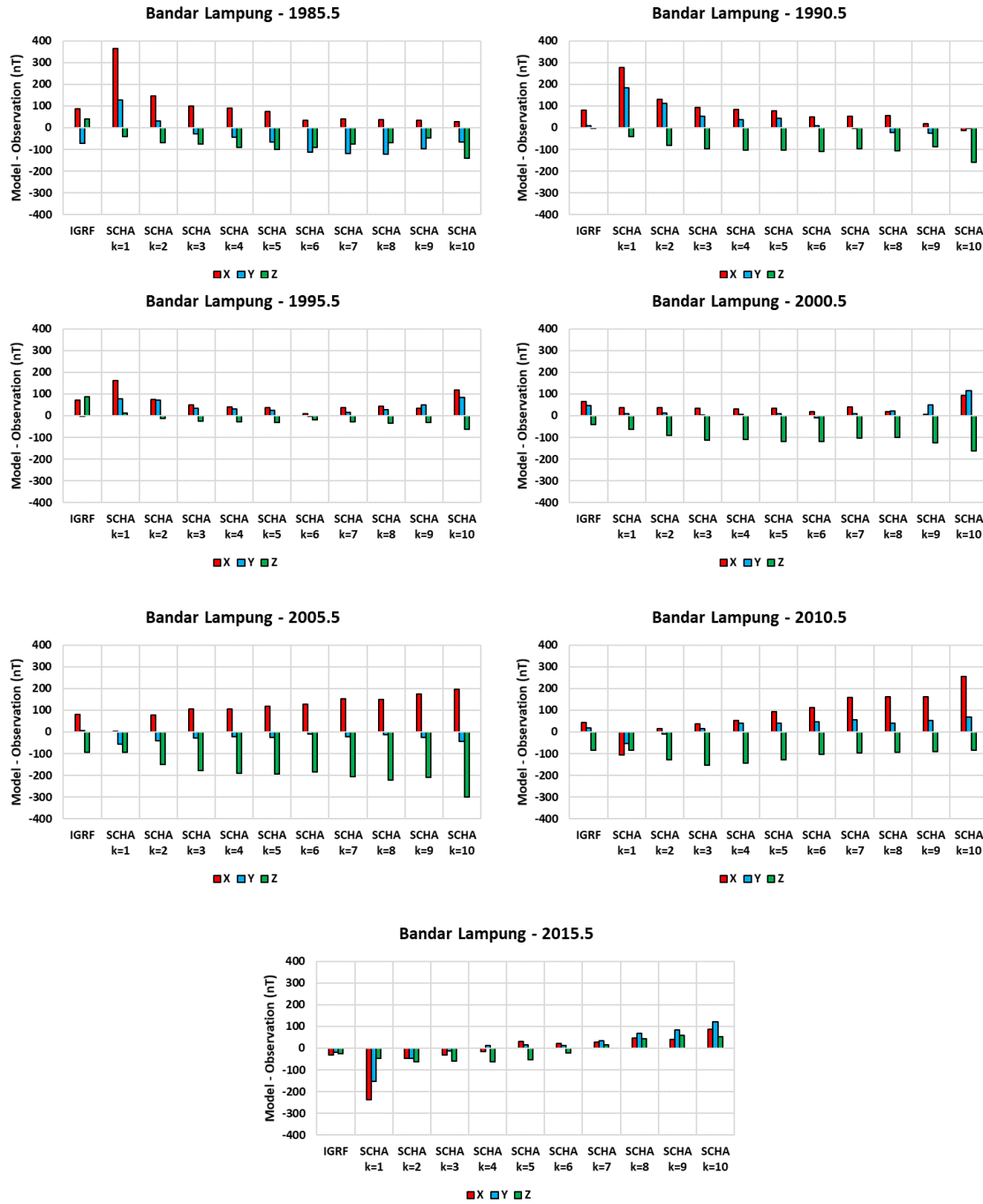


Figure 26. Differences between models and the observed field at Bandar Lampung from 1985.5 – 2015.5. The far-left graph is the difference between the IGRF model and the observation data, while the rest is the difference between the SCHa with different index k and the observation data.

However, looking at the difference in each component is quite complicated. The RMSD on each model is calculated to make it straightforward. The RMSD is calculated by the square root of the sum of the squares of the difference of all the components. The results of this calculation are shown in Figure 27. A value that is smaller than the red bar is considered better than the IGRF model.

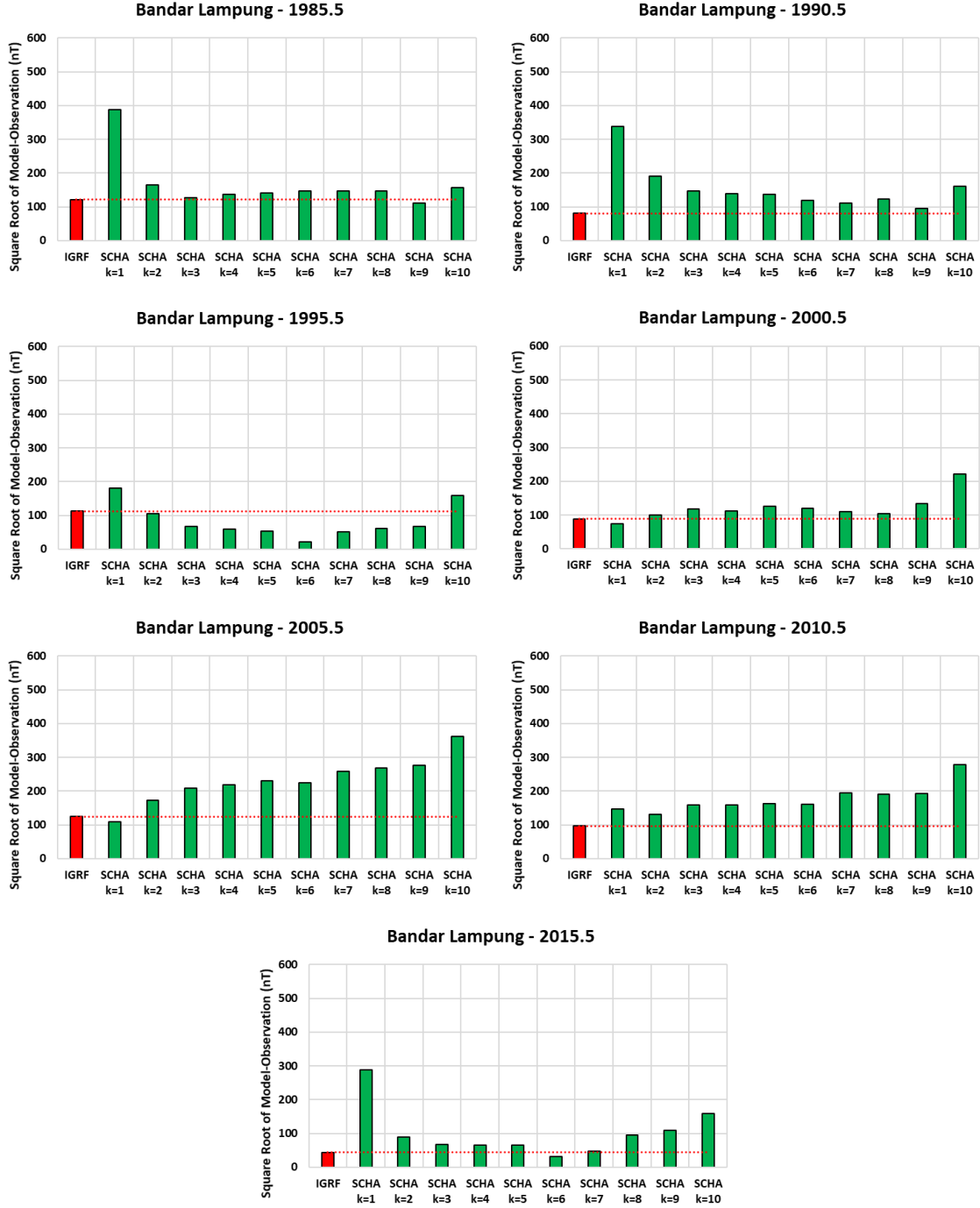


Figure 27. The RMSD between models and the observed field at Bandar Lampung from 1985.5 – 2015.5. The dotted red line is obtained from the RMSD of the IGRF model and used to determine whether the RMSD of the SCHA model is better than the RMSD of the IGRF model or not.

Again, in the main, the results show that the IGRF is still better than the SCHA model at Bandar Lampung. The graphs show that the pattern is still inconsistent for each year. Some graphs have an increasing pattern as the index k increases, but the graph in other years shows a decreasing trend. The minimum value of the RMSD at

the SCHA model occurs in 1995.5 at index $k = 6$. This RMSD is about three times smaller than the RMSD of the IGRF model. In contrast, the maximum RMSD at the SCHA model occurs in 1985.5 at index $k = 1$. The RMSD is approaching 400 nT, whereas the RMSD at the IGRF model in the same epoch is only a little more than 100 nT.

From this analysis, it can be concluded that the IGRF model at Bandar Lampung is still better than the SCHA model for every spatial truncation index k for every epoch. However, to ascertain whether this situation only happens at Bandar Lampung or might happen at the repeat stations or geomagnetic observatories located around it, an in-depth analysis will be discussed later. This is needed to determine whether this case occurs due to modelling error or because the data itself is not accurate in representing the real geomagnetic field at this location.

4.1.3.2 Tual

The next location which is used to validate the reliability of the model is Tual which is located in the eastern part of Indonesia. Several repeat stations surround Tual, but there is no geomagnetic observatory nearby. The nearest geomagnetic observatories are Kakadu (KDU), an INTERMAGNET observatory that is located more than 700 km to the south, and Kupang (KPG) which is located more than 1000 km to the west. The difference between the IGRF and SCHA models with the real observed data on each component in Tual is shown in Figure 28.

Again, from Figure 28, the pattern is inconsistent for each component in each epoch. The difference value between the real observed data and the SCHA model fluctuates as the index k increases. However, in 2005.5, the difference in all the components gradually increases as the index k increases. A distinctive feature can be seen in 1995.5 in the SCHA model with index $k = 10$ when the value increases dramatically for all components. Nevertheless, the difference value between the real observed data and the SCHA model at Tual is slightly smaller than the similar value at Bandar Lampung.

Besides that, most of the SCHA models for all the epochs give a smaller difference value than the IGRF model. Even in 1995.5, when there is an increase in the difference value in the SCHA model with index $k = 10$, the rest of the SCHA models with index $k < 10$ give a better result than the IGRF model.

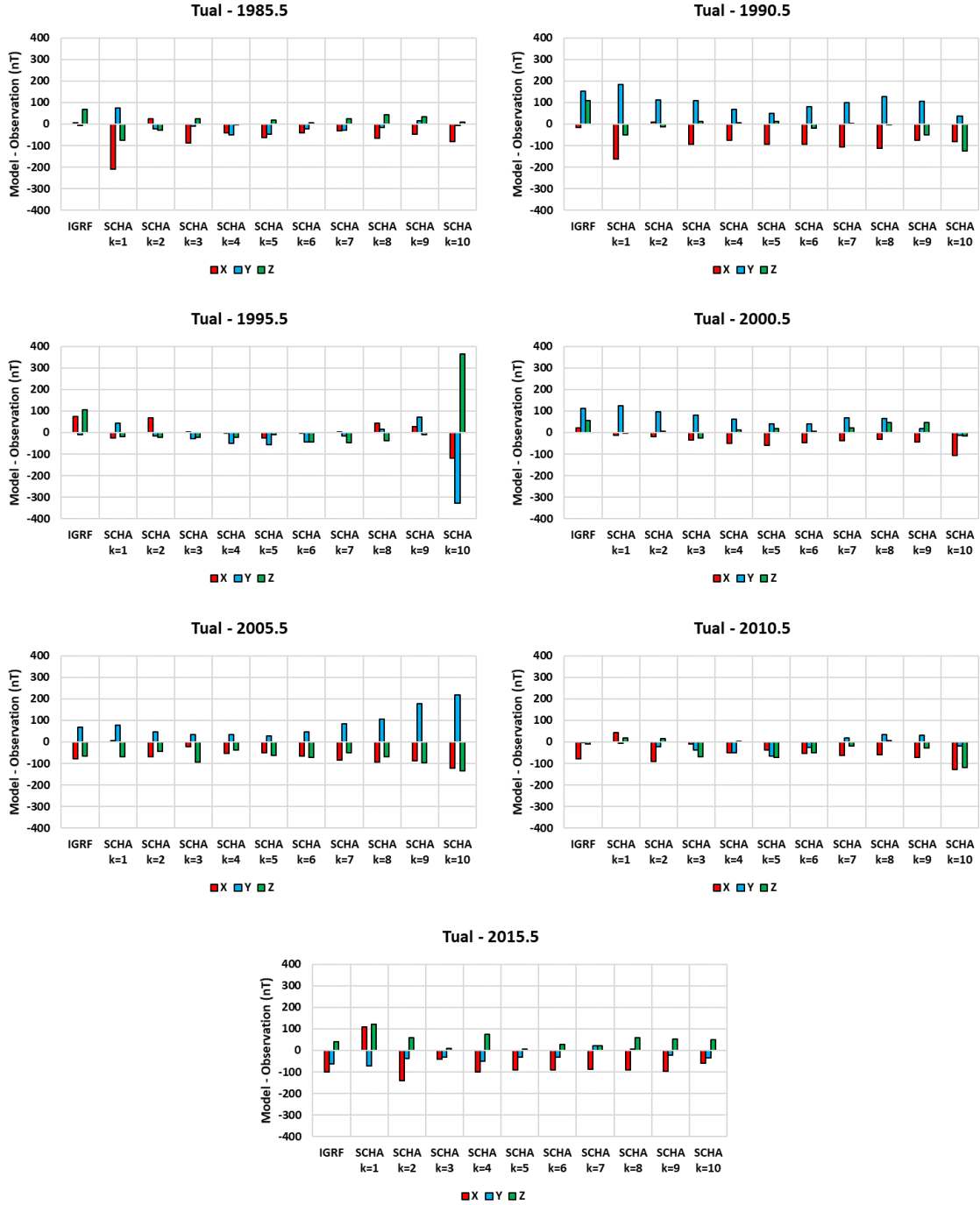


Figure 28. Differences between models and the observed field at Tual from 1985.5 – 2015.5. The far-left graph is the difference between the IGRF model and the observation data, while the rest is the difference between the SCHA with different index k and the observation data.

As has been done at Bandar Lampung, the RMSD between each model and the real observed data in Tual is calculated. The results of this calculation are shown in Figure 29. The red bar is the value for the IGRF model, and it is considered as a baseline to conclude that the SCHA model is better than the IGRF model.

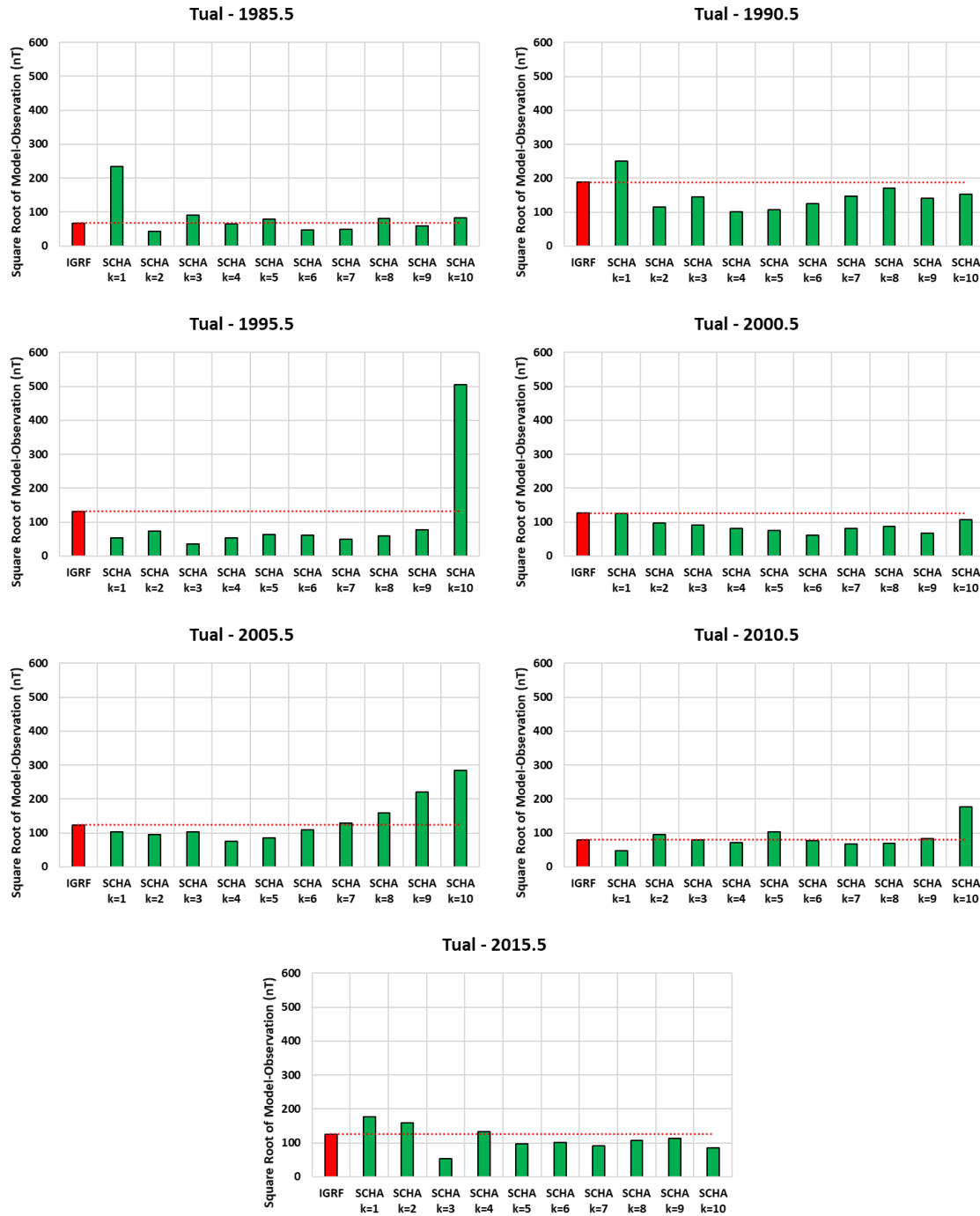


Figure 29. The RMSD between models and the observed field at Tual from 1985.5 – 2015.5. The dotted red line is obtained from the RMSD of the IGRF model and used to determine whether the RMSD of the SCHA model is better than the RMSD of the IGRF model or not.

From Figure 29, it is unquestionably the case that most of the SCHA models are better than the IGRF models. An exception is in 2005.5 when the IGRF model gives a better result than several SCHA models. The large jump for $k = 10$ in 1995.5 is still visible, and its value is more than four times higher than the value of the IGRF

model. Apparently, in this case, the regional model behaves better than the global model when comparing its RMSD.

By comparing Figure 27 and Figure 29, on average, over the 30 years, the RMSD of the IGRF model at Bandar Lampung and Tual are mostly similar, that is approximately 100 nT. On the other hand, the RMSD of the SCHA models at Bandar Lampung is higher than the RMSD of the SCHA models at Tual. Therefore, the problem that arises lies in the SCHA model. The geomagnetic observatories near Tual, e.g., Kakadu (KDU) and Kupang (KPG), although located far away from Tual, can still be modelled well by the IGRF calculation.

Although the RMSD of the SCHA model is smaller than the RMSD of the IGRF model at Tual, the minimum value position is inconsistent for each epoch. However, frequently, at index $k > 7$, the RMSD of the SCHA model is higher than the RMSD of the IGRF model.

4.1.4 Difference between the Regional and Global Model

The difference between SCHA and the IGRF model is calculated to understand the model reliability around the research area. The calculation is done by creating a grid with a size of 30 minutes arc on all the models and calculating the difference between the SCHA and the IGRF model. The calculation and mapping are done using Generic Mapping Tools version 6 (Wessel et al., 2019). As has been done on the analysis of the stability of the model, this analysis will be done on the SCHA model with index $k = 7 - 9$ only.

Areas shown in yellow indicate where the SCHA model gives similar values to the IGRF model. The negative values, shown as green to blue, mean the SCHA model gives smaller values than the IGRF model, and the positive values, shown as orange to red, mean conversely that the SCHA model gives larger values than the IGRF model. The colour scale is limited to ± 500 nT to maintain the readability of the contours. Any value which is higher than 500 nT is shown as a black area, while a value that is lower than -500 nT is shown as a white area. However, to know whether the difference is an improvement or not, both models need to be plotted against the observed data.

4.1.4.1 Difference between SCHA Index $k = 7$ and the IGRF Model

The contour map of the difference between SCHA with index $k = 7$ and the IGRF model is shown in Figure 30 – 33. From Figure 30, it seems that the X component of the SCHA model is mostly similar to the IGRF model. The difference is mostly negative from 1985.5 – 2000.5. However, from 2005.5 and onwards, the difference becomes positive. In 2005.5, the white area in the northern part of the research area indicates that the boundary problem of regional geomagnetic modelling still exists, although it is minimised.

For the Y component, the difference is not so significant. Mostly, the contour map is dominated by yellow. An interesting fact is that while the extreme difference value in the X component appears in the northern and southern side of the research area, the extreme difference value in the Y component appears in the eastern and western side of the research area. Moreover, the extreme difference value on one side is always the opposite of the other side.

The difference contour on the Z component does not show an east-west or north-south pattern. The positive and negative values are mostly balanced in each epoch. An interesting feature in this contour map is in 2005.5 when an extremely negative area occurs around Vietnam and an area with extremely positive values is on the eastern side of the research area.

For the F component, the difference value is dominantly positive from 2005.5 onwards, while the rest of it is mostly balanced. The F component in the SCHA model is calculated as the resultant vector from the X, Y, and Z components.

Nonetheless, the IGRF itself is designed to exclude the contribution of the crustal magnetic field (Thébault et al., 2015). Qamili et al. (2010) explained that more detailed results were accomplished by regional modelling when they compared the IGRF and the SCHA model in Albania and Southeast Italy.

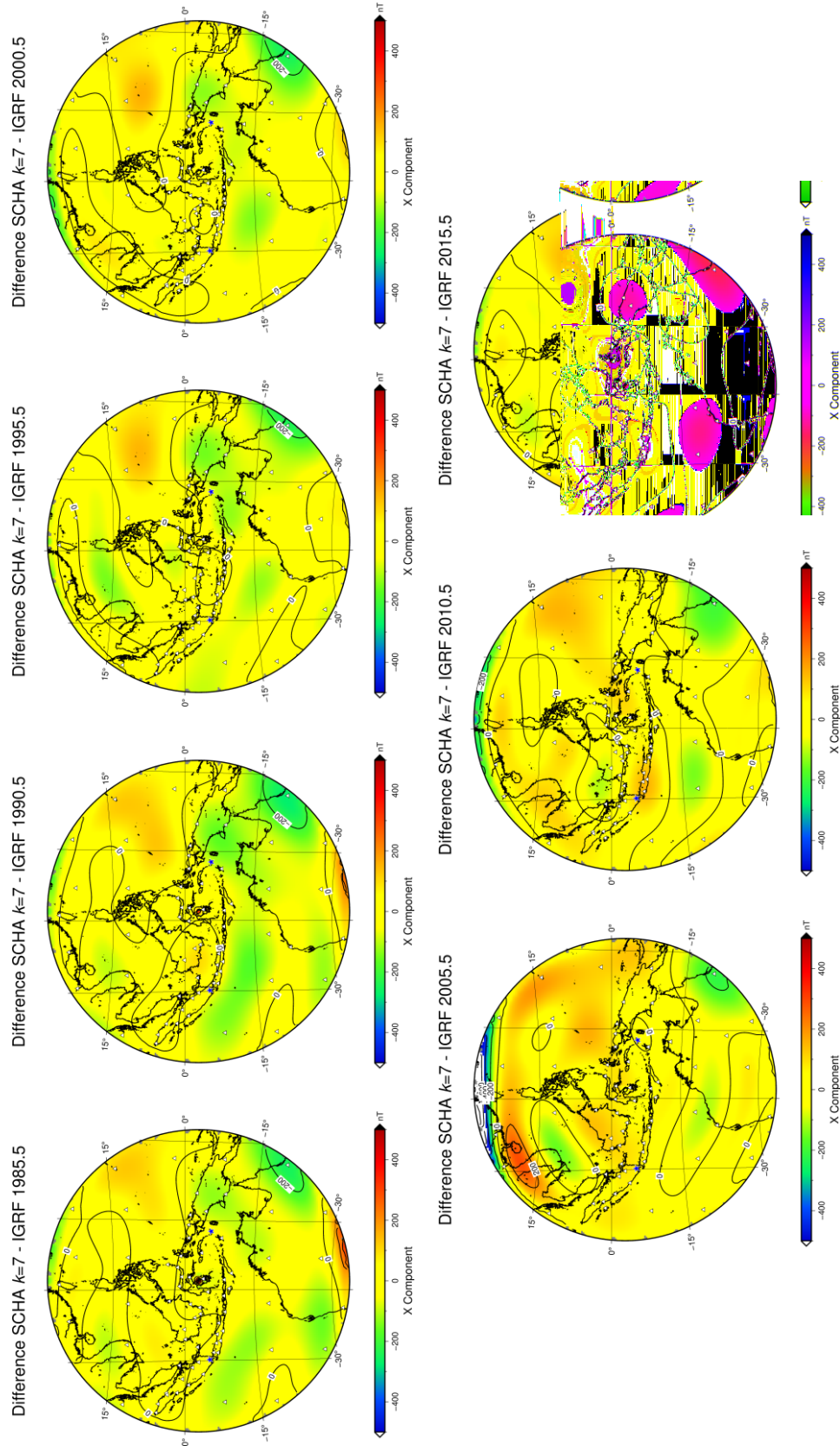
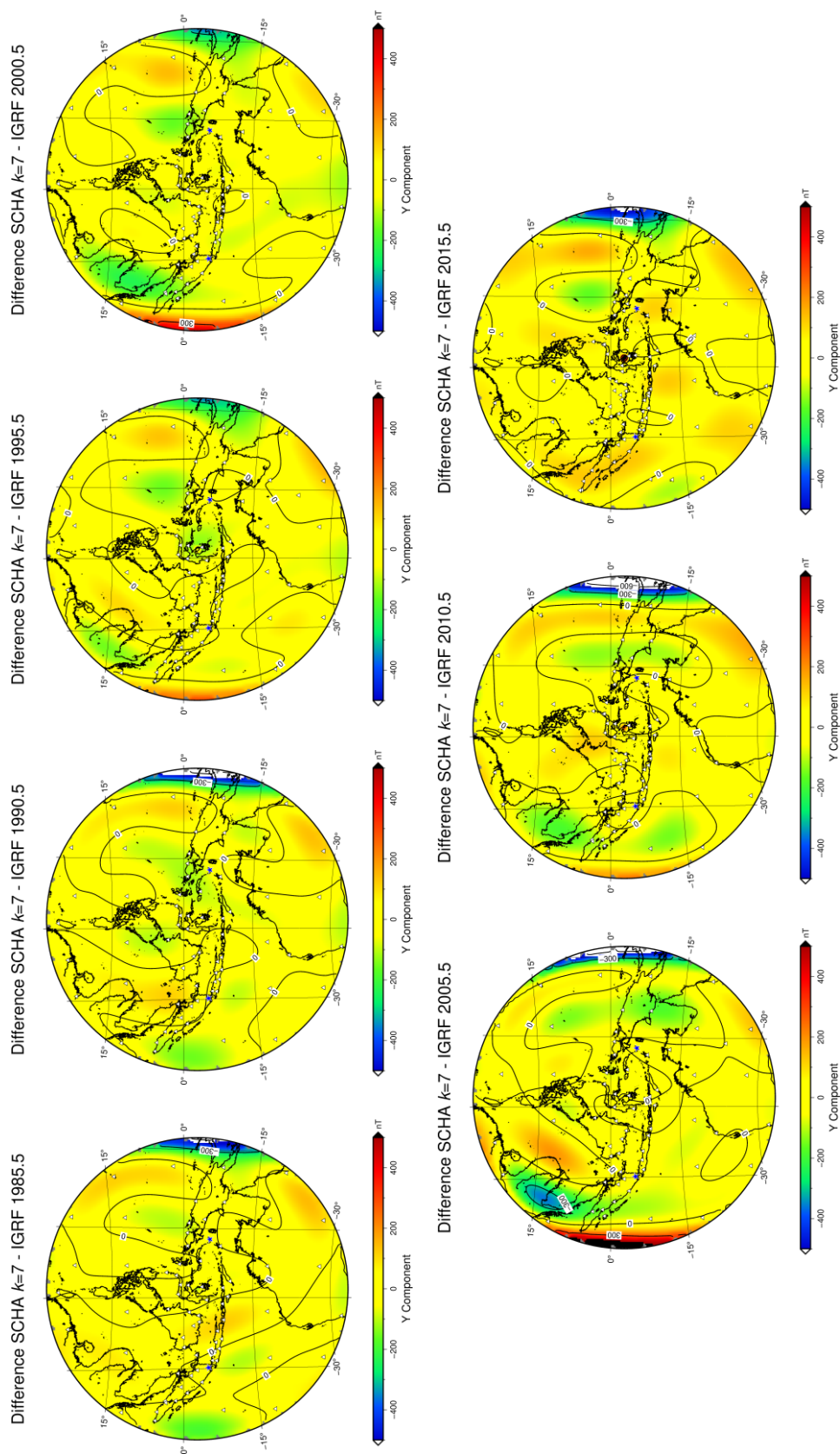


Figure 30. Difference contour at X component between the SCHA using index $k=7$ and IGRF from 1985.5-2015.5.

Figure 31. Difference contour at Y component between the SCH A using index $k=7$ and IGRF from 1985.5-2015.5.

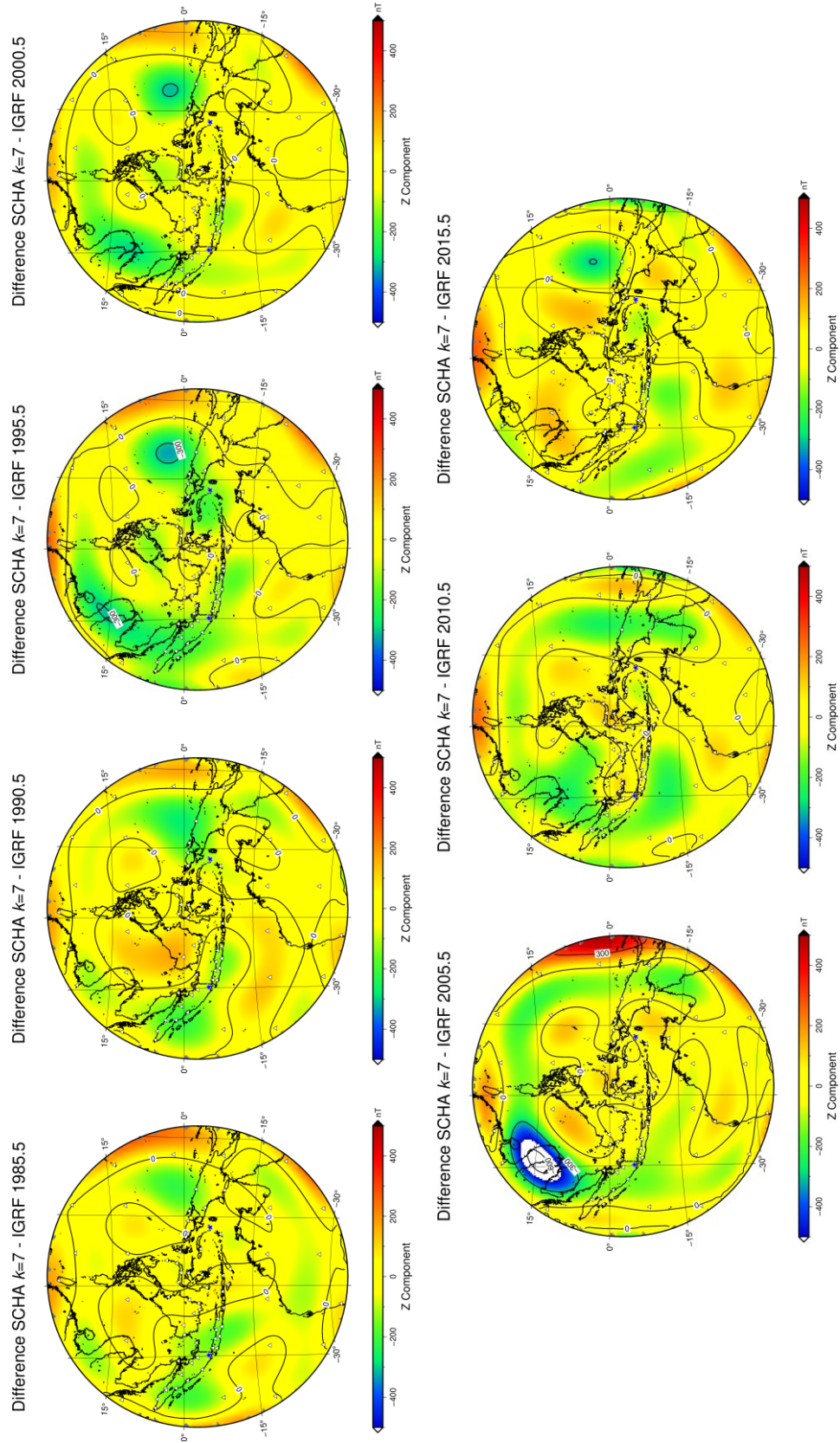


Figure 32. Difference contour at Z component between the SCHa using index $k=7$ and IGRF from 1985.5-2015.5.

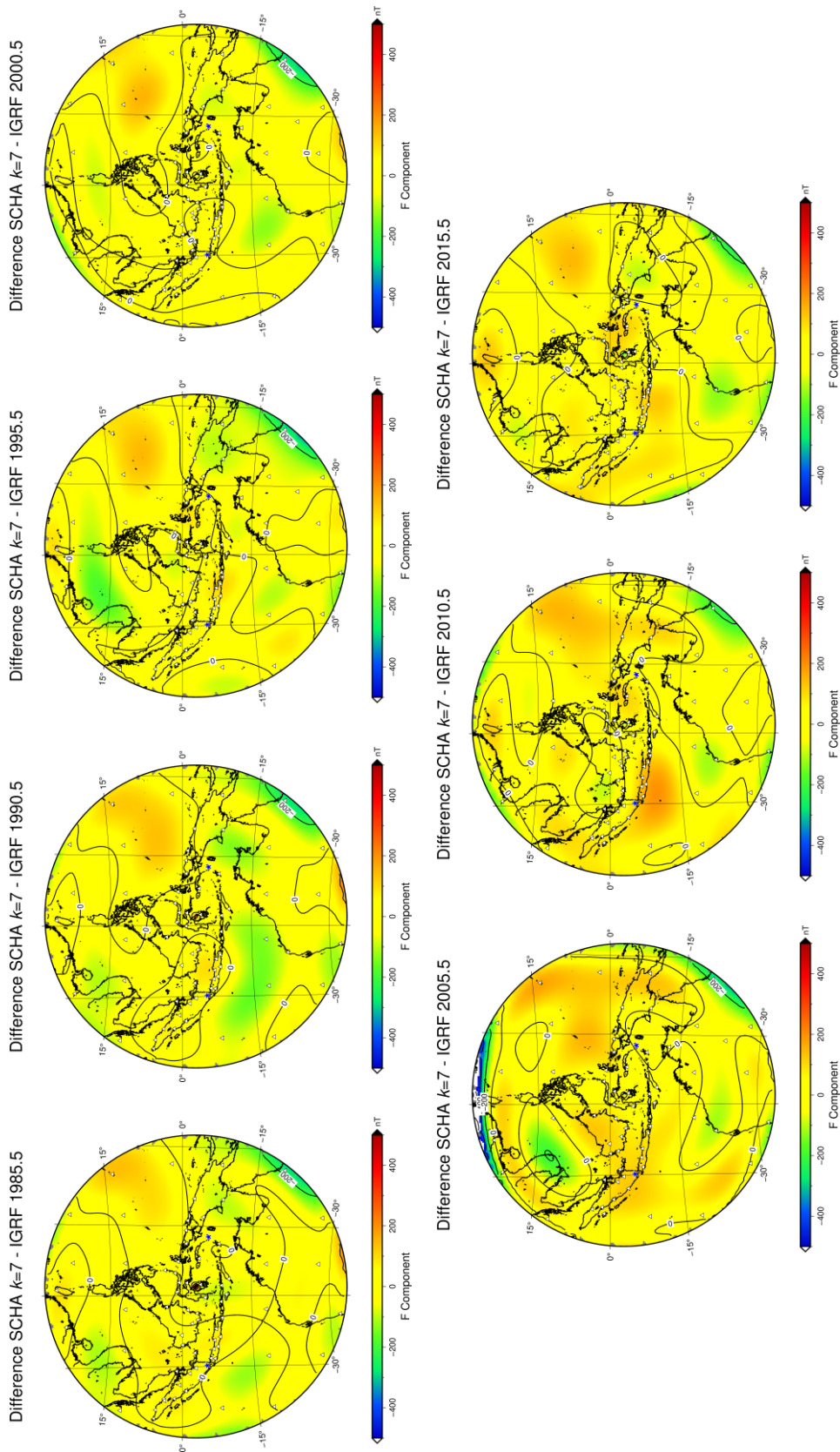


Figure 33. Difference contour at F component between the SCH A using index $k=7$ and IGRF from 1985.5-2015.5.

4.1.4.2 Difference between SCHA Index $k = 8$ and the IGRF Model

The contour map of the difference between SCHA with index $k = 8$ and the IGRF model is shown in Figure 34 – 37. For the X component, the difference contour maps between SCHA from 1985.5 – 2000.5 show mostly a balance between positive and negative values but become dominantly positive from 2005.5 onwards. The extreme negative area in the northern part of the research area in 2005.5 is also still visible, although the area is minimised compared to the difference map with index $k = 7$. However, extremely positive and negative areas arise around Vietnam.

In the Y component, the extreme values show on the eastern and western boundaries in all epochs, except for 2015.5. The extremely negative area always arises on the eastern boundary side, while the extremely positive area always shows on the western boundary side. In 2005.5, the extreme values only appear on the eastern side boundary side with extremely negative values. An extreme negative area also appears around Thailand. The value on this area is just lower than -500 nT compared to the difference map between SCHA with index $k = 7$ and the IGRF model as seen in Figure 31.

For the Z component, in Figure 36, the difference contour values become more significant than for the same map on the previous model, which can be seen in Figure 32. Most of the extreme values occur around the boundary, except in 2005.5 when an extremely negative values area occurred around Vietnam. Compared to Figure 32, the size of this area is larger on the SCHA model with index $k = 8$.

The difference contours for the F component also fluctuate more compared with the previous model in Figure 33. The extremely negative area in 2005.5 is still visible on the northern side, while a positive area appears on the eastern side. This positive area is not visible in the previous model in Figure 33. A positive area also appears on the northern side of the research area in 2010.5, while this feature does not exist in the previous model.

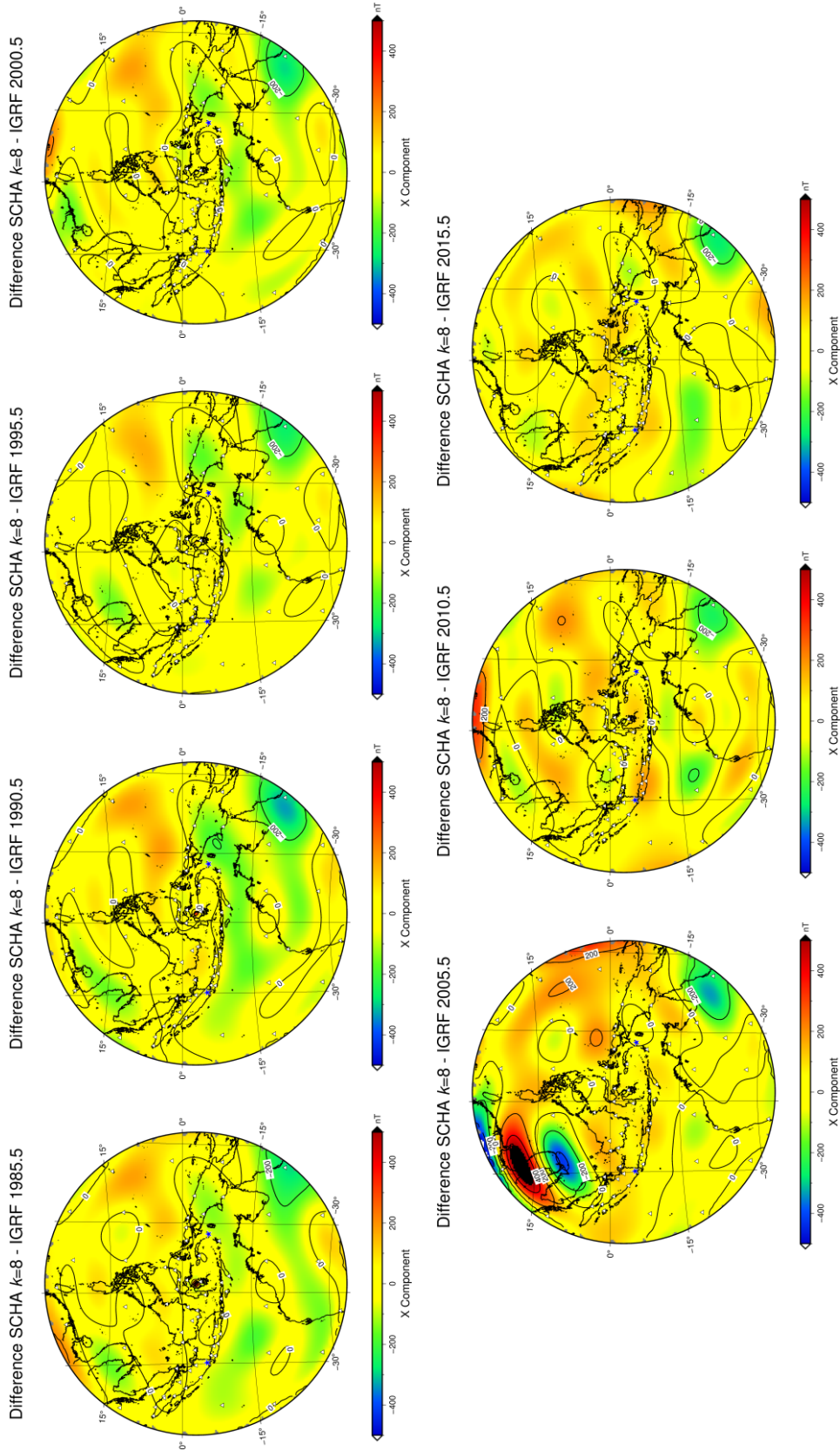


Figure 34. Difference contour at X component between the SCH A using index $k=8$ and IGRF from 1985.5-2015.5.

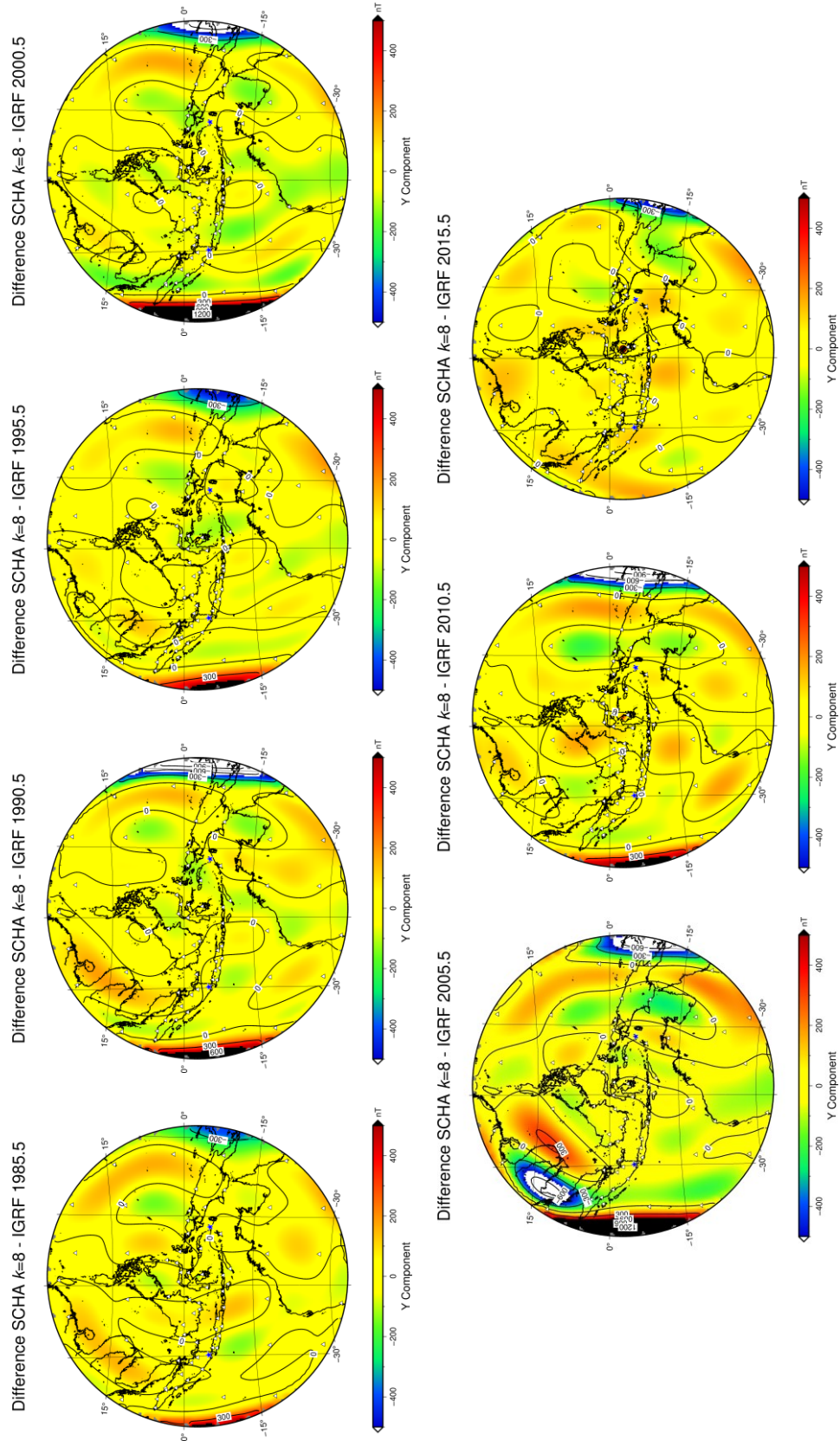


Figure 35. Difference contour at Y component between the SCHA using index $k=8$ and IGRF from 1985.5-2015.5.

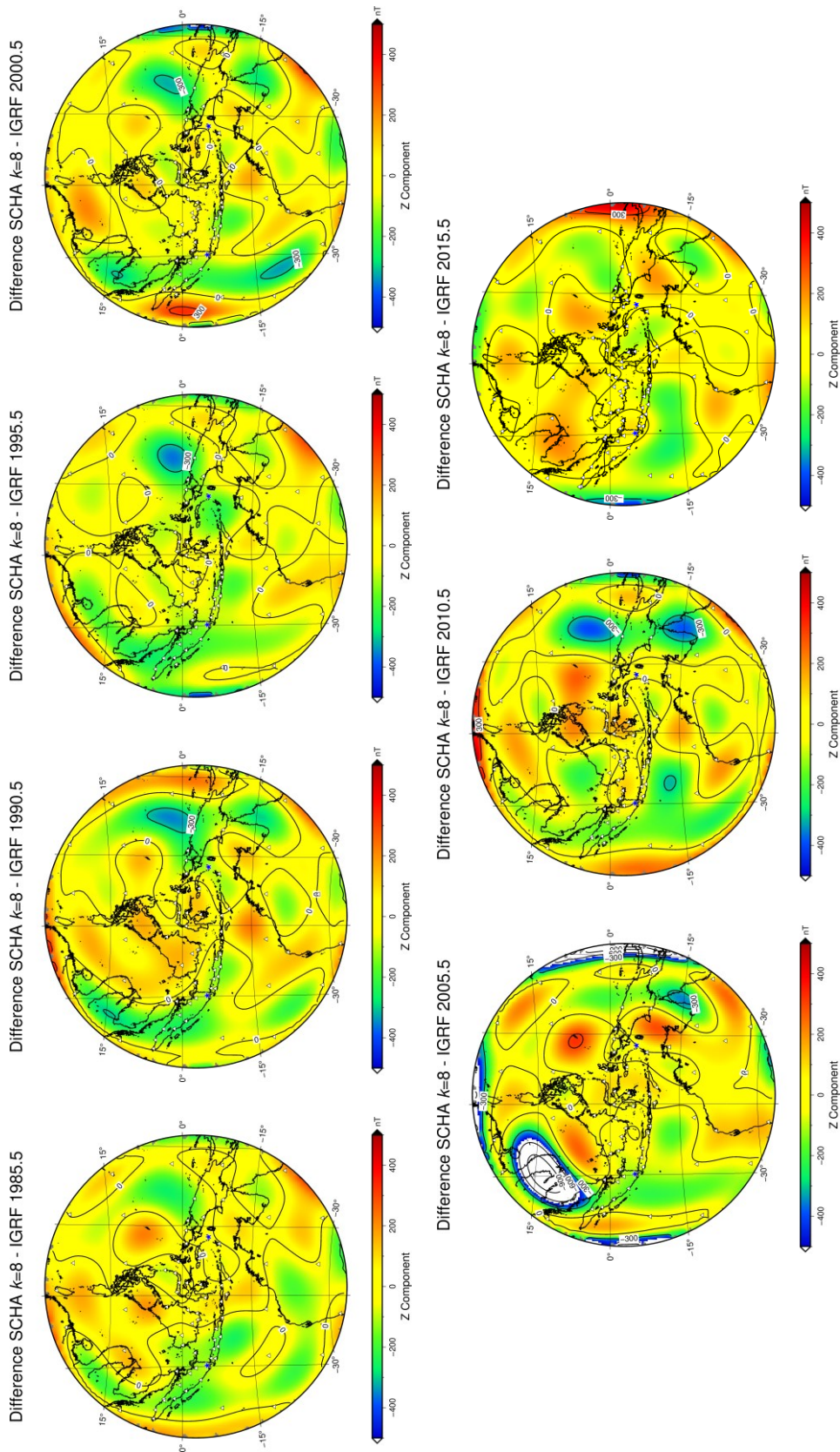


Figure 36. Difference contour at Z component between the SCH A using index $k=8$ and IGRF from 1985.5-2015.5.

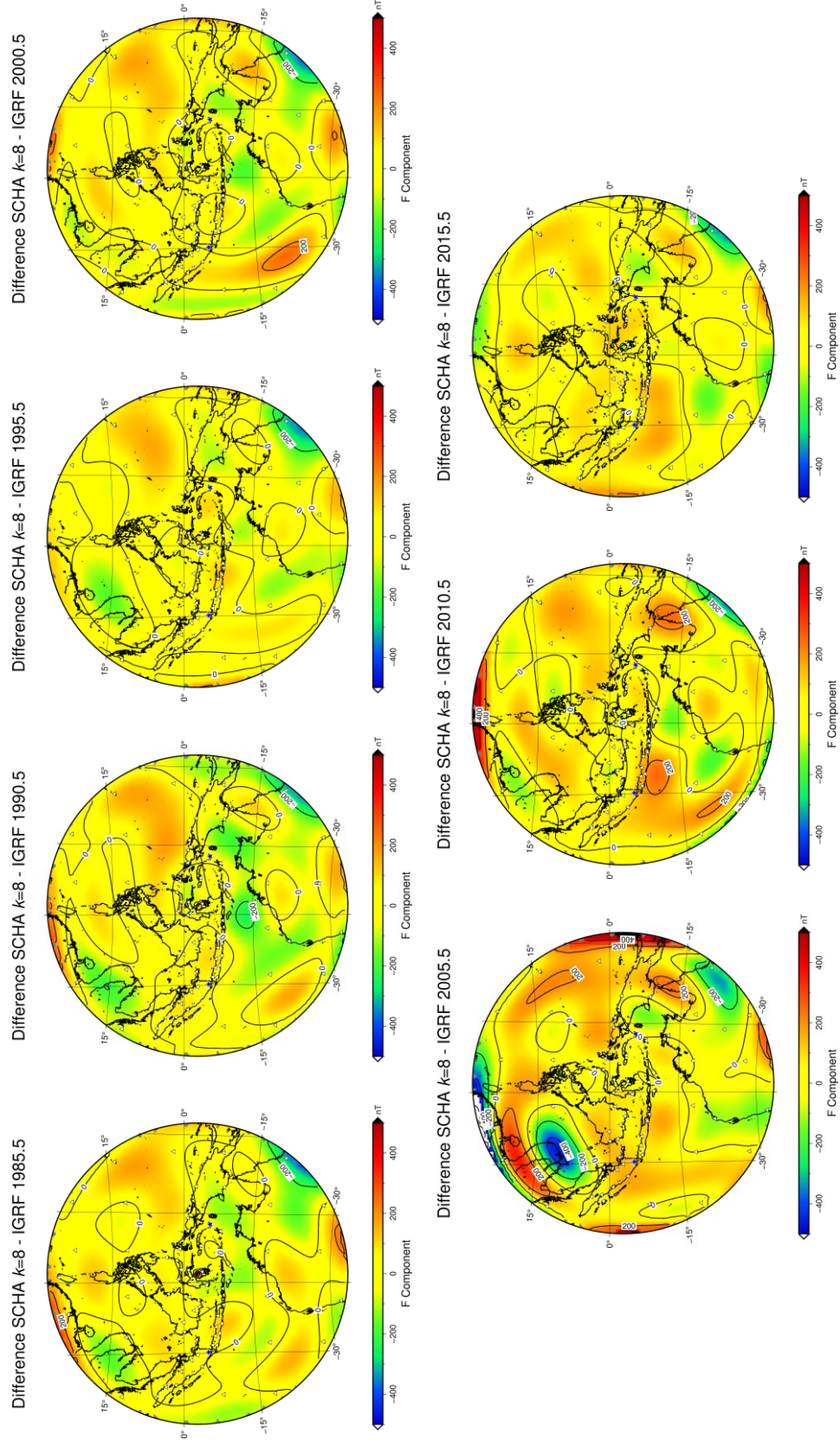


Figure 37. Difference contour at F component between the SCH A using index $k=8$ and IGRF from 1985.5-2015.5.

4.1.4.3 Difference between SCHA Index $k = 9$ and the IGRF Model

The contour map of the difference between SCHA with index $k = 9$ and the IGRF model is shown in Figure 38 – 41. Overall, the difference value is more significant compared to the previous models with lower index k . For all the components, the extreme value areas appear more frequently for all epochs and the boundary problem becomes obvious.

On the X component, in Figure 38, the extreme values area around Vietnam in 2005.5 is still visible. The same feature also appears in other areas in the other epochs. Most of these features are located near the boundary and in the area where there are no observation data available. However, by looking only at the immediate Indonesian region, from 2005.5 and onwards, the difference value is dominated by a positive value.

For the Y component, in Figure 39, the extreme values on the eastern and western boundaries appear in all epochs. However, in 1985.5, 1995.5, and 2015.5, this area is minimised. Extremely negative values still appear around Thailand but are now followed by an area of extremely positive value on its eastern side.

The difference contour map on the Z and F components continues to show more variation. Extreme values approach Indonesia at the centre of the research area. On the Z component, the area of extremely negative values around Vietnam has grown and intersected Indonesia in 2005.5.

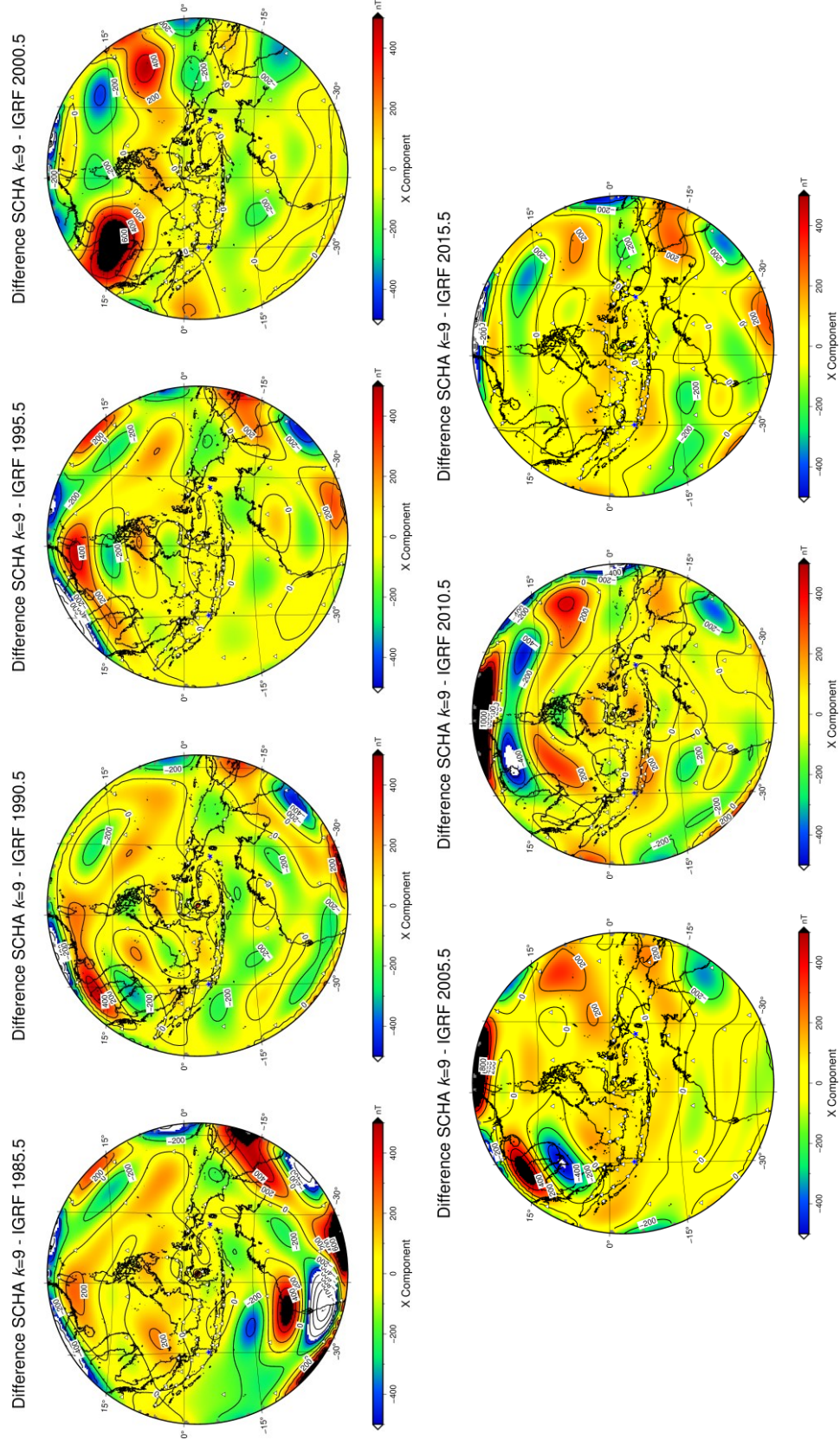


Figure 38. Difference contour at X component between the SCHA using index $k=9$ and IGRF from 1985.5-2015.5.

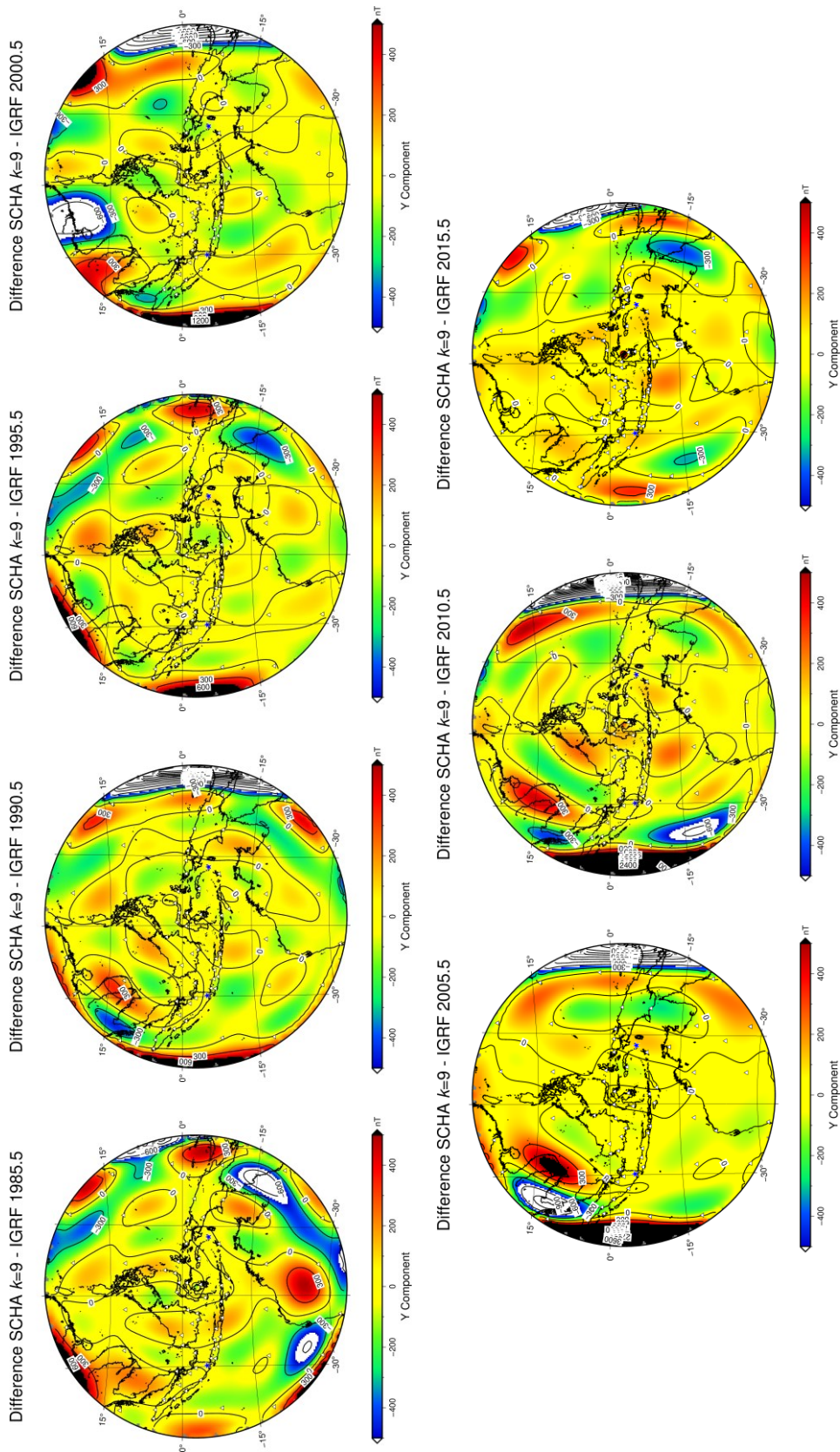


Figure 39. Difference contour at Y component between the SCH A using index $k=9$ and IGRF from 1985.5-2015.5.

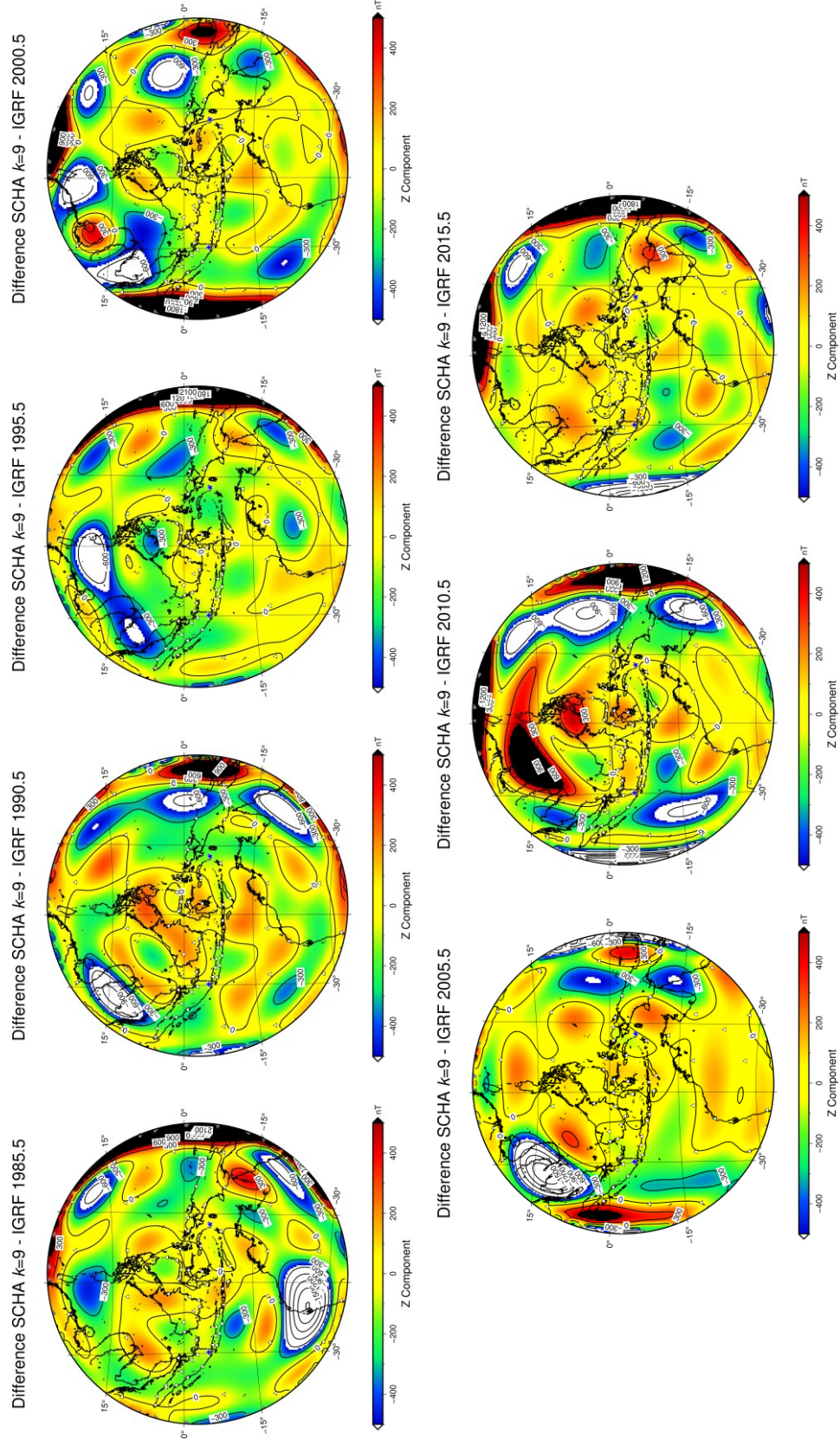


Figure 40. Difference contour at Z component between the SCH A using index $k=9$ and IGRF from 1985.5-2015.5.

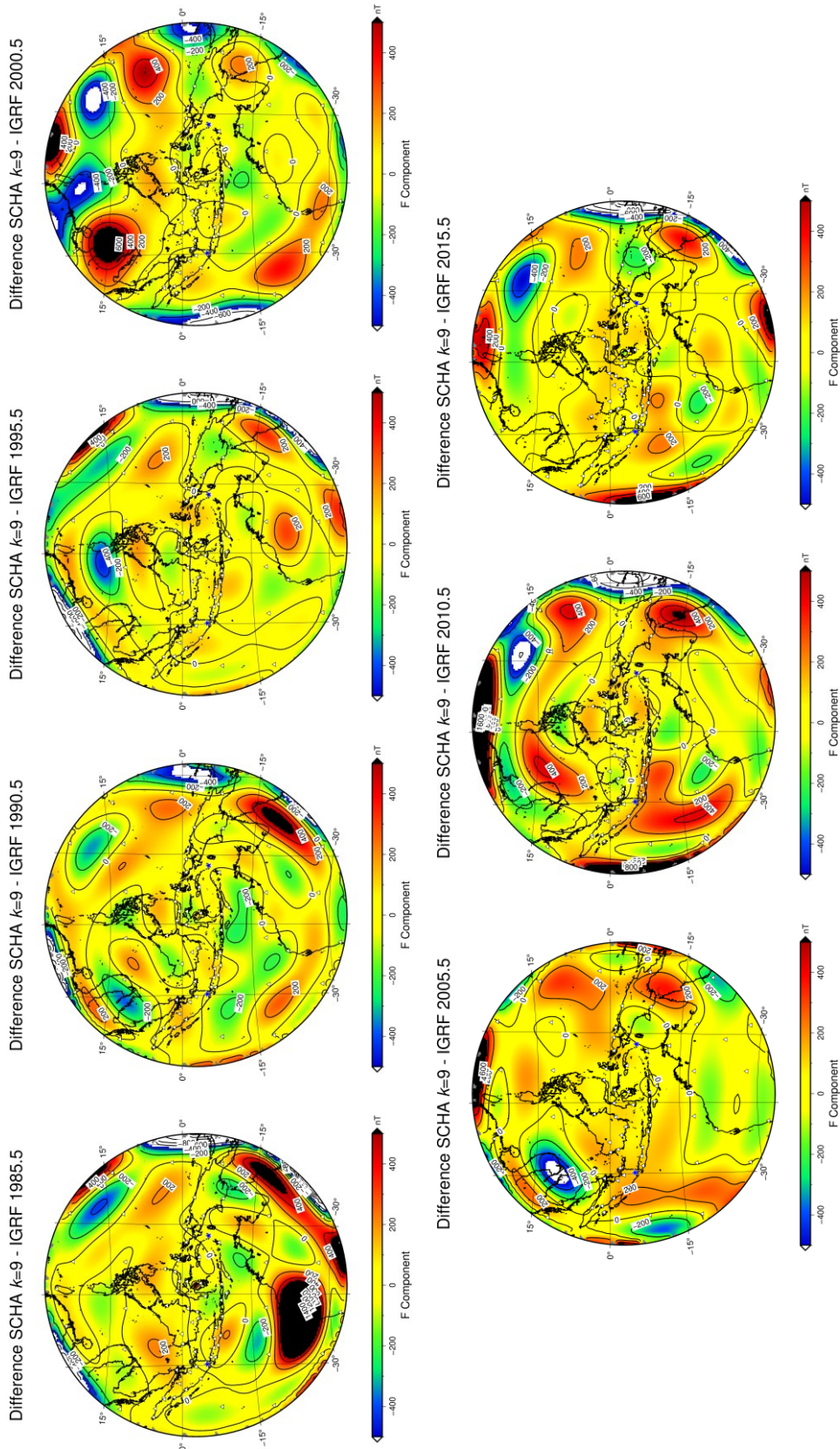


Figure 41. Difference contour at F component between the SCH A using index $k=9$ and IGRF from 1985.5-2015.5.

4.1.5 Selection of the Truncation Level

By analysing the RMSD of the model, checking the model stability, validating the model, and checking the difference between the SCHA and the IGRF model, the best model of the SCHA can be determined. The SCHA model is considered the best if its RMSD is lower than the RMSD of the IGRF model, the contour map is stable, the difference between the model and the real observed value is minimised, the difference between the selected model and the IGRF model is not overly high or low, and the spatial resolution is better than the spatial resolution of the IGRF model.

From Figure 9, the RMSD of the SCHA model can be seen to improve as the index k increases. Overall, at index $k = 4$, the RMSD of the SCHA model is better than the RMSD of the IGRF model. Besides, the SCHA with index $k = 4$ is comparable with the IGRF in terms of the minimum wavelength which can be represented. With that spatial truncation index, the SCHA and the IGRF model represent a minimum wavelength of around 3000 km. In conclusion, models with index $k = 4$ and higher can potentially be selected as the best SCHA model.

However, although the RMSD is small, this does not guarantee that the model is the best, and checking the contour map is essential. This is because the RMSD is only calculated at the observational data locations. At index $k = 9$, the contour map of the SCHA model starts to present a new structure that does not exist on the contour map of the IGRF model. As there is no observational data available at the location of this new structure, it is unconstrained by actual data. However, although the contour map is slightly distorted at index $k = 7$ and 8, there is no new structure formed in Indonesia. In conclusion, by comparing with the contour map of the IGRF model, the SCHA with index $k < 9$ suggests promising results.

Validation of the model has already been discussed, and it is recognised that the SCHA model is not always better than the IGRF model at fitting data from individual locations, for example, at Bandar Lampung. However, the validation of the SCHA model at Tual suggests, that in general, the model gives promising results. From Figure 29, the SCHA with index $k = 7$ seems the model which gives a better result than the IGRF model, although the minimum value position is inconsistent for each epoch.

Another aim to determine the best SCHA model is by calculating its differences with the IGRF model. From the previous section, the SCHA with index $k = 9$ shows a high number of areas with extreme difference values and some large differences on

the boundary of the spherical cap. For the SCHA with index $k = 8$, the number of extreme difference values is reduced significantly, but the boundary problem is still obviously visible. Both the extreme difference value area and the boundary problem are minimised for the SCHA with index $k = 7$. In summary, as the index k increases, the boundary problem and the extreme difference values become more prominent.

Overall, from the analysis already performed, the SCHA model with spatial truncation index $k \geq 4$ and $k \leq 7$ is capable of being chosen as the best model. Nevertheless, considering its minimum wavelength representation, the SCHA model with index $k = 7$ is considered as the best SCHA model. The SCHA model with index $k = 7$ can represent a minimum wavelength of 1821 km. The same wavelength would need the IGRF of degree 22. By looking at Table 5, it can be seen that from 1985.5 – 2015.5, the average RMSD of the SCHA with index $k = 7$ on the X, Y, and Z components are approximately 117 nT, 86 nT, and 134 nT, respectively. On the other hand, the average RMSD of the IGRF on the X, Y, and Z components are approximately 145 nT, 97 nT, and 168 nT, respectively, as can be seen in Table 4. Thus, compared to the RMSD of the IGRF model, the RMSD of the SCHA model with index $k = 7$ is reduced by 28 nT, 11 nT, and 34 nT for X, Y, and Z components, respectively. This suggests that the SCHA method is far more efficient in geomagnetic modelling than the SHA method regarding its accuracy and minimum wavelength property.

4.1.6 Misfit Analysis between Observed and Predicted Data

After finding the best SCHA model, it is still necessary to check the misfit between the observed data and the selected SCHA model. The misfit is calculated by subtracting the SCHA model from the real observation data at all the observation data locations. As Indonesia is the main focus of this research, the map is only plotted in the Indonesian region. The results are given in Figure 42 – 44 for the X, Y, and Z components, respectively.

By looking at Figure 42, the misfit for the X component varies for all epochs. Nevertheless, most of the locations give a small misfit, shown by the colour green. However, some of the locations show larger positive or negative misfit values. Such a location, e.g., Jayapura at the eastern side of Indonesia on Papua Island, shows a positive misfit for all the epochs. The local geomagnetic field causes a misfit of this kind if its wavelength is smaller than the minimum wavelength representation by the SCHA model. Another location, such as Jambi, on the northern side of Bandar

Lampung on Sumatera Island, shows a very negative misfit in 1995.5. However, the misfit in the other epoch is not so large. Inaccuracy in the data itself might cause this kind of misfit.

On the Y component in Figure 43, the misfit varies for all epochs. The number of locations with very positive or negative misfit values is much lower than for the X component.

For the Z component, it is evident that the misfit at the locations around Bandar Lampung is much more significant than the misfit around Tual. The misfit at Bandar Lampung itself is not so large, apart from in 2005.5. However, the data around Bandar Lampung vary between positive and negative misfits over a short distance. This suggests that the misfit of the SCHA model around Bandar Lampung is caused by a combination of the inaccuracy of the observation data itself and the small-scale local geomagnetic field, which cannot be modelled by the SCHA model. This also explains why the RMSD of the SCHA models at Bandar Lampung is much higher than at Tual, as discussed in the model validation section.

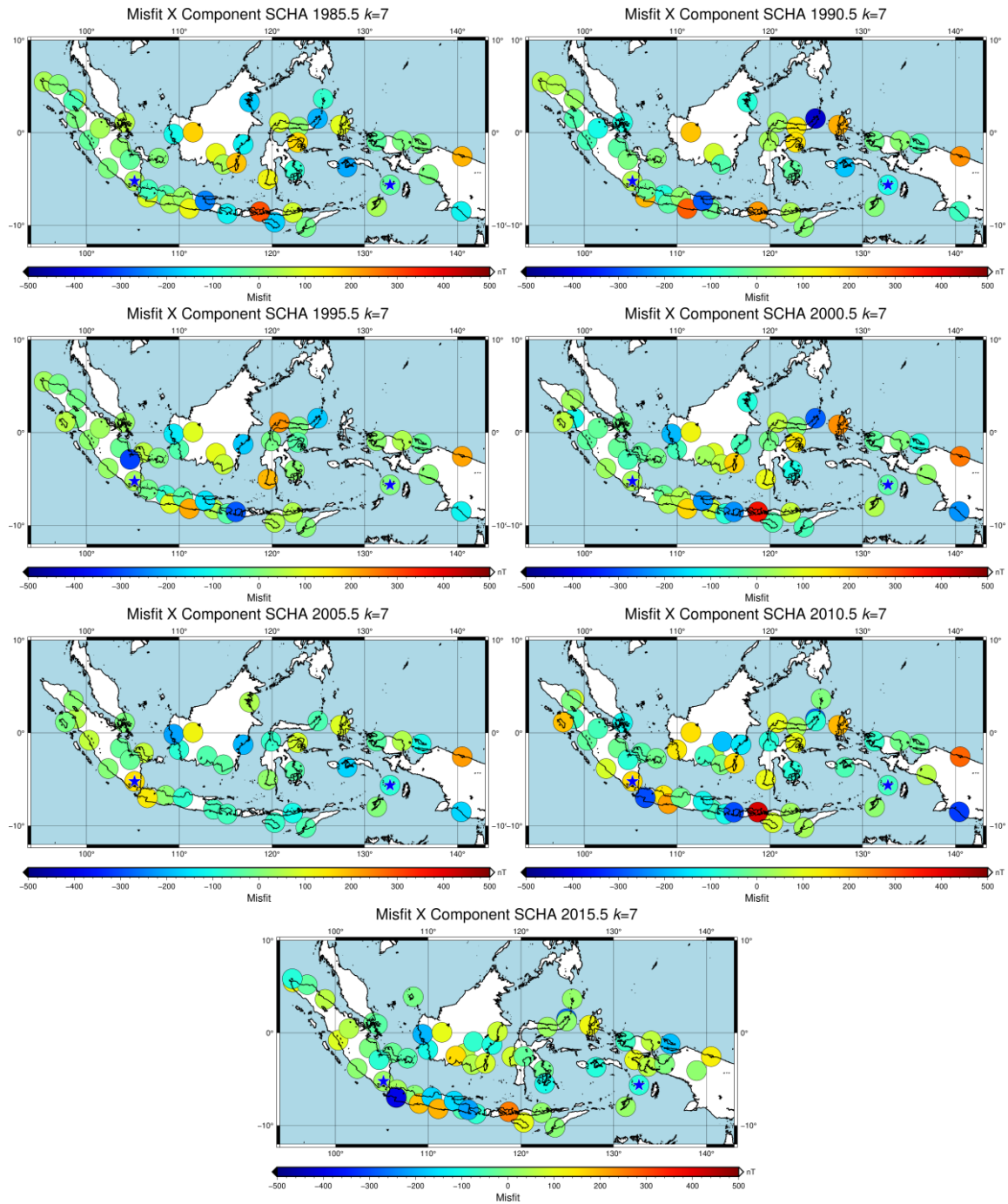


Figure 42. Misfit map at X component between the SCHA using index $k=7$ and observed data from 1985.5-2015.5. The misfit is calculated by subtracting the SCHA model from the real observation data at all the observation data locations. The good misfit is indicated by the value approaching zero or showed as the light green circle.

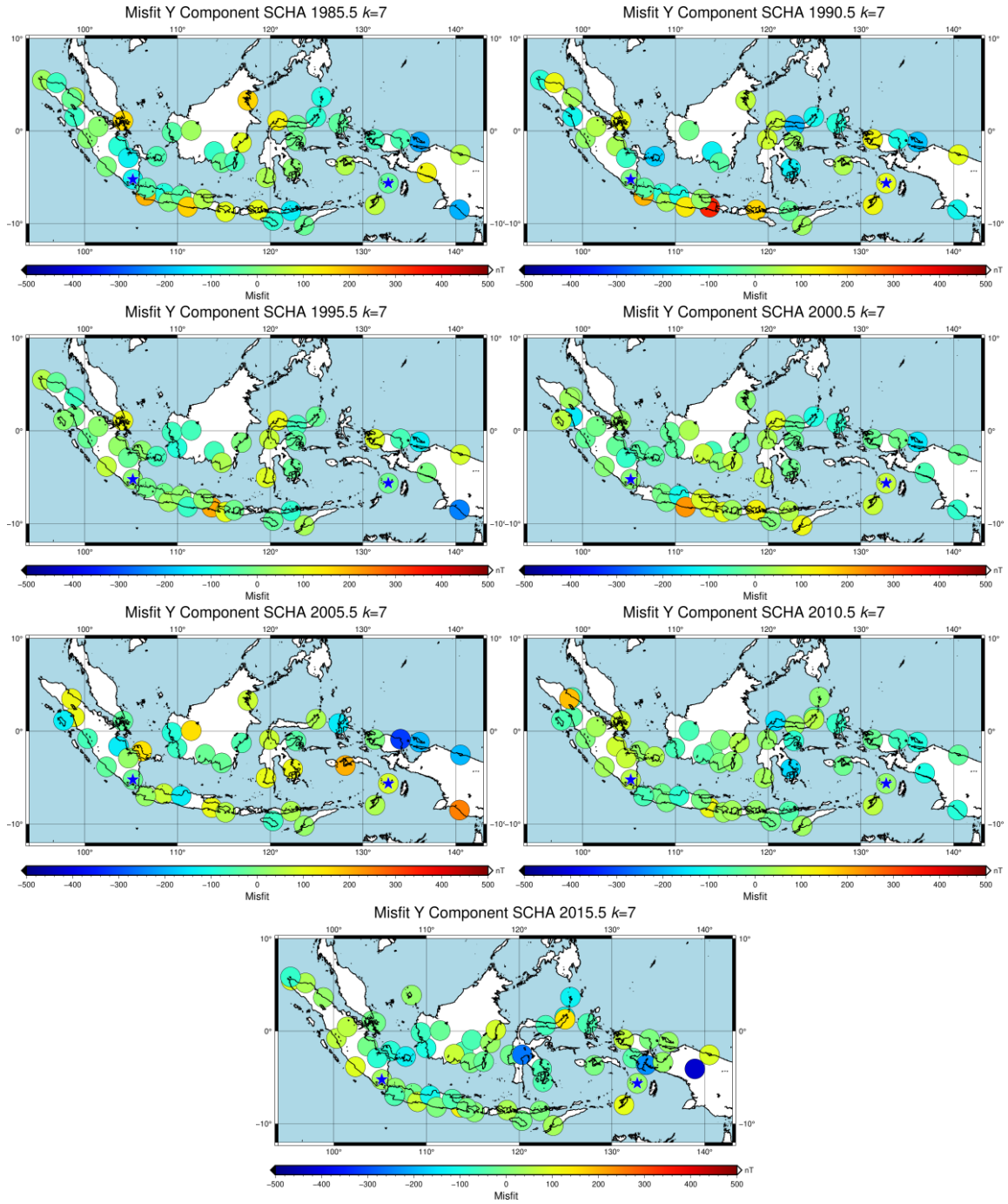


Figure 43. Misfit map at Y component between the SCHA using index $k=7$ and observed data from 1985.5-2015.5. The misfit is calculated by subtracting the SCHA model from the real observation data at all the observation data locations. The good misfit is indicated by the value approaching zero or showed as the light green circle.

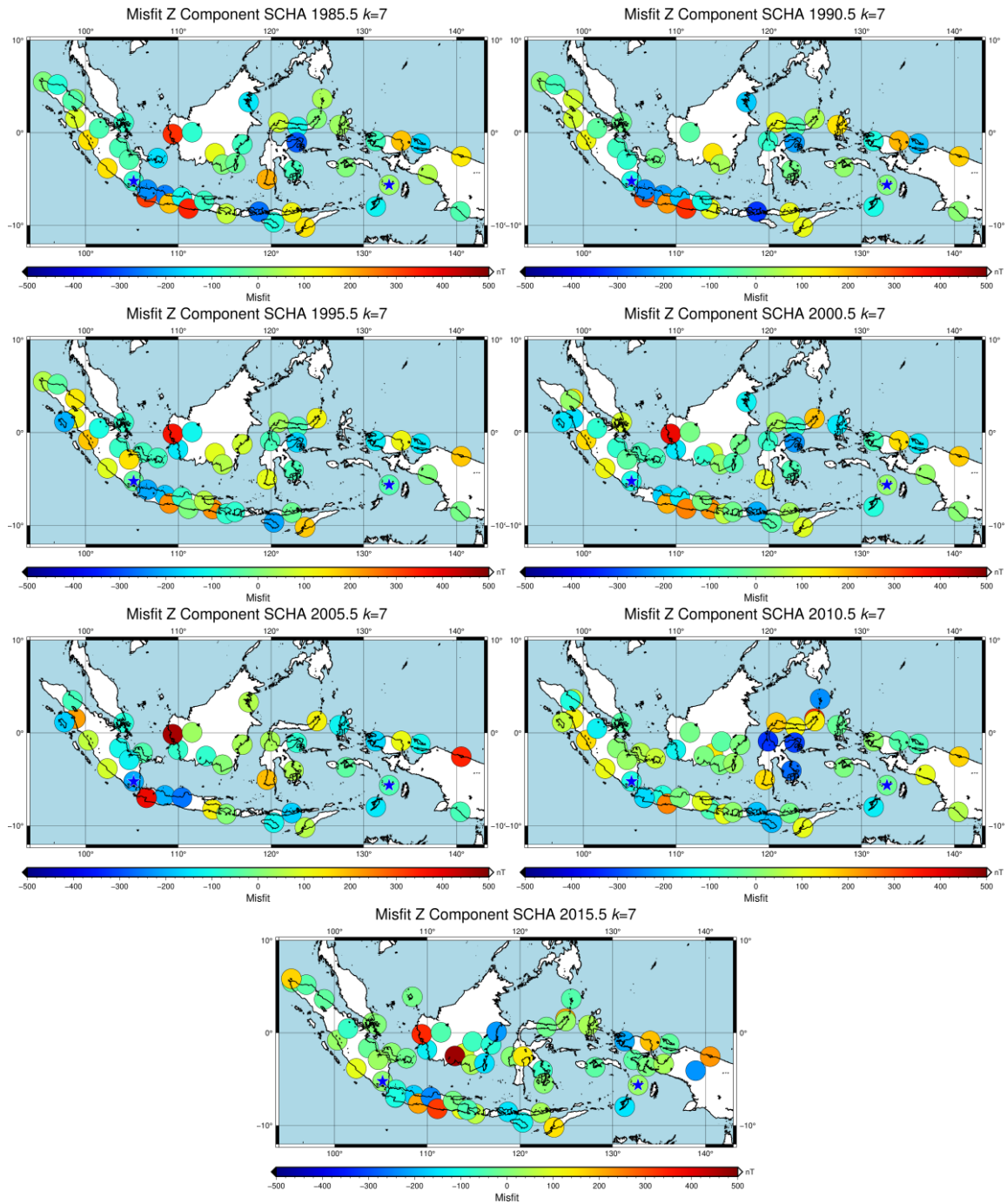


Figure 44. Misfit map at Z component between the SCHA using index $k=7$ and observed data from 1985.5-2015.5. The misfit is calculated by subtracting the SCHA model from the real observation data at all the observation data locations. The good misfit is indicated by the value approaching zero or showed as the light green circle.

4.2 Temporal Modelling Analysis

After analysing the spatial model and selecting the best spatial truncation index, temporal modelling and secular variation forecasting needed to be performed. Temporal modelling is needed for estimating the geomagnetic field in the years when there is no observational data available.

The problem with temporal modelling is how to fit a smooth function to time-series observation data. Hulot et al. (2015) mentioned that for long period data, longer than five years for instance, a spline representation is preferred. In this chapter, the spline representation is used to calculate annual field models. However, it also can be used to calculate a shorter time separation model by changing the knot points. After obtaining the annual geomagnetic field model, the annual rate of changes can be calculated. Later, based on this annual field model, the secular variation can be forecasted for the next five years using an autoregressive forecasting method.

4.2.1 Temporal Modelling using B-splines

To show the performance of the regional model over the period spanned by the data, the SCHA models at five-year intervals need to be interpolated. One of the common methods in time-dependent field modelling is using B-splines as the temporal basis functions. The B-splines used in this calculation is the fourth-order (or cubic) B-splines as the basis functions for the expansion of the Gauss coefficients, i.e. g_k^m and h_k^m . The interpolated Gauss coefficients are used to calculate the SCHA model on an annual basis. A similar method was used by Verbanac (2007) when he developed the main geomagnetic field in Europe using synthetic data. The annual mean graphs at Tuntungan (TUN), Bandar Lampung, and Tual are shown in Figure 45. Bandar Lampung and Tual are chosen to validate the model because the data from both locations are not used to calculate the SCHA model, while Tuntungan is chosen to validate the data at a geomagnetic observatory.

Overall, there is good agreement between the SCHA and IGRF models, and both models represent the geomagnetic field quite appropriately. The trend is smoothly represented, and there are no sudden increments or fast decreases in either model. The annual mean values and the trend seen in both models are quite similar. However, the annual mean values of the SCHA model provide better results than the annual mean of the IGRF model at Tuntungan (TUN), particularly for the Y component. Hence, although the TUN data is used in the IGRF-11 and IGRF-10 calculations, it seems that the IGRF is not as well modelled at Tuntungan as it is by the SCHA.

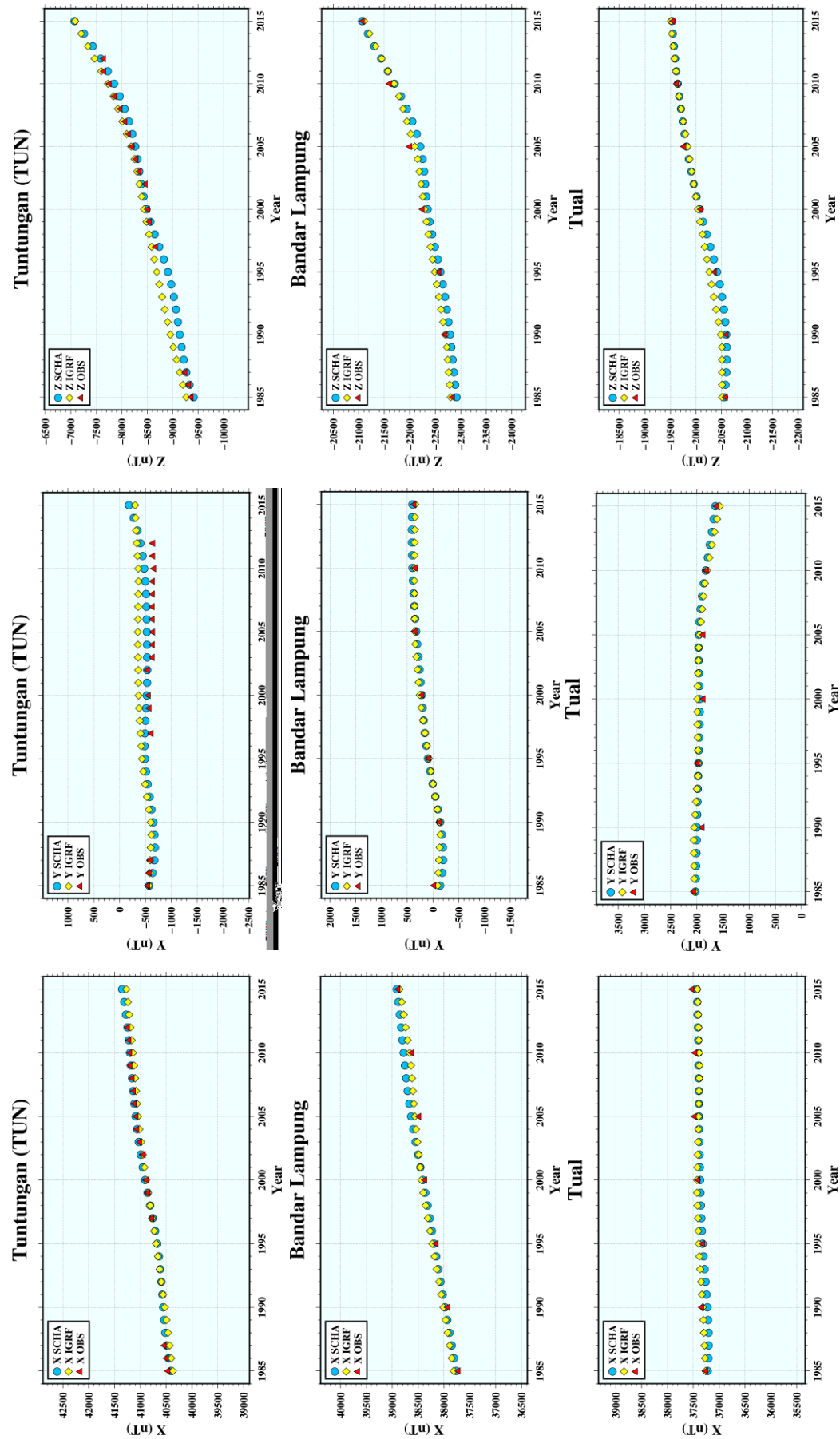


Figure 45. Annual mean values at Tuntungan, Bandar Lampung and Tual.

The SCHA model at Bandar Lampung and Tual are well behaved, and the trend of the annual mean values is well modelled. Nevertheless, surprisingly, the IGRF model also represents the geomagnetic field appropriately at these locations, which are repeat stations and not used to calculate the IGRF.

A comparison of the annual rate of changes of the SCHA and IGRF is shown in Figure 46 to illustrate the fidelity of the models in representing the observation data. The annual rate of changes is calculated by subtracting the current annual means values by their previous annual means, i.e., the first differences of the annual means. Again, the illustrations are only shown for Tuntungan (TUN), Bandar Lampung, and Tual as examples. However, Bandar Lampung and Tual are repeat stations, thus the observation data from these locations are only available every 5 years. As a result, the annual rate of changes from observation data on these locations is not shown in Figure 46.

Despite the limited geographical extension of the research area, the regional model follows the temporal variation of the different components better than the IGRF does. The IGRF has constant secular variation over the given five-year intervals, while the SCHA model follows the secular variation accurately on an annual basis and shows a smooth trend. This shows that the secular variation of very small portions of the Earth is not captured by the global model (Pegini et al., 2018). Similar results were obtained by De Santis et al. (2003) when they modelled the annual means and predicted the secular variation of the Italian Geomagnetic Reference Field (ITGRF).

The interesting fact in Figure 46 is the rate of annual changes in Tuntungan (TUN), specifically at the end of the Y and Z components. Because there are no data in 2015.5 used in the creation of the SCHA model, there is a mismatch at the end of the data series, compared with the rate of changes of the IGRF model. The rate of changes of the SCHA model increases continuously, while the IGRF shows a lower rate of changes and a small decrease at the end of the data series. However, there are no data available for the last several years of the data series to determine which model is considered better. This shows the difficulty in regional geomagnetic modelling when there are not sufficient data available.

On the other hand, the annual means at Bandar Lampung and Tual are well represented by the SCHA and IGRF model, and the annual rate of changes are almost similar for both models in all components. The data series at both locations do not contribute to the SCHA model calculation; instead are used as an external quality check. In this case, it suggests that the data series from the locations around

Bandar Lampung and Tual influence and control the SCHA model for both locations such that they approach the observational data.

Finally, based on both annual means and the annual rate of changes graphs, the B-splines basis functions provide reliable results in the time-dependent field modelling. The results are smooth and generally match with the observations data. Nevertheless, a complete and evenly distributed data set will provide a significant benefit in temporal modelling.

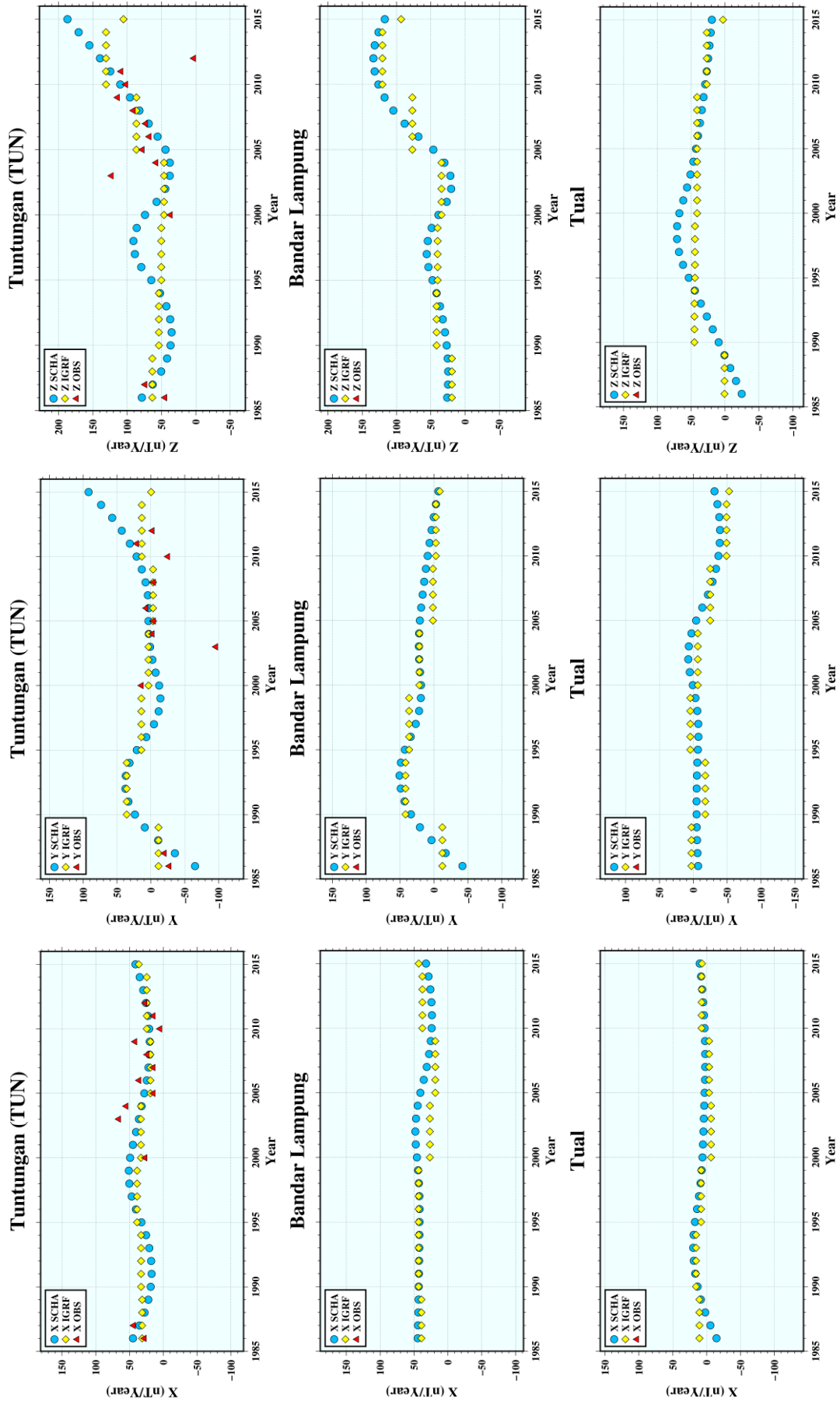


Figure 46. Annual rate of change at Tuntungan, Bandar Lampung and Tual.

4.2.2 Secular Variation Forecasting using Autoregressive Forecasting

The last goal of this research is to extend the secular variation model to 2020. The interpolated data using B-splines in the previous section is used as the input for the extrapolation of the SCHA model using the autoregressive (AR) forecasting method. De Santis et al. (2003) used a similar method when developing the Italian geomagnetic reference field and obtained satisfying results.

The AR method is a time series model that uses observations from previous time steps as input to a regression equation to predict the value at the next time step. It is generally restricted to stationary data, meaning the property of the mean, variance, and autocorrelation structure do not change over time (NIST/SEMATECH, 2012). Thus, time series with trends or with seasonal changes are not stationary. By looking at Figure 45, it is clear that the annual mean time series is not stationary as there is an increasing or decreasing trend. One way to make the data series stationary is by computing the differences between consecutive observations data, also known as differencing. By removing changes in the level of a time series, differencing can stabilise the mean of time series and eliminate or reduce the trend.

Subsequently, after differencing the data, the order (p) of the AR is determined by looking at its partial autocorrelation graph. The order of the AR shows the number of prior values (lag) of the series that will be regressed against the current value. The partial autocorrelation of an $AR(p)$ process is zero at lag $p + 1$ and greater. Therefore, the order (p) of the AR is identified by looking for the point where the partial autocorrelation becomes zero. Placing a 95% confidence interval for statistical significance is helpful for this purpose. The Statsmodel (Seabold & Perktold, 2010), a library for statistical and econometric analysis in Python, is used to make the partial autocorrelation plots for the Y component, which are shown in Figure 47.

The partial autocorrelation plot shows clear statistical significance for lags 1, 2, and 3. The next few lags are at the borderline of statistical significance and considered insignificant. Similar plots are also seen for the other components in all locations. Thus, the partial autocorrelation plot indicates that an AR with order $p = 3$ is the correct model. The lag at 0 is not shown because the partial autocorrelation coefficient is always 1.

After the order of the AR model is identified, the forecasting calculations can be accomplished and are limited to the next five years from the last data series to

maintain their accuracy. The differenced time series is used as the input data and are transformed back to the original values after the AR(3) calculations. The forecasting calculation is performed at all the data locations. However, the results of the AR(3) calculations, shown in Figure 47, only show the results at Tuntungan, Bandar Lampung, and Tual.

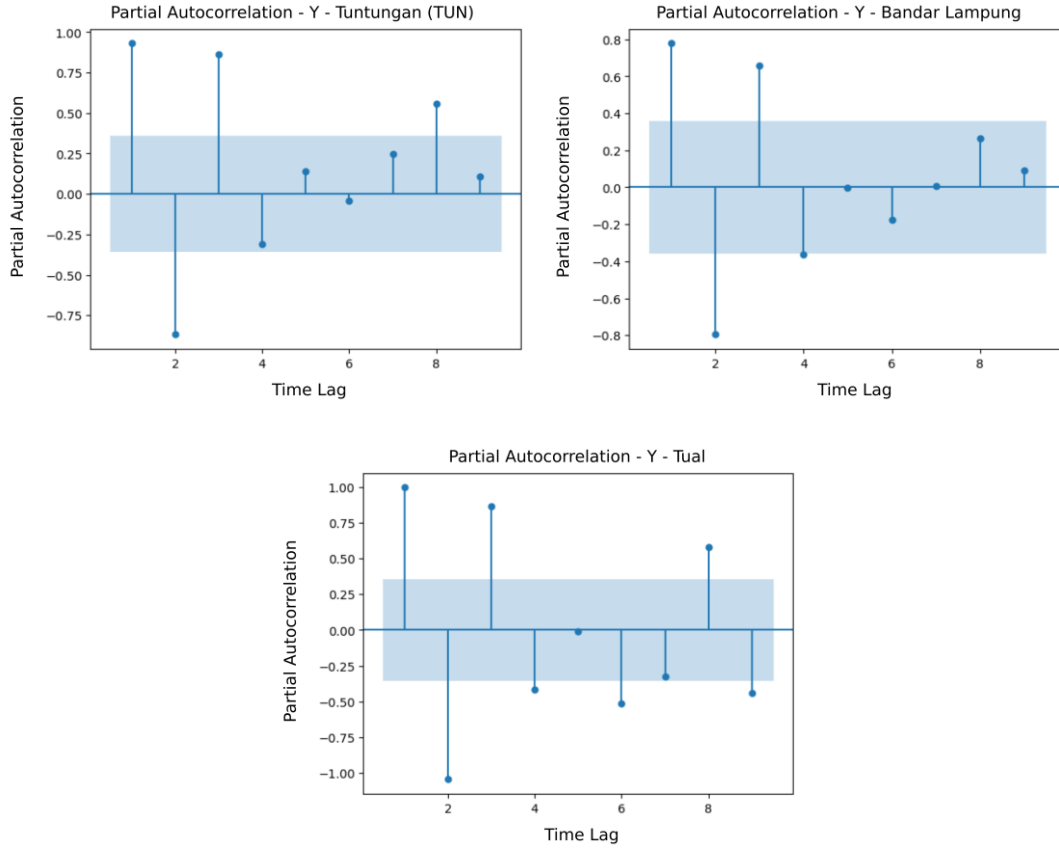


Figure 47. Partial autocorrelation plot at Tuntungan, Bandar Lampung, and Tual. The 95% confidence interval bands are shown as the blue area.

In Figure 48, the annual means from 1985.5 – 2015.5, as shown in Figure 45, are combined with the forecasted values from 2016.5 – 2020.5. The results of the AR forecasting model appear quite adequate. Values are mostly similar to the observational data and the prediction from the IGRF model (which of necessity give a linear trend over the 5 year period). However, for some locations and for some components, the trend of the forecasted SCHA model is different from the IGRF model.

At Tuntungan (TUN), the forecasted SCHA model is adequate for the X component. However, despite the better results for the annual mean calculations from 1985.5 – 2015.5, the forecast model does not perform well, specifically for the Y component. It increases sharply compared to the IGRF model, and the observational data are

closer to the IGRF than the forecasted SCHA model. Similar poor performance is visible in the Z component. This may result from the lack of Tuntungan data towards the end of the data span, especially at 2015.5.

Nevertheless, at Bandar Lampung and Tual, the SCHA model is well behaved and gives a similar trend to the IGRF model, although there is a slightly different trend at Tual in the Y component. The observational data in 2015.5 are available at the data locations around both locations. They confirm that the forecasting method will perform better when there are data available at the end of the data-spanning period.

The RMSD from both the SCHA and the IGRF model are calculated to validate the forecasting method. However, from 2016.5 – 2020.5, the observational data is minimal, and it is not well distributed both spatially and temporally. Nonetheless, from the 41 observational data from all sites, the RMSD of the forecasted SCHA model is 154.92 nT, 200.87 nT, and 104.39 nT for X, Y, and Z components, respectively, while the RMSD of the IGRF model is 172.62 nT, 95.52 nT, and 117.55 nT for X, Y, and Z components respectively. The RMSD of the IGRF is calculated for similar years and locations as the RMSD of the SCHA. It is noticeable that the SCHA model gives a smaller RMSD than the IGRF for the X and Z components but a larger one for the Y component. However, for the total field (F) component, the RMSD of the SCHA is 135.81 nT, and it is better than the RMSD of the IGRF, which is 162.38 nT.

The intention of the regional geomagnetic modelling is mainly for navigation, which uses the declination, and for the reduction of survey results to enable the production of the anomaly maps, which uses the total field component. The declination component is determined by the vector content of the X and Y components from the relationship $D = \tan^{-1}(Y/X)$, while the total field component is obtained from the relationship $F = \sqrt{X^2 + Y^2 + Z^2}$. Therefore, from the results of the RMSD, the forecasted SCHA model in this research is less suitable for navigation use, but it is more appropriate for data reduction than the IGRF model.

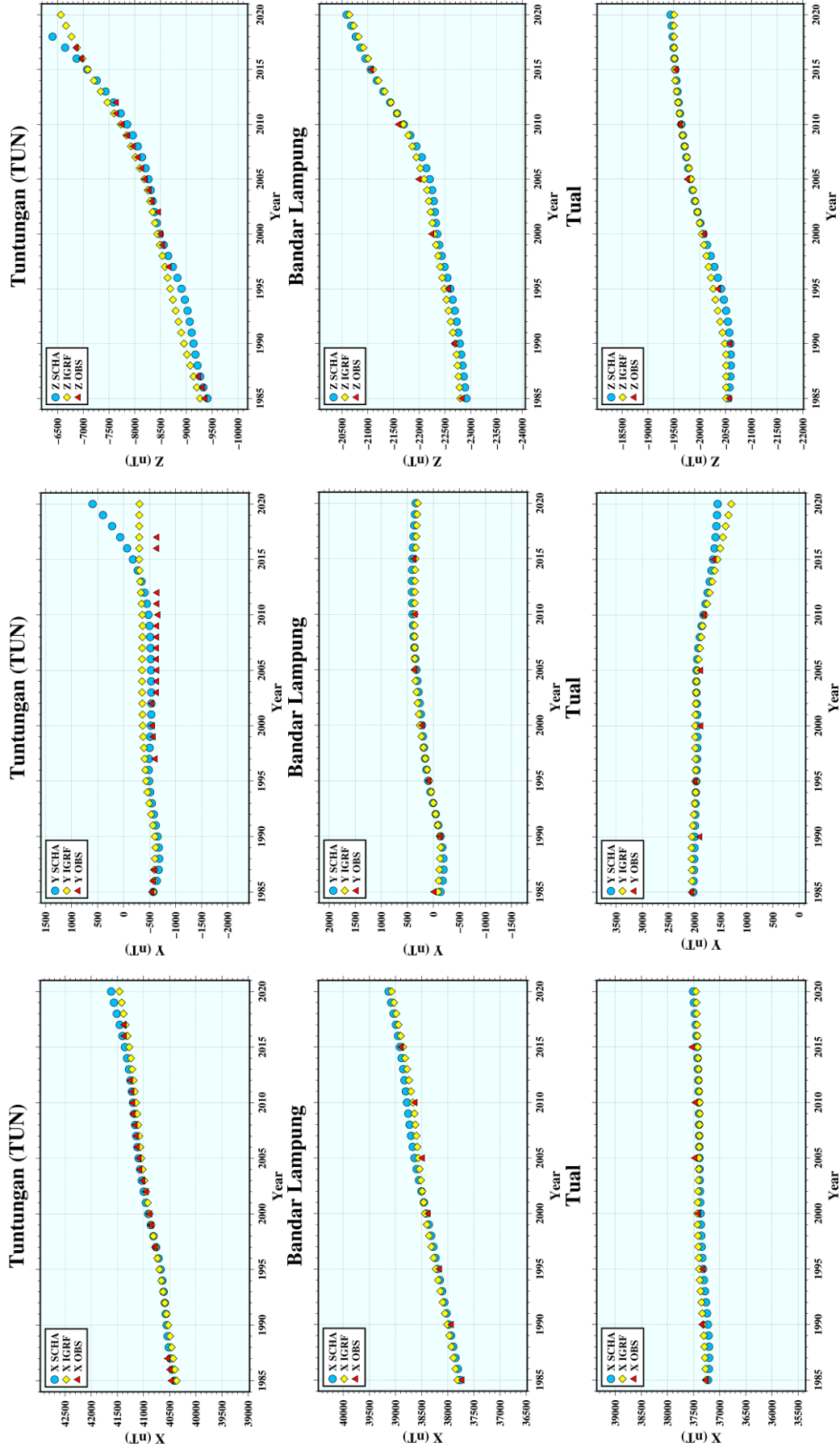


Figure 48. Annual mean values at Tuntungan, Bandar Lampung, and Tual together with the values predicted by the SCHA and the IGRF.

In the same way, a synthetic grid of the SCHA model from 1985.5 to 2015.5 at annual intervals was constructed using B-splines basis functions and used as the input data in the AR(3) forecasting. This grid covers the whole research area at a 1° interval. The contour maps of the AR(3) forecasting calculations are shown in Figure 49 through to Figure 52.

The behaviour of the AR(3) forecasting method is different for each component. However, overall, the contour map is slightly distorted as time moves forward, and likewise, the accuracy of the forecasted SCHA model deteriorates. The contour map of the X component in Figure 49 shows smooth contours. Distortions are most significant in the Y component and become stronger as time moves forward. This agrees with the results of the AR(3) forecasting in Figure 48, where the forecasted Y component deviates from the forecasted IGRF model. Moreover, the RMSD of the forecasted SCHA model for the Y component is larger compared to the RMSD of the IGRF model. On the other hand, the forecasted contour maps of the Z and F components are reasonable, and the distortion is minimal compared to the other components. The F component on this map is the resultant of the X, Y, and Z components.

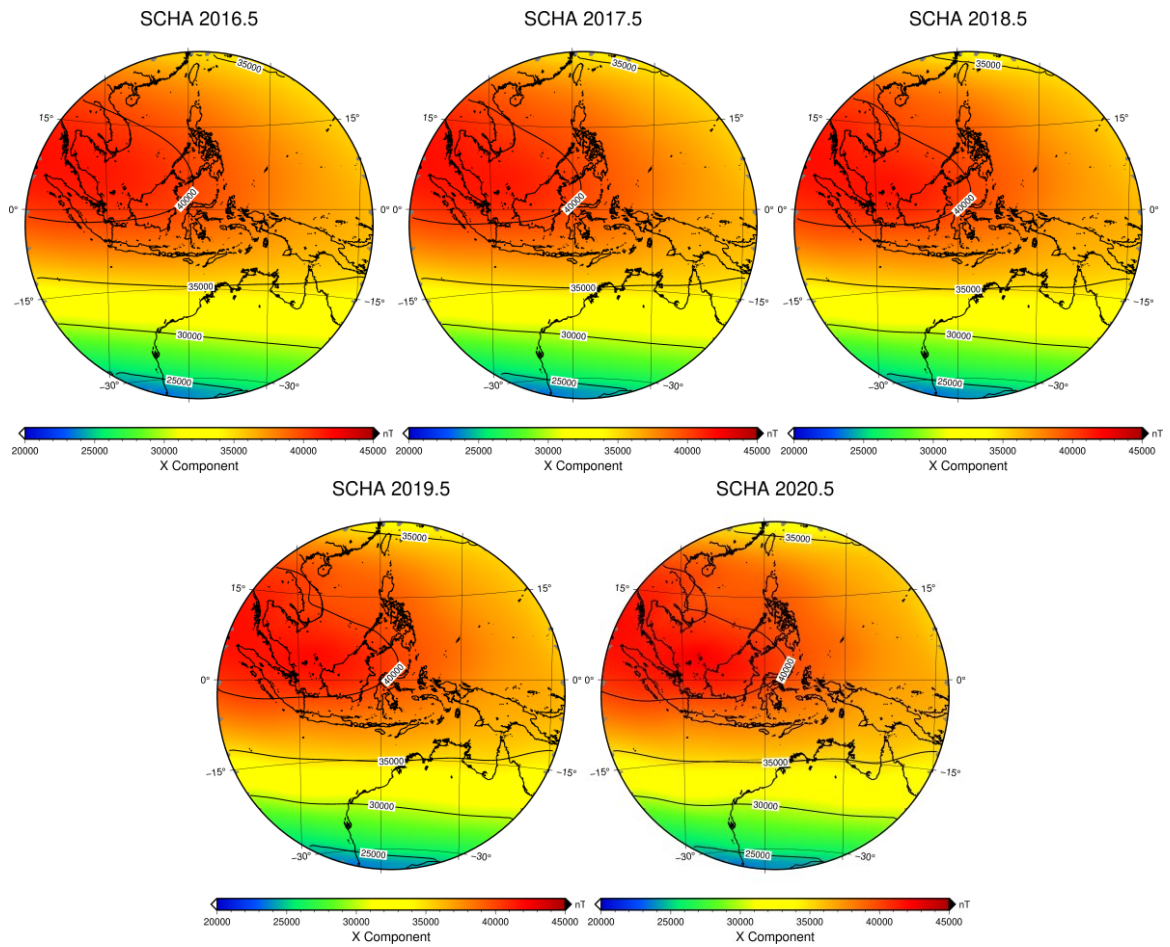


Figure 49. Forecasted X component contour from 2016.5 – 2020.5.

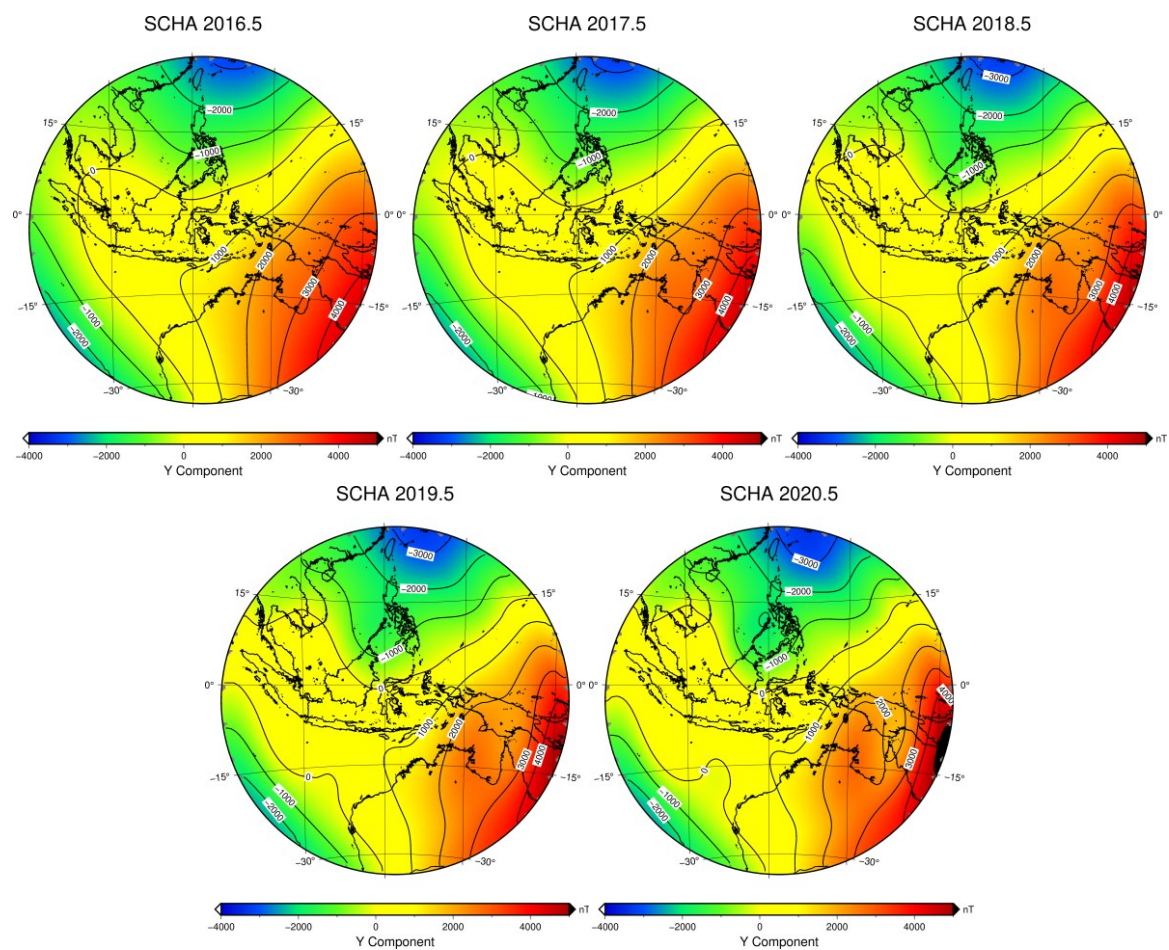


Figure 50. Forecasted Y component contour from 2016.5 – 2020.5.

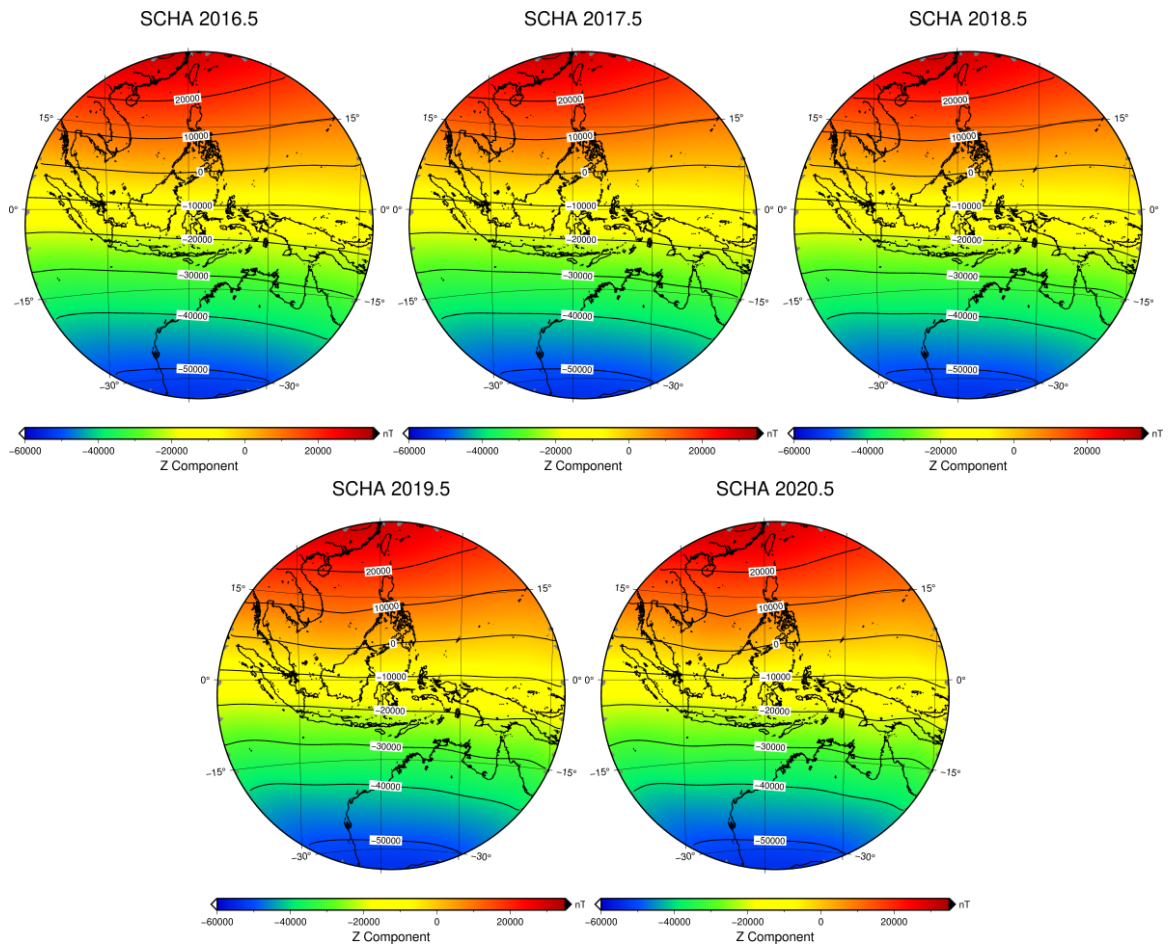


Figure 51. Forecasted Z component contour from 2016.5 – 2020.5.

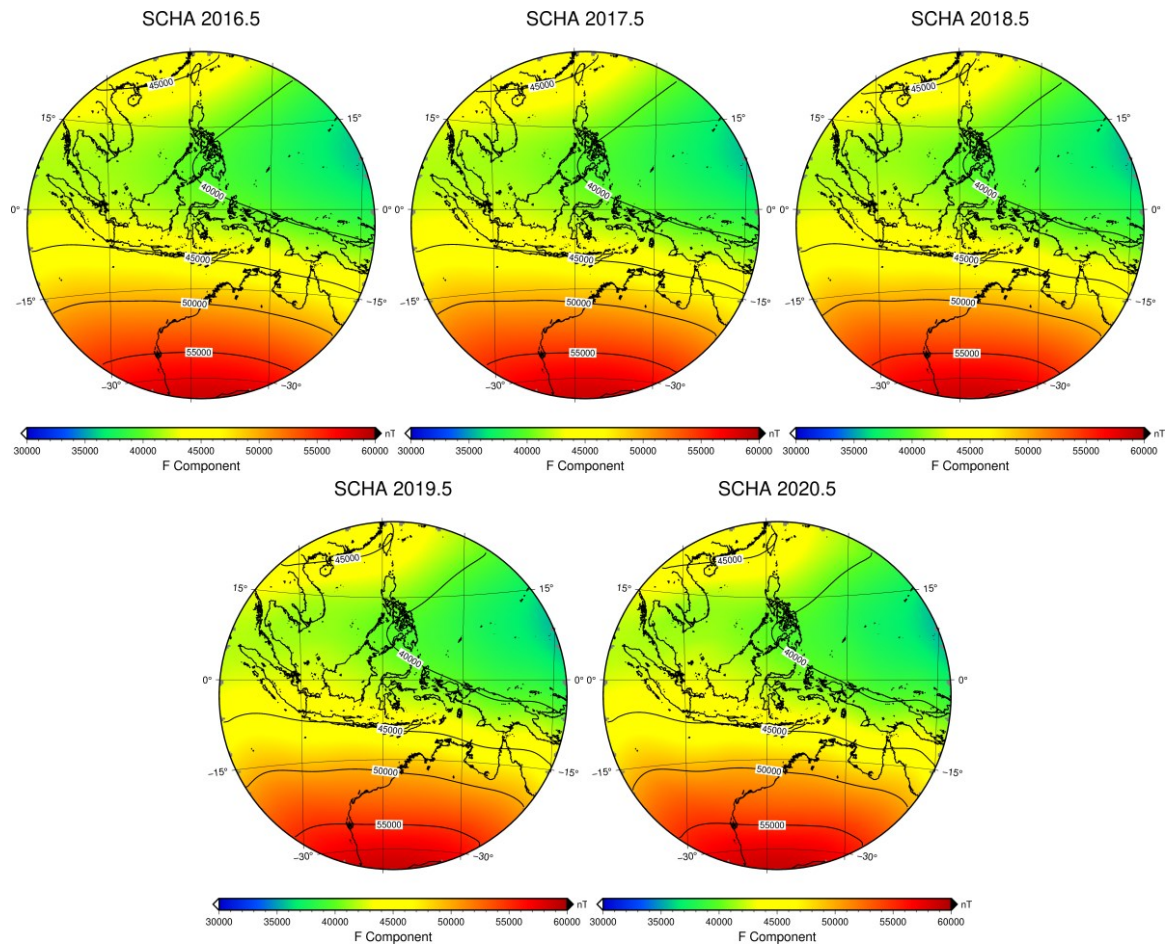


Figure 52. Forecasted F component contour from 2016.5 – 2020.5.

The last thing that needs to be studied is the uncertainty of the forecast model. As in any forecasting calculation, the forecasted model always contains unpredictable elements. To determine this uncertainty, the confidence interval of the SCHA model is calculated in 95% and 80% confidence levels. The results of the confidence interval calculation at Bandar Lampung and Tual are shown in Figure 53. Note that the ordinate of each graph is on a different scale.

Figure 53 shows that the uncertainty is increasing as time moves forward, and likewise, the accuracy of the forecasted SCHA model deteriorates. As an example, in 2020.5, it is 95% confident that the forecasted SCHA model at Bandar Lampung for the X component is between 39113.0 nT and 39152.1 nT, and 80% confident that it is between 39119.7 nT and 39145.3 nT. The widest range in the 2020.5 predictions is in the Z component at Bandar Lampung, where the 95% confidence level is in the range of 192.27 nT and 125.72 nT for the 80% confidence level. This is also related to the SCHA calculation at Bandar Lampung, which is worse than in Tual, as discussed in the validation of the model section.

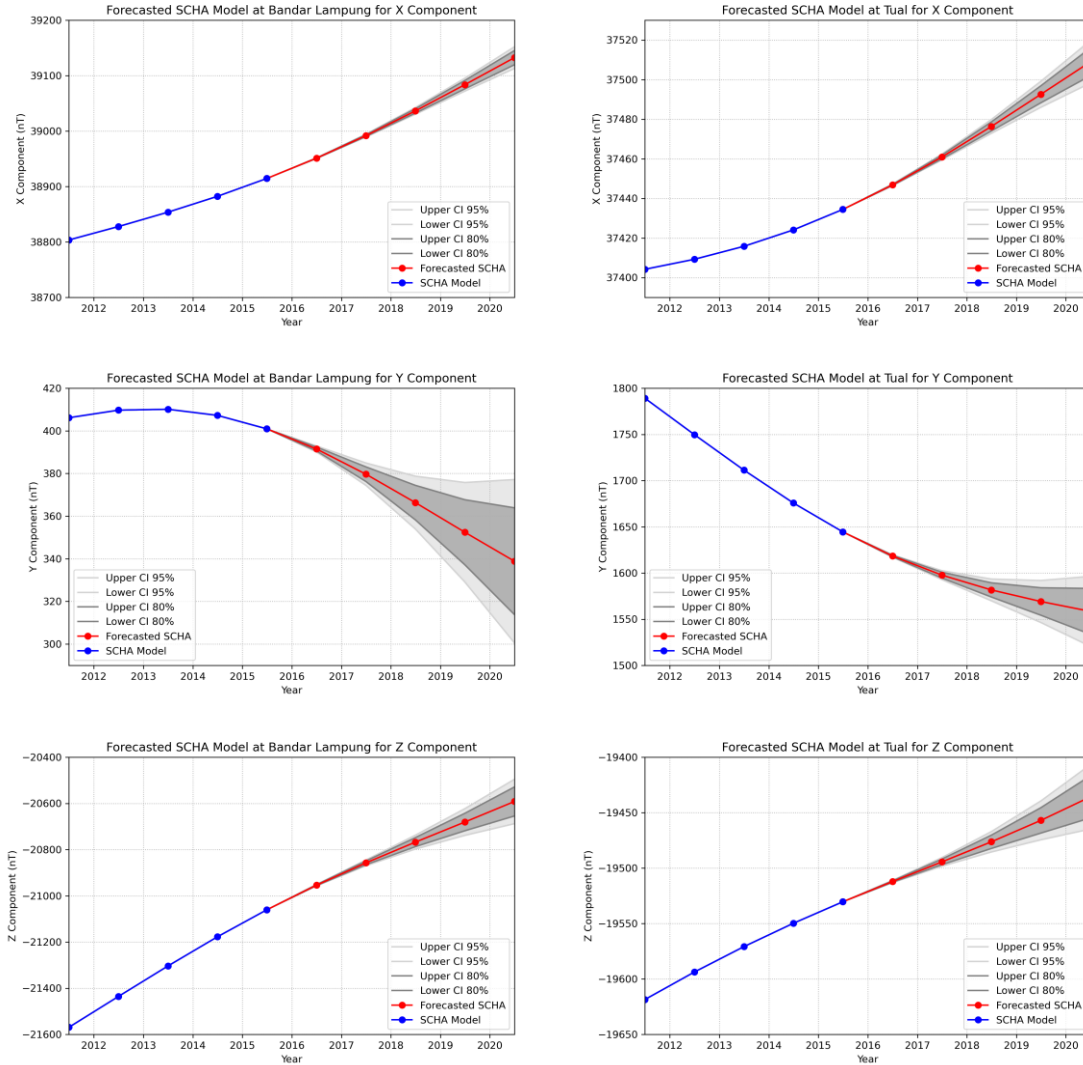


Figure 53. 95% and 80% confidence interval from the forecasted SCHA model at Bandar Lampung and Tual. The light grey shaded area is the 95% confidence level, and the darker grey shaded area is the 80% confidence level. The horizontal axis is bounded between 2011.5 and 2020.5 to focus the charts.

Chapter 5

Summary

5.1 Summary

Ten regional geomagnetic reference fields have been developed using the SCHA technique covering the Indonesian region. The SCHA model uses the spatial truncation index $k = 1 - 10$. The SCHA models have benefited from the availability of a larger amount of observational data than the IGRF. These models have been calculated using observational data from 68 geomagnetic repeat stations in Indonesia, covering the period 1985.5 – 2015.5 from BMKG, definitive data from five BMKG's geomagnetic observatories, combined with the definitive data from 13 INTERMAGNET observatories, and completed with the synthetic data at sea level at 17 fixed points calculated from IGRF-13 from the similar period.

By statistical analysis and comparison with the IGRF model, the SCHA model with index $k = 7$ is considered the best. The average RMSD of the SCHA with index $k = 7$ for the X, Y, and Z components are approximately 117 nT, 86 nT, and 134 nT respectively, while the average RMSD of the IGRF on the X, Y, and Z components is approximately 145 nT, 97 nT, and 168 nT respectively. The SCHA model with index $k = 7$ represents a minimum wavelength of 1821 km, while the similar wavelength representation by the IGRF model needs the degree of 22. SHA-based models, such as IGRF, will require high computational loads to obtain an equal wavelength representation. Using equation (11), the SHA model will require 528 Gauss coefficients, compared to only 64 Gauss coefficients in the SCHA model with index $k = 7$ to represent this wavelength. Even the calculation of the SHA model in this degree will be limited by the accuracy of the observatory data. Thus, the SCHA method shows excellence in terms of its accuracy and minimum wavelength property compared with the IGRF. Nevertheless, inaccuracies in the observational data and small scale variations in the local geomagnetic field, where the wavelength is shorter than the minimum wavelength representation of the SCHA model and cannot be modelled, cause inaccuracies in the SCHA model calculations.

A model from interpolation of the SCHA with index $k = 7$ using B-splines basis functions for the year 1985.5 – 2015.5 and its forecast until 2020.5 using autoregressive forecasting method has been obtained. The annual means plots show that the SCHA and IGRF models are well behaved, but the SCHA model gives better results than the IGRF, as can be seen in Tuntungan. For the annual rate of changes

plots, the SCHA is better than the IGRF model as it shows a smooth trend and follows the secular variation accurately on an annual basis. However, where there is a lack of data at the end of the data series, the annual rate of changes of the SCHA shows a mismatch with the annual rate of changes of the IGRF model. Nevertheless, the B-splines basis functions provide reliable results in time-dependent modelling, but complete and evenly distributed data will provide a significant benefit.

By looking at its RMSD, the forecasting calculation suggests that the AR(3) of the SCHA with index $k = 7$ gives better results than the forecasting of the IGRF model, especially on the X, Z, and F components. However, for the Y component, the IGRF is still better than the SCHA model. The RMSD of the forecasted SCHA model is 154.92 nT, 200.87 nT, 104.39 nT, and 135.81 nT for X, Y, Z, and F components, respectively, while the RMSD of the IGRF model is 172.62 nT, 95.52 nT, 117.55 nT, and 162.38 nT for X, Y, Z, and F components respectively. Thus, the forecasted SCHA model is suitable for data reduction in geomagnetic surveys in the Indonesian region but not for navigation.

The geomagnetic surveys for mapping magnetic anomalies measure the sum of the global and local contributions, and the total field magnetic anomaly is the residual when the magnitude of the predicted field is subtracted from the observed field magnitude (Turner et al., 2015). Therefore, the total magnetic anomaly will be mapped more accurately if the predicted field of the SCHA model is used. However, the accuracy of the forecasted SCHA deteriorates as time moves forward.

Nevertheless, with the availability of this SCHA model, more accurate and detailed maps of some geological features can be obtained in the future in the Indonesian region. These geological features, including mineral deposits that might be considered to have economic potential or seismic faults that can be the potential source of the earthquake. Moreover, it is expected that more research in geomagnetic modelling in the Indonesian region can further develop in order to obtain better results compared with the results of this research.

5.2 Recommendations for Further Research

Recently, regional geomagnetic modelling techniques have been developed. A revised version of the SCHA has been published, such as the Revised Spherical Cap Harmonic Analysis (R-SCHA) by Thébault et al. (2006). This technique is suitable for calculating the regional geomagnetic field using data recorded at different altitudes, e.g., combining ground-based geomagnetic data with satellite data. Thus, several sources of satellite data, such as the European Swarm satellite mission, could be used

to improve the accuracy of the regional geomagnetic model in Indonesia. This would allow the regional geomagnetic model to be modelled in 3D format. Moreover, more data could be added to the observational data. The last repeat stations' survey in Indonesia is conducted at the end of 2020, which is when this thesis is being written. The addition of this data will improve the current SCHA model and extend the forecasting model into 2025.5. Furthermore, to solve the reliability of the forecasted geomagnetic model, the monthly quasi definitive data or annual definitive data at the geomagnetic observatories used in this research can be used to update the forecast regularly. Lastly, another recommendation is to combine several small or regional geomagnetic models into a larger region, as has been discussed by Schachtschneider et al. (2012).

References

- Allredge, L. R. (1981). Rectangular harmonic analysis applied to the geomagnetic field. *Journal of Geophysical Research*, 86(B4), 3021–3026. <https://doi.org/10.1029/JB086iB04p03021>
- Barton, C. E. (1988). Global and regional geomagnetic reference fields. In *Exploration Geophysics* (Vol. 19, Issue 3, pp. 401–416). <https://doi.org/10.1071/EG988401>
- Beggan, C., & Whaler, K. (2010). Forecasting secular variation using core flows. *Earth, Planets and Space*, 62(10), 821–828. <https://doi.org/10.5047/eps.2010.07.004>
- Blakely, R. J. (1995). Potential Theory in Gravity and Magnetic Applications. In *Cambridge University Press*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511549816>
- Bloxham, J., & Jackson, A. (1991). Fluid flow near the surface of Earth's outer core. *Reviews of Geophysics*, 29(1), 97. <https://doi.org/10.1029/90RG02470>
- Brown, W. J., Mound, J. E., & Livermore, P. W. (2013). Jerks abound: An analysis of geomagnetic observatory data from 1957 to 2008. *Physics of the Earth and Planetary Interiors*, 223, 62–76. <https://doi.org/10.1016/j.pepi.2013.06.001>
- Campbell, W. H. (2003). *Introduction to Geomagnetic Fields*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139165136>
- Chulliat, A., & Maus, S. (2014). Geomagnetic secular acceleration, jerks, and a localized standing wave at the core surface from 2000 to 2010. *Journal of Geophysical Research: Solid Earth*, 119(3), 1531–1543. <https://doi.org/10.1002/2013JB010604>
- Chulliat, A., Thébaud, E., & Hulot, G. (2010). Core field acceleration pulse as a common cause of the 2003 and 2007 geomagnetic jerks. *Geophysical Research Letters*, 37(7), n/a-n/a. <https://doi.org/10.1029/2009GL042019>
- Constable, C. G., & Constable, S. C. (2004). *Satellite magnetic field measurements: Applications in studying the deep Earth* (pp. 147–159). <https://doi.org/10.1029/150GM13>
- Cox, G. A., Brown, W. J., Billingham, L., & Holme, R. (2018). MagPySV: A Python Package for Processing and Denoising Geomagnetic Observatory Data. *Geochemistry, Geophysics, Geosystems*, 19(9), 3347–3363. <https://doi.org/10.1029/2018GC007714>
- Creer, K. M. (1988). Geomagnetic Field and Radiocarbon Activity Through Holocene Time. In *Secular Solar and Geomagnetic Variations in the Last 10,000 Years* (pp. 381–397). Springer Netherlands. https://doi.org/10.1007/978-94-009-3011-7_24
- De Santis, A. (1991). Translated origin spherical cap harmonic analysis. *Geophysical Journal International*, 106(1), 253–263. <https://doi.org/10.1111/j.1365-246X.1991.tb04615.x>

- De Santis, Angelo, Gaya-Piqué, L., Dominici, G., Meloni, A., Torta, J. M., & Tozzi, R. (2003). Italian Geomagnetic Reference Field (ITGRF): Update for 2000 and secular variation model up to 2005 by autoregressive forecasting. *Annals of Geophysics*, 46. <https://doi.org/doi.org/10.4401/ag-3422>
- De Santis, Angelo, Kerridge, D. J., & Barraclough, D. R. (1989). A Spherical Cap Harmonic Model of the Crustal Magnetic Anomaly Field in Europe Observed by Magsat. In F. J. Lowes, D. W. Collinson, J. H. Parry, S. K. Runcorn, D. C. Tozer, & A. Soward (Eds.), *Geomagnetism and Palaeomagnetism* (pp. 1–17). Springer Netherlands. <https://doi.org/10.1007/978-94-009-0905-2>
- Demetrescu, A. N., & Dobric, V. (2005). *Recent Secular Variation of the Geomagnetic Field . New Insights From Long Series of Observatory Data*. 63–72.
- Dumberry, M., & Finlay, C. C. (2007). Eastward and westward drift of the Earth's magnetic field for the last three millennia. *Earth and Planetary Science Letters*, 254(1–2), 146–157. <https://doi.org/10.1016/j.epsl.2006.11.026>
- Düzgüt, Z., & Malin, S. R. C. (2000). Assessment of regional geomagnetic field modelling methods using a standard data set: Spherical cap harmonic analysis. *Geophysical Journal International*, 141(3), 829–831. <https://doi.org/10.1046/j.1365-246X.2000.00099.x>
- Feng, Y., Holme, R., Cox, G. A., & Jiang, Y. (2018). The geomagnetic jerk of 2003.5-characterisation with regional observatory secular variation data. *Physics of the Earth and Planetary Interiors*, 278(March), 47–58. <https://doi.org/10.1016/j.pepi.2018.03.005>
- Feng, Y., Jiang, Y., Jiang, Y., Liu, B. J., Jiang, J., Liu, Z. W., Ye, M. C., Wang, H. S., & Li, X. M. (2016). Spherical cap harmonic analysis of regional magnetic anomalies based on CHAMP satellite data. *Applied Geophysics*, 13(3), 561–569. <https://doi.org/10.1007/s11770-016-0567-8>
- Finlay, C. C., Maus, S., Beggan, C. D., Bondar, T. N., Chambodut, A., Chernova, T. A., Chulliat, A., Golovkov, V. P., Hamilton, B., Hamoudi, M., Holme, R., Hulot, G., Kuang, W., Langlais, B., Lesur, V., Lowes, F. J., Lühr, H., Macmillan, S., Manda, M., ... Zvereva, T. I. (2010). International Geomagnetic Reference Field: The eleventh generation. *Geophysical Journal International*, 183(3), 1216–1230. <https://doi.org/10.1111/j.1365-246X.2010.04804.x>
- Fiori, R.A.D., Boteler, D. H., Koustov, A. V., Knudsen, D., & Burchill, J. K. (2014). Investigation of localized 2D convection mapping based on artificially generated Swarm ion drift data. *Journal of Atmospheric and Solar-Terrestrial Physics*, 114, 30–41. <https://doi.org/10.1016/j.jastp.2014.04.004>
- Fiori, Robyn A. D. (2020). Spherical Cap Harmonic Analysis Techniques for Mapping High-Latitude Ionospheric Plasma Flow—Application to the Swarm Satellite Mission. In *Ionospheric Multi-Spacecraft Analysis Tools* (pp. 189–218). Springer International Publishing. https://doi.org/10.1007/978-3-030-26732-2_9
- Fournier, A., Eymin, C., & Alboussière, T. (2007). A case for variational geomagnetic data assimilation: Insights from a one-dimensional, nonlinear, and sparsely observed MHD system. *Nonlinear Processes in Geophysics*, 14(2), 163–180.

- <https://doi.org/10.5194/npg-14-163-2007>
- Glatzmaiers, G. A., & Roberts, P. H. (1995). A three-dimensional self-consistent computer simulation of a geomagnetic field reversal. *Nature*, 377(6546), 203–209. <https://doi.org/10.1038/377203a0>
- Haines, G. V. (1985a). Spherical cap harmonic analysis. *Journal of Geophysical Research*, 90(B3), 2583–2591. <https://doi.org/10.1029/JB090iB03p02583>
- Haines, G. V. (1985b). Spherical cap harmonic analysis of geomagnetic secular variation over Canada 1960–1983. *Journal of Geophysical Research*, 90(B14). <https://doi.org/10.1029/jb090ib14p12563>
- Haines, G. V. (1988). Computer programs for spherical cap harmonic analysis of potential and general fields. *Computers and Geosciences*, 14(4), 413–447. [https://doi.org/10.1016/0098-3004\(88\)90027-1](https://doi.org/10.1016/0098-3004(88)90027-1)
- Haines, G. V., & Fiori, R. A. D. (2013). Modeling by singular value decomposition and the elimination of statistically insignificant coefficients. *Computers and Geosciences*, 58, 19–28. <https://doi.org/10.1016/j.cageo.2013.04.021>
- Hampel, F. R. (1974). The influence curve and its role in robust estimation. *Journal of the American Statistical Association*, 69(346), 383–393. <https://doi.org/10.1080/01621459.1974.10482962>
- Holschneider, M., Chambodut, A., & Manda, M. (2003). From global to regional analysis of the magnetic field on the sphere using wavelet frames. *Physics of the Earth and Planetary Interiors*, 135(2–3), 107–124. [https://doi.org/10.1016/S0031-9201\(02\)00210-8](https://doi.org/10.1016/S0031-9201(02)00210-8)
- Hulot, G., Sabaka, T. J., Olsen, N., & Fournier, A. (2015). The Present and Future Geomagnetic Field. In *Treatise on Geophysics* (pp. 33–78). Elsevier. <https://doi.org/10.1016/B978-0-444-53802-4.00096-8>
- Hwang, C., & Chen, S. (1997). Fully normalized spherical cap harmonics: application to the analysis of sea-level data from TOPEX/POSEIDON and ERS-1. *Geophysical Journal International*, 129(2), 450–460. <https://doi.org/10.1111/j.1365-246X.1997.tb01595.x>
- Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: Principles and Practice. *Principles of Optimal Design*, 421–455. <https://doi.org/10.1017/9781316451038.010>
- Ingham, L. (2009). *The Geomagnetic Field in the South Pacific During the 19th Century*.
- Jackson, A., & Finlay, C. (2015). Geomagnetic Secular Variation and Its Applications to the Core. In *Treatise on Geophysics* (pp. 137–184). Elsevier. <https://doi.org/10.1016/B978-0-444-53802-4.00099-3>
- Jackson, Andrew, Jonkers, A. R. T., & Walker, M. R. (2000). Four centuries of geomagnetic secular variation from historical records. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 358(1768), 957–990. <https://doi.org/10.1098/rsta.2000.0569>
- Jankowski, J., & Sucksdorff, C. (1996). Guide for Magnetic Measurements and

- Observatory Practice. In *IAGA*. <http://iugg.org/IAGA/iaga/pages/pdf/IAGA-Guide-Observatories.pdf>
- Kono, M. (2015). Geomagnetism: An Introduction and Overview. In *Treatise on Geophysics* (Vol. 5, pp. 1–31). Elsevier. <https://doi.org/10.1016/B978-0-444-53802-4.00095-6>
- Korte, M., & Haak, V. (2000). Modeling European magnetic repeat station and survey data by SCHA in search of time-varying anomalies. *Physics of the Earth and Planetary Interiors*, 122(3–4), 205–220. [https://doi.org/10.1016/S0031-9201\(00\)00194-1](https://doi.org/10.1016/S0031-9201(00)00194-1)
- Korte, M., & Holme, R. (2003). Regularization of spherical cap harmonics. *Geophysical Journal International*, 153(1), 253–262. <https://doi.org/10.1046/j.1365-246X.2003.01898.x>
- Korte, Monika, & Manda, M. (2019). Geomagnetism: From Alexander von Humboldt to Current Challenges. *Geochemistry, Geophysics, Geosystems*, 20(8), 3801–3820. <https://doi.org/10.1029/2019GC008324>
- Kotzé, P. B. (2019). Geomagnetic secular variation changes in southern Africa during the SWARM period 2013–2018. *Annals of Geophysics*, 62, 1–13. <https://doi.org/10.4401/ag-8126>
- Kuang, W., Tangborn, A., Wei, Z., & Sabaka, T. (2009). Constraining a numerical geodynamo model with 100 years of surface observations. *Geophysical Journal International*, 179(3), 1458–1468. <https://doi.org/10.1111/j.1365-246X.2009.04376.x>
- Langel, R. A., Kerridge, D. J., Barraclough, D. R., & Malin, S. R. C. (1986). Geomagnetic temporal change: 1903–1982, A spline representation. *Journal of Geomagnetism and Geoelectricity*, 38(6), 573–597. <https://doi.org/10.5636/jgg.38.573>
- Lanza, R., & Meloni, A. (2006). *The Earth's Magnetism An Introduction for Geologists*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-540-27980-8>
- Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49(4), 764–766. <https://doi.org/10.1016/j.jesp.2013.03.013>
- Love, J. J. (2008). Magnetic monitoring of earth and space. *Physics Today*, 61(2), 31–37. <https://doi.org/10.1063/1.2883907>
- Macmillan, S., & Finlay, C. (2011). The International Geomagnetic Reference Field. In *Geomagnetic Observations and Models* (pp. 265–276). Springer Netherlands. https://doi.org/10.1007/978-90-481-9858-0_10
- Macmillan, S., & Maus, S. (2005). International Geomagnetic Reference Field—the tenth generation. *Earth, Planets and Space*, 57(12), 1135–1140. <https://doi.org/10.1186/BF03351896>
- Maier, T., & Mayer, C. (2003). Multiscale Downward Continuation of CHAMP

- FGM-Data for Crustal Field Modelling. In *First CHAMP Mission Results for Gravity, Magnetic and Atmospheric Studies* (pp. 288–295). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-38366-6_42
- Malin, S. R. C. (1983). Modelling the geomagnetic field. *Geophysical Journal International*, 74(1), 147–157. <https://doi.org/10.1111/j.1365-246X.1983.tb01874.x>
- Malin, S. R. C. (1985). On the unpredictability of geomagnetic secular variation. *Physics of the Earth and Planetary Interiors*, 39(4), 293–296. [https://doi.org/10.1016/0031-9201\(85\)90142-6](https://doi.org/10.1016/0031-9201(85)90142-6)
- Mandea, M., Holme, R., Pais, A., Pinheiro, K., Jackson, A., & Verbanac, G. (2010). Geomagnetic jerks: Rapid core field variations and core dynamics. *Space Science Reviews*, 155(1–4), 147–175. <https://doi.org/10.1007/s11214-010-9663-x>
- McKnight, J. D. (1990). *The New Zealand Regional Geomagnetic Field Mapping, Modelling, and Interpretation* (Issue April). Victoria University of Wellington.
- McKnight, J. D. (1996). *An Updated Regional Geomagnetic Field Model for New Zealand*.
- Merrill, R. T., McElhinny, M. W., & McFadden, P. L. (1996). The Magnetic Field of the Earth: Paleomagnetism, the Core, and the Deep Mantle. In *International Geophysics* (Vol. 63). Academic Press.
- Metodiev, M., & Trifonova, P. (2017). Bulgarian Geomagnetic Reference Field (BulGRF) for 2015.0 and secular variation prediction model up to 2020. *Annales Geophysicae*, 35(5), 1085–1092. <https://doi.org/10.5194/angeo-35-1085-2017>
- Nevanlinna, H., Rynö, J., Haines, G. V., & Borg, K. (1988). Spherical cap harmonic analysis applied to the Scandinavian geomagnetic field 1985.0. *Deutsche Hydrographische Zeitschrift*, 41(3–6), 177–186. <https://doi.org/10.1007/BF02225927>
- NIST/SEMATECH. (2012). NIST/SEMATECH e-Handbook of Statistical Methods. In *e-Handbook of Statistical Methods*. <https://doi.org/10.18434/M32189>
- Olson, P. (2015). Core Dynamics: An Introduction and Overview. In *Treatise on Geophysics: Second Edition* (Vol. 8). <https://doi.org/10.1016/B978-0-444-53802-4.00137-8>
- Pavón-Carrasco, F. J., Torta, J. M., Catalán, M., Talarn, Æ., & Ishihara, T. (2013). Improving total field geomagnetic secular variation modeling from a new set of cross-over marine data. *Physics of the Earth and Planetary Interiors*, 216, 21–31. <https://doi.org/10.1016/j.pepi.2013.01.002>
- Peqini, K., Duka, B., Egli, R., & Leichter, B. (2018). Crustal geomagnetic field and secular variation by regional and global models for Austria. *Austrian Journal of Earth Sciences*, 111(1), 48–63. <https://doi.org/10.17738/ajes.2018.0004>
- Purucker, M. E., & Whaler, K. A. (2015). Crustal Magnetism. In *Treatise on Geophysics* (pp. 185–218). Elsevier. <https://doi.org/10.1016/B978-0-444-53802-4.00100-7>

- Qamili, E., De Santis, A., Cianchini, G., Duka, B., Gaya-Piqué, L. R., Dominici, G., & Hyka, N. (2010). Two geomagnetic regional models for Albania and south-east Italy from 1990 to 2010 with prediction to 2012 and comparison with IGRF-11. *Earth, Planets and Space*, *62*(10), 833–841. <https://doi.org/10.5047/eps.2010.07.011>
- Ryskin, G. (2009). Secular variation of the Earth's magnetic field: Induced by the ocean flow? *New Journal of Physics*, *11*. <https://doi.org/10.1088/1367-2630/11/6/063015>
- Santis, A. De, Battelli, O., & Kerridge, D. J. (1990). Spherical cap harmonic analysis applied to regional field modelling for Italy. *Journal of Geomagnetism and Geoelectricity*, *42*(9), 1019–1036. <https://doi.org/10.5636/jgg.42.1019>
- Schachtschneider, R., Holschneider, M., & Mande, M. (2012). Error distribution in regional modelling of the geomagnetic field. *Geophysical Journal International*, *191*(3), 1015–1024. <https://doi.org/10.1111/j.1365-246X.2012.05675.x>
- Schott, J.-J., & Thébaud, E. (2011). Modelling the Earth's Magnetic Field from Global to Regional Scales. In *Geomagnetic Observations and Models* (pp. 229–264). Springer Netherlands. https://doi.org/10.1007/978-90-481-9858-0_9
- Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and Statistical Modeling with Python. *Proc. of the 9th Python in Science Conf, December*, 92–96. <http://statsmodels.sourceforge.net/>
- Simons, F. J., & Dahlen, F. A. (2006). Spherical slepian functions and the polar gap in geodesy. *Geophysical Journal International*, *166*(3), 1039–1061. <https://doi.org/10.1111/j.1365-246X.2006.03065.x>
- Sun, Z., Tangborn, A., & Kuang, W. (2007). Data assimilation in a sparsely observed one-dimensional modeled MHD system. *Nonlinear Processes in Geophysics*, *14*(2), 181–192. <https://doi.org/10.5194/npg-14-181-2007>
- Syirojudin, M., Murjaya, J., Zubaidah, S., Hasanudin, H., Ahadi, S., Efendi, N., & Suroyo, T. (2018). Indonesian Geomagnetic Maps for Epoch 2015.0 to cover of Indonesian Regions. *IOP Conference Series: Earth and Environmental Science*, *132*(1), 0–6. <https://doi.org/10.1088/1755-1315/132/1/012003>
- Syirojudin, Muhamad, & Bijaksana, S. (2013). Comparing the accuracy of BMKG geomagnetic maps with those of IGRF over the Indonesian region. *AIP Conference Proceedings*, *1554*, 249–252. <https://doi.org/10.1063/1.4820332>
- Thébaud, E. (2008). A proposal for regional modelling at the Earth's surface, R-SCHA2D. *Geophysical Journal International*, *174*(1), 118–134. <https://doi.org/10.1111/j.1365-246X.2008.03823.x>
- Thébaud, E., Finlay, C. C., Alken, P., Beggan, C. D., Canet, E., Chulliat, A., Langlais, B., Lesur, V., Lowes, F. J., Manoj, C., Rother, M., & Schachtschneider, R. (2015). Evaluation of candidate geomagnetic field models for IGRF-12. *Earth, Planets and Space*, *67*(1). <https://doi.org/10.1186/s40623-015-0273-4>
- Thébaud, E., Finlay, C. C., Beggan, C. D., Alken, P., Aubert, J., Barrois, O., Bertrand, F., Bondar, T., Boness, A., Brocco, L., Canet, E., Chambodut, A.,

- Chulliat, A., Coïsson, P., Civet, F., Du, A., Fournier, A., Fratter, I., Gillet, N., ... Zvereva, T. (2015). International Geomagnetic Reference Field: the 12th generation. *Earth, Planets and Space*, 67(1), 79. <https://doi.org/10.1186/s40623-015-0228-9>
- Thébault, E., Schott, J. J., & Mandea, M. (2006). Revised spherical cap harmonic analysis (R-SCHA): Validation and properties. *Journal of Geophysical Research: Solid Earth*, 111(1), 1–17. <https://doi.org/10.1029/2005JB003836>
- Toh, H., Kanezaki, H., & Ichiki, M. (2007). A regional model of the geomagnetic field over the Pacific Ocean for epoch 2002. *Geophysical Research Letters*, 34(9), 1–5. <https://doi.org/10.1029/2007GL029341>
- Torta, J. M., & Da Santis, A. (1996). On the derivation of the Earth's conductivity structure by means of spherical cap harmonic analysis. *Geophysical Journal International*, 127(2), 441–451. <https://doi.org/10.1111/j.1365-246X.1996.tb04732.x>
- Torta, J. M., García, A., Curto, J. J., & De Santis, A. (1992). New representation of geomagnetic secular variation over restricted regions by means of spherical cap harmonic analysis: application to the case of Spain. *Physics of the Earth and Planetary Interiors*, 74(3–4), 209–217. [https://doi.org/10.1016/0031-9201\(92\)90011-J](https://doi.org/10.1016/0031-9201(92)90011-J)
- Torta, J. M., García, A., & Santis, A. DE. (1993). A Geomagnetic Reference Field for Spain at 1990. *Journal of Geomagnetism and Geoelectricity*, 45(7), 573–588. <https://doi.org/10.5636/jgg.45.573>
- Torta, J. Miquel. (2019). Modelling by Spherical Cap Harmonic Analysis: A Literature Review. In *Surveys in Geophysics* (Issue 0123456789). Springer Netherlands. <https://doi.org/10.1007/s10712-019-09576-2>
- Torta, J. Miquel, Gaya-Piqué, L. R., & De Santis, A. (2006). Spherical Cap Harmonic Analysis of the Geomagnetic Field with Application for Aeronautical Mapping. In J. L. Rasson & T. Delipetrov (Eds.), *Geomagnetics for Aeronautical Safety* (pp. 291–307). Springer Netherlands. https://doi.org/10.1007/978-1-4020-5025-1_23
- Torta, J. Miquel, Pavón-Carrasco, F. J., Marsal, S., & Finlay, C. C. (2015). Evidence for a new geomagnetic jerk in 2014. *Geophysical Research Letters*, 42(19), 7933–7940. <https://doi.org/10.1002/2015GL065501>
- Turner, G. (2011). *North Pole, South Pole The Epic Quest to Solve the Great Mystery of Earth's Magnetism*. The Experiment.
- Turner, G. M., Rasson, J., & Reeves, C. (2015). Observation and Measurement Techniques. In *Treatise on Geophysics* (pp. 91–135). Elsevier. <https://doi.org/10.1016/B978-0-444-53802-4.00098-1>
- Verbanac, G. (2007). On regional modeling of the main geomagnetic field. *Geofizika*, 24(1), 1–27.
- Walker, J. K. (1989). Spherical cap harmonic modelling of high latitude magnetic activity and equivalent sources with sparse observations. *Journal of Atmospheric*

and Terrestrial Physics, 51(2), 67–80. [https://doi.org/10.1016/0021-9169\(89\)90106-2](https://doi.org/10.1016/0021-9169(89)90106-2)

Wardinski, I., & Holme, R. (2011). Signal from noise in geomagnetic field modelling: Denoising data for secular variation studies. *Geophysical Journal International*, 185(2), 653–662. <https://doi.org/10.1111/j.1365-246X.2011.04988.x>

Wessel, P., Luis, J. F., Uieda, L., Scharroo, R., Wobbe, F., Smith, W. H. F., & Tian, D. (2019). The Generic Mapping Tools Version 6. *Geochemistry, Geophysics, Geosystems*, 20(11), 5556–5564. <https://doi.org/10.1029/2019GC008515>