



OPEN

## The accuracy of geostatistics for regional geomagnetic modeling in an archipelago setting

Muhamad Syirojudin<sup>1,3,4</sup>, Eko Haryono<sup>2,4</sup>✉ & Suaidi Ahadi<sup>3,4</sup>

Indonesia as an archipelago country relies on a limited number and clustered distributed repeat station networks. This paper explores the use of geostatistical modeling to overcome this data limitation. The model data set consisted of repeat station data from 1985 to 2015 epoch. The geostatistical methods utilized included ordinary kriging (OK), collocated cokriging (CC), and kriging with external drift (KED). The model generated using these geostatistical methods was then compared to spherical cap harmonic analyses (SCHA) and polynomial models. The geostatistical model was shown to perform better, with greater accuracy in declination, inclination, and total intensity, as indicated by the root mean square error (RMSE). We have demonstrated that the geostatistical method is a promising approach in the modeling of regional geomagnetic field, especially in areas with limited and clustered distributed data.

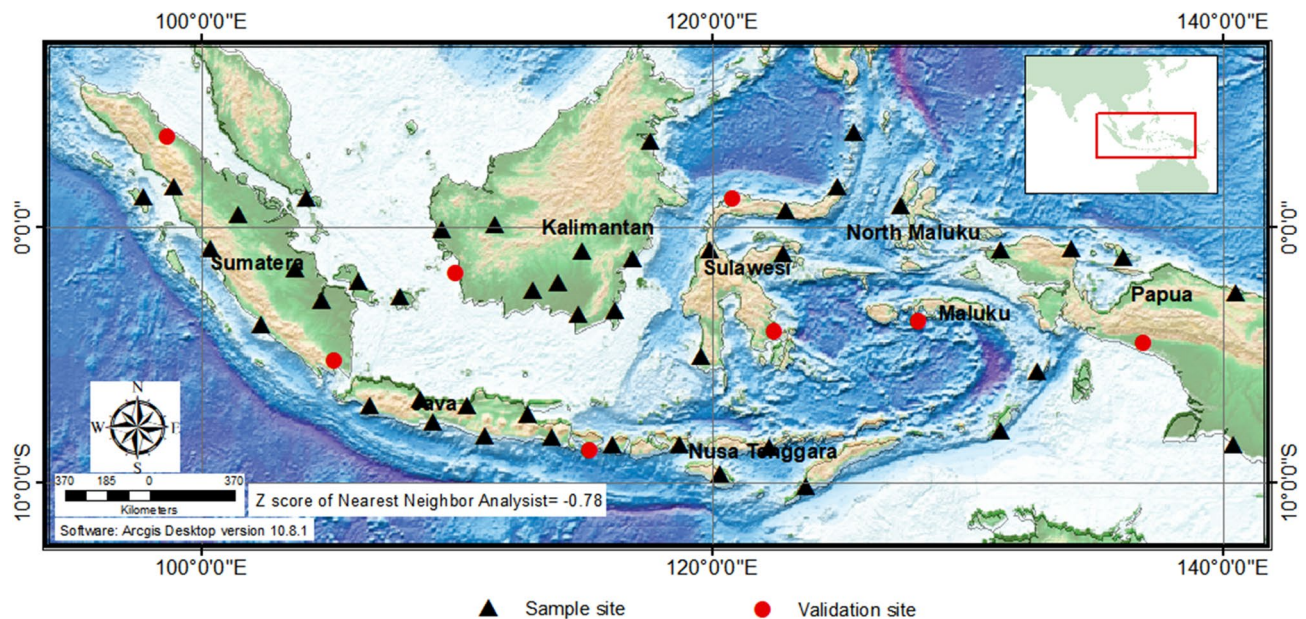
### Abbreviations

BMKG	Badan Meteorologi, Klimatologi dan Geofisika, or the Meteorological, Climatological, and Geophysical Agency of Indonesia
CC	Collocated cokriging
DGRF	Definitive geomagnetic reference field
IAGA	International association of geomagnetism and aeronomy
IDW	Inverse distance weighting
IGRF	International geomagnetic reference field
IUGG	International union of geodesy and geophysics
IQSY	International years of the quiet sun
KED	Kriging with external drift
OK	Ordinary kriging
SCHA	Spherical cap harmonic analysis
RMSE	Root means square error
WMS	World magnetic survey

The geomagnetic field is dynamic as a result of fluid and material movement in the earth's core, mantle, and crust<sup>1,2</sup>. Therefore, periodic observation and modeling are needed to produce up-to-date geomagnetic field maps. An isogonic chart of declination components was made for the first time by Halley in 1701<sup>3</sup>, and Gauss initiated the global geomagnetic field model in 1832 based on spherical harmonic analysis<sup>4</sup>. Modern global geomagnetic field mapping began in 1957 at the International Union of Geodesy and Geophysics (IUGG) meeting in Toronto. The conference marked the birth of the World Magnetic Survey<sup>5</sup>. In 1960, a mathematical model based on spherical harmonic analysis was recommended. By international agreement, the International Geomagnetic Reference Field (IGRF) model<sup>6,7</sup> of the geomagnetic field and its secular variations<sup>8</sup> was thus established.

The IGRF model is updated every five years with oversight from the International Association of Geomagnetism and Aeronomy (IAGA). Improvements to the IGRF model's accuracy have been achieved by incorporating satellite data and adding the truncation order of the spherical harmonic analysis equation. However, there are still limitations involving accuracy and spatial resolution<sup>9</sup>. These limitations are caused by the truncation order of the spherical harmonic analysis equation<sup>7</sup>, asymmetric distribution of repeat stations and geomagnetic observatories<sup>10</sup>, and the inability to model lithospheric geomagnetic anomaly due to its long spectrum<sup>11</sup>.

<sup>1</sup>Doctoral Program in Geography, Faculty of Geography, Gadjah Mada University, Yogyakarta, Indonesia. <sup>2</sup>Department of Environmental Geography, Gadjah Mada University, Yogyakarta, Indonesia. <sup>3</sup>Meteorological, Climatological, and Geophysical Agency, Jakarta 10720, Indonesia. <sup>4</sup>These authors contributed equally: Muhamad Syirojudin, Eko Haryono and Suaidi Ahadi. ✉email: e.haryono@ugm.ac.id



**Figure 1.** Distribution of the 45 sample locations and the 8 locations designated for data validation, Epoch 2010. The Z score is the nearest-neighbor analysis value; a negative score indicates cluster distribution. The base map is from DEMNAS (the national digital elevation model) published by the National Mapping Agency of Indonesia (<https://tanahair.indonesia.go.id/demnas/#/>). The national bathymetry is generated from the inversion of gravity anomaly data incorporated with additional single and multibeam sounding by the National Mapping Agency of Indonesia, National Geophysical Data Center, British Oceanographic Data Centre, Agency for the Assessment and Application of Technology Indonesia, Indonesian Institute of Sciences, and Research and Development Center for Marine Geology of Indonesia.

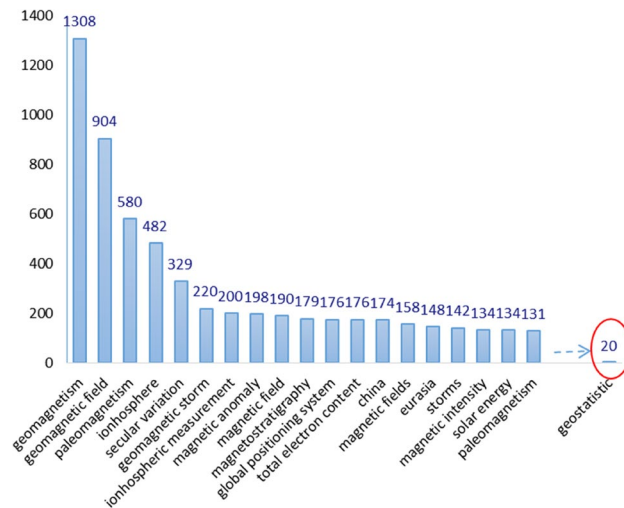
Comparing IGRF and regional geomagnetic data in China shows that the difference in total intensity can reach 143 nT<sup>12</sup>. For research and exploration, geomagnetic data require an accuracy of 0.1° in the declination component and 50 nT in the total intensity<sup>13</sup>. Furthermore, the spherical harmonic analysis equation cannot be used for a geomagnetic field in the form of an orthogonal area<sup>14</sup>. This has led to the introduction of new regional geomagnetic field models in the hope of producing a more accurate model.

Based on IAGA Resolution No. 23 (1963), “Repeat station measurements during the International Years of the Quiet Sun (IQSY) in support of the World Magnetic Survey (WMS),” the Badan Meteorologi, Klimatologi dan Geofisika (BMKG), or the Meteorological, Climatological, and Geophysical Agency of Indonesia have conducted measurements in repeat stations to cover the region of Indonesia starting in 1985. These measurements were routinely carried out every five years. The locations of the Indonesian repeat stations comply with some of the criteria set forth by the IAGA<sup>15</sup>. The cluster distribution of the repeat stations of the Indonesian archipelago is characterized by nearest-neighbor analysis with a negative value (Z score = −0.78). Repeat station distribution is difficult to make uniform<sup>16</sup>. The repeat station locations cannot be placed on the water, as those sites cannot be accessed. No locations were found that met the standards for repeat stations<sup>17</sup>. This situation problematizes the regional geomagnetic field modeling framework in Indonesia (Fig. 1).

The methods generally used in regional geomagnetic field modeling include polynomial<sup>18,19</sup>, spline, Taylor expansion<sup>20,21</sup>, spherical cap harmonic analysis (SCHA)<sup>22</sup>, revised spherical cap harmonic analysis<sup>23</sup>, and revised spherical cap harmonic analysis for the Earth’s surface<sup>24</sup>. These methods typically require dense data for modeling. The boundary effects of the existing methods are also affected by the data configuration. Numerical instability due to wide data gaps resulting from sparse data patterns needs to be balanced with statistical or physical regularization to increase model reliability<sup>25</sup>.

We propose the geostatistical method as an alternative for regional geomagnetic field modeling using limited and clustered data patterns. A geostatistical method predicts the value of data property which is not covered by the sample data distribution or a sparse data pattern<sup>26</sup>. Originally developed to predict the distribution percentages of ores for mining<sup>27</sup>, geostatistical methods also very accurately model natural resources<sup>28</sup> such as underground rivers<sup>29</sup> and reservoirs of oil and gas<sup>30,31</sup>.

Geostatistical methods are classified into two types of algorithms: traditional and geostatistical<sup>32</sup>. The weighted values in the traditional algorithm are based on the geometrical distance of the surrounding estimator data. This includes inverse distance weighting (IDW), nearest neighbor or polygon, and triangulation methods. Geostatistical algorithms, on the other hand, use the structure or statistics of distances from the surrounding estimator data to weigh the data property. The geostatistical kriging algorithm evaluates the distance and direction of the array of sample data points to reflect spatial correlations. Kriging can produce good estimates for anisotropic data and clustered distributions<sup>33</sup>. There are many kinds of kriging algorithms: e.g., simple kriging, ordinary kriging (OK)<sup>34,35</sup>, universal kriging, kriging with external drift (KED)<sup>36</sup>, and collocated cokriging (CC)<sup>37</sup>.



**Figure 2.** The occurrences of words in the titles, keywords, and abstracts of the 3,524 Scopus-indexed papers using the keywords “geomagnetic” and “regional.” The keyword “geostatistical” has 20 occurrences. However, out of those 20, only 3 papers discuss geomagnetic regional modeling.

However, geostatistical methods are rarely used for regional geomagnetic field modeling. A search on Scopus for indexed papers using the keywords “geomagnetic” and “regional” resulted in only 3 of 3524 papers referencing geomagnetic regional modeling (Fig. 2). No paper discusses the geostatistical method for geomagnetic regional modeling with special emphasis on clustered distributed data. Accordingly, the purpose of this paper is to explore the accuracy of the geostatistical method in regional geomagnetic field modeling and mapping.

## Results

The geostatistical method generally yields a smaller root mean square error (RMSE) than do other existing methods for regional geomagnetic modeling in the Indonesian archipelago. The smallest RMSE for the declination component (7.03 minutes) is yielded by the collocated cokriging (CC) method and an identical range of 2.048.97 minutes using both KED and OK (Fig. 3a). The KED method yields a slightly higher RMSE of 7.19 minutes. The OK method, using stable and exponential variograms, yields RMSE values of 7.67 and 7.68 minutes, respectively. On the other hand, the polynomial and SCHA methods show an RMSE of 7.96, and 9.26 minutes, and have widerange values of 1.93–9.67, and 2.9311.97 minutes, respectively.

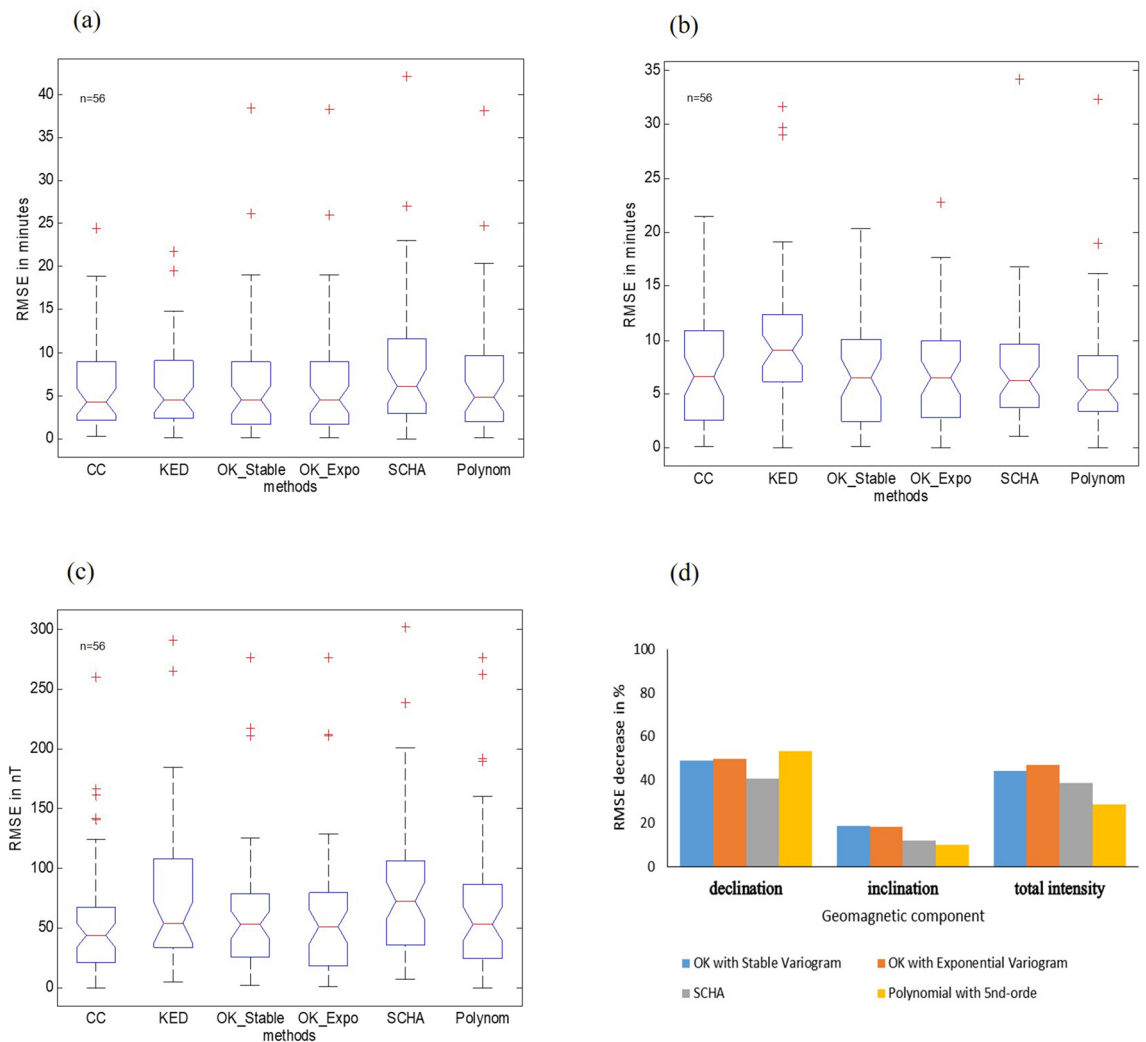
However, for the inclination component, not all geostatistical methods yield small RMSE values. As shown in Fig. 3b, the smallest RMSE is returned using OK, with a stable variogram (6.77 minutes). This is followed by OK with an exponential variogram (6.90 minutes) and has range values of 2.6210.83 minutes. CC has a higher RMSE than the polynomial method, and KED has a higher RMSE than both the polynomial and SCHA methods. The polynomial method has an RMSE of 6.91 minutes with a range of 3.448.55 minutes, the CC has an RMSE of 7.19 minutes and has an identical range with OK, the SCHA has an RMSE of 8.15 minutes with a range of 3.789.63 minutes, and the KED has an RMSE of 10.28 minutes with a range of 6.1112.36 minutes.

As expected, for total intensity the geostatistical method yields smaller RMSE values as shown in Fig. 3c. The smallest RMSE is obtained by the CC method (57.07 nT) with the shortest range of 21.6567.75 nT, followed by OK (63.58 nT) with a range of 25.6078.20 nT. On the other hand, KED (74.54 nT) has a higher RMSE than the polynomial method (68.97 nT), but smaller than SCHA (82.24 nT).

## Discussion

The results show that the geostatistical method is a good model of the regional geomagnetic field based on data from typical clustered distribution patterns, as exemplified in Indonesia. The geostatistical model using CC was 2.23 minutes more accurate than the SCHA and polynomial models of the declination component. The CC model also was 25.17 nT more accurate than the SCHA and the polynomial models of the total intensity component. However, for the inclination component, the OK model was more accurate (0.43 minutes smaller) than the polynomial and SCHA models. The dataset in Indonesia consists of the archipelago’s clustered data with average distances among the repeat stations of 250 km (Fig. 1). This distance and distribution are not spherical caps enough. The distance among the repeat station points is quite large, as the sea forms a gap among the repeat stations. The large distances between the sample data causes the accuracy of the SCHA to be lower<sup>38,39</sup> because additional statistical or physical regulation are required to compensate for the numerical instability<sup>25</sup>. The same problem occurs with the polynomial method, though typically this method will produce good accuracy if the data distribution is uniform<sup>16</sup>.

The RMSE values of the geomagnetic regional modeling parameters are affected by the spatial distribution of the repeat stations. The same phenomenon was also reported in China<sup>40</sup>, where a randomly distributed pattern produced a noticeable effect. Importantly, the SCHA and polynomial methods do not produce effective models

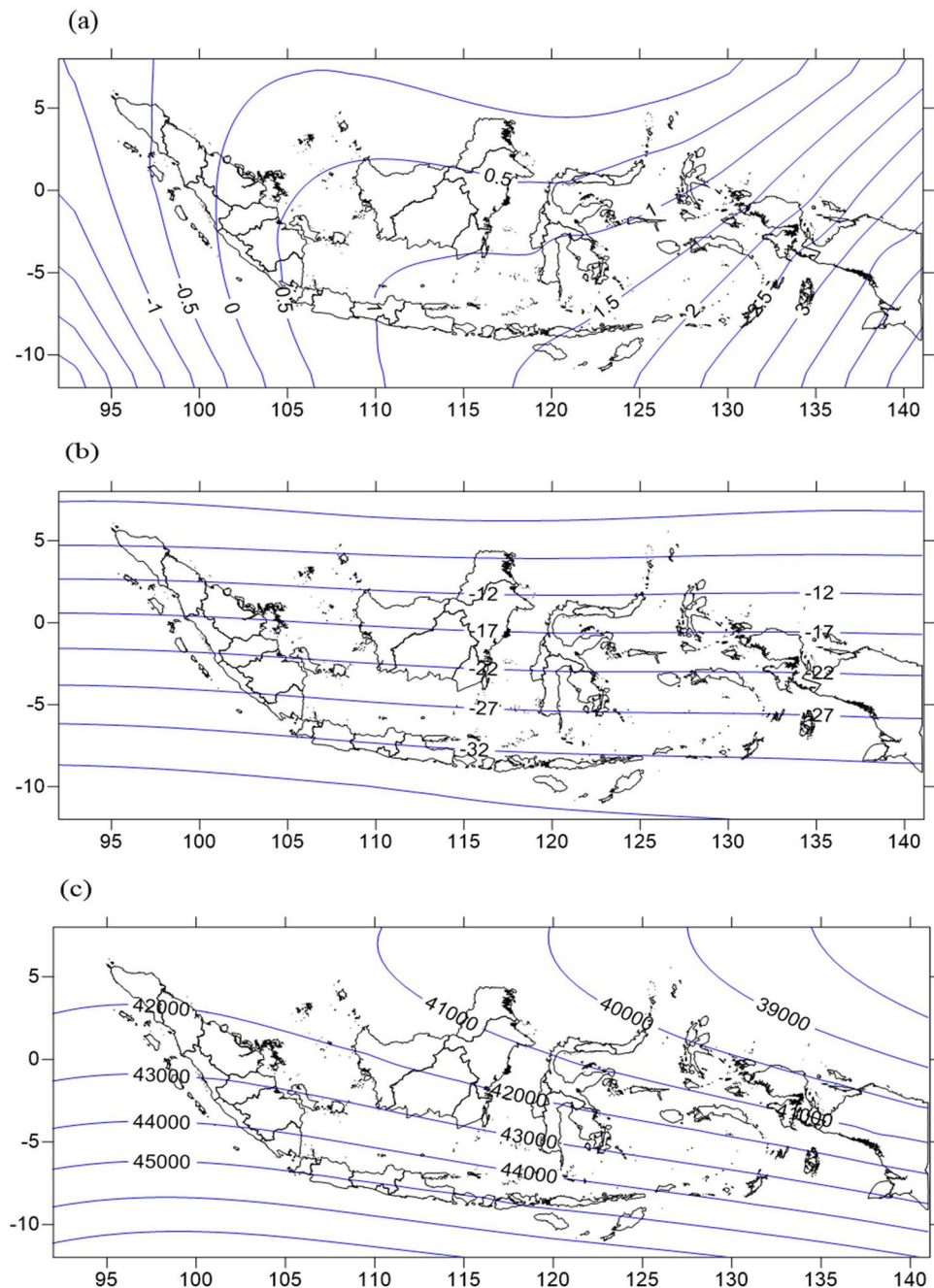


**Figure 3.** Boxplot of root mean square error (RMSE) ( $n=56$  data) between data estimation and validation from the CC, KED, OK methods with stable variogram, OK with exponential variogram, SCHA, and polynomial methods data for epoch 1985–2015 for each geomagnetic component. **(a)** Declination Component (minutes); **(b)** Inclination Component (minutes); **(c)** Total Intensity (nT); **(d)** RMSE decreased below normalized value due to the addition of 28% more sample data.

based on clustered distributed patterns. Both these methods require more regularly and densely distributed data<sup>25,41,42</sup>. That the CC and OK methods might better fit a wider range of data distributions was also suggested by Webster and Oliver<sup>26</sup> and Zhao et al.<sup>43</sup>. The CC and OK methods are suitable for random, clustered, and anisotropic distributions<sup>26,43</sup>. They are also able to deal with different weights between clusters using varied numbers of data sets, which can reduce bias in the estimations<sup>32</sup>. Meanwhile, the SCHA and polynomial methods cannot accept a different number of data sets between clusters.

The average RMSE of the inclination component from OK is slightly lower than with the polynomial method (0.13 minutes). For the declination and total intensity components whose contours are a curving isoline, the CC method produces a smaller average RMSE than the OK, SCHA, and polynomial methods. This is ideal for the curving isoline of geomagnetic components that results from additional secondary data (IGRF-13). This result was also in agreement with the work of Rivoirard<sup>37</sup> and Han et al.<sup>44</sup> that showed CC could result in a lower RMSE for complex data property. The larger RMSE of the SCHA and polynomial methods probably results from the curving isoline of the geomagnetic declination. Indonesia, which lays in the equator, always has a banding isoline; the declination's contours in Indonesia deviate along the longitude (Fig. 4). That the SCHA and polynomial methods have a larger RMSE with a curving isoline was also found by Gu et al.<sup>45</sup> in China and in Europe by Korte et al.<sup>11</sup>. The polynomial method is more accurate for a straight geomagnetic isoline<sup>46</sup>.

Our results also reveal that when adding more repeat station data sets, the geostatistical method results in better increment accuracy than the SCHA and polynomial methods. The research used the 2015 data set (with 28% additional sample data), compared to the data set from 1990 to 2010 (Fig. 3d). The accuracy of the OK method increased by 47% for the total intensity component. This increment of OK accuracy is 18% higher than that of the SCHA and polynomial methods. The accuracy of OK also increases by 19% for the inclination component,



**Figure 4.** Geomagnetic chart for epoch 2015.0 using the collocated cokriging method in Indonesia; (a) Declination Comp. with  $\Delta = 0.5^\circ$ , (b) Inclination Comp. with  $\Delta = 5^\circ$ , and (c) Total Intensity Comp. with  $\Delta = 1000$  nT.

which is 9% higher than SCHA and polynomial values. For the declination component, the accuracy of OK also increases by 50%, or 9% higher than SCHA, but a bit lower (3%) than polynomial. However, the accuracy of the CC and KED methods did not increase as significantly as the OK method. The use of a secondary variable<sup>36,37,47,48</sup> in the CC and KED methods may account for the smaller increase. This result also showed that with additional data sets, the SCHA and polynomial methods do not result in an accuracy increase as significant as those seen in previous research from Europe<sup>49,50</sup>, where the data set was regularly distributed.

## Methods

We used repeat station data from 1985 to 2015 published by BMKG, and the Indonesian authority granted permission to conduct geomagnetic measurements under the IAGA resolution. Most of the repeat stations are located at the airports of the Indonesian islands. This has the advantage that, in addition to fulfilling the special requirements for the location of the repeat stations, the stations can also support the calibration of runway

azimuth, which is important for aircraft navigation<sup>51,52</sup>. These measurements were then routinely carried out every five years. The data has been reduced by diurnal and secular variation to get the same time for every period.

To validate the regional geomagnetic model produced by the geostatistical methods (OK, CC, and KED) and the geomagnetic methods (SCHA and polynomial), we used 15% of the total existing repeat station data gathered from 1985 to 2015. The data used for model validation purposes was excluded from the modeling process. The composition and distribution of the validation data were selected randomly (Fig. 1). For the main geomagnetic field, we used the Definitive Geomagnetic Reference Field (DGRF) data generated from IGRF-13, since the global model is more accurate than the regional one<sup>53</sup>.

This research tested the accuracy of the geostatistical method and the existing methods to model the regional geomagnetic field within the Indonesian territory. From the geostatistical method group, we used the OK, CC, and KED methods. The OK equation is as follows<sup>32</sup>:

$$Z_{OK}(x_0) = \sum_{i=1}^n \lambda_i Z(x_i). \quad (1)$$

where  $Z_{OK}$  is the data estimation result,  $\lambda$  is the weight of each primary data value, and  $Z(x_i)$  is the data value in the  $i$ th location. This method is easily applied and considered the best linear unbiased estimator (BLUE). This method has been used in many fields of earth science<sup>54–57</sup> and produces reasonable estimates.

The CC equation is as follows<sup>37</sup>:

$$Z_{CC}(x_0) = \sum_{i=1}^n \lambda_i Z_1(x_i) + \mu Z_2(x_0). \quad (2)$$

where  $Z_{CC}$  is the data estimation result,  $\lambda$  is the weight of the measurement datum in the  $i$ th location,  $Z_1$  is the measurement datum in the  $i$ th location,  $\mu$  is the weight of the available complementary datum from IGRF-13, and  $Z_2$  is the available complementary datum from IGRF-13.

The KED equation is as follows<sup>36</sup>:

$$Z_{KED}(x_0) = \sum_{i=1}^n \lambda_i Z_1(x_i) + a \sum_{i=1}^n \lambda_i Z_2(x_i) + a Z_2(x_0). \quad (3)$$

where  $Z_{KED}(x_0)$  is the estimation result in location  $x_0$ ,  $Z_1(x_i)$  is the measurement datum from the BMKG in the  $i$ th repeat station,  $\lambda_i$  is the weight of  $Z_1(x_i)$  and  $Z_2(x_i)$ , and  $a$  is the slope of the available complementary data from IGRF-13 in location  $x_0$ .

The geostatistical method was then compared with the SCHA and polynomial methods. The SCHA method was chosen because it has been used frequently by researchers in recent decades<sup>25</sup>. The potential field equation  $V$  in the SCHA method is the following:

$$\begin{aligned} V(r, \theta, \lambda) = & a \sum_{k \geq 0} \sum_{m \geq 0} \left(\frac{a}{r}\right)^{n_k(m)+1} (g_{n_k}^{i,m} \cos(m\lambda) + h_{n_k}^{i,m} \sin(m\lambda)) P_{n_k}^m \theta \\ & + a \sum_{k \geq 0} \sum_{m \geq 0} \left(\frac{r}{a}\right)^{n_k(m)} (g_{n_k}^{e,m} \cos(m\lambda) + h_{n_k}^{e,m} \sin(m\lambda)) P_{n_k}^m \theta. \end{aligned} \quad (4)$$

where  $V$  = geomagnetic potency,  $a$  = earth's radius (6,371.2 km),  $r$  = radial distance of the set location from the Earth's core,  $\theta$  = colatitude ( $90^\circ$  - latitude),  $\lambda$  = longitude,  $P_{n_k}^m \theta$  = Legendre function of the associated Schmidt non-integer  $n_k(m)$  and  $m$ ,  $g_{n_k}^{i,m}$  and  $h_{n_k}^{i,m}$  are the Gaussian Earth's core magnetic field coefficients,  $g_{n_k}^{e,m}$  and  $h_{n_k}^{e,m}$  are the Gaussian electric current coefficients.

The polynomial method was also selected for comparison because it is usually more accurate for regional geomagnetic modeling<sup>45,58</sup>. The Taylor polynomial normal field model equation is given below<sup>14,45,58</sup>:

Epoch	Total repeat stations	Number of sample locations (repeat stations)	Number of validation locations
1985	59	50	8
1990	53	45	8
1995	53	45	8
2000	53	45	8
2005	54	46	8
2010	53	45	8
2015	68	58	8

**Table 1.** The number of sample and validation data sets used in each epoch 1985–2015.

$$C(\theta, \varphi) = \sum_{n=0}^N \sum_{m=0}^n A_{nm} (\theta - \theta_0)^{n-m} (\varphi - \varphi_0)^m. \quad (5)$$

where  $C$  is the geomagnetic field component,  $A_{nm}$  is the Taylor polynomial coefficient,  $\theta$  is the latitude,  $\phi$  is the longitude, and  $\theta_0$  and  $\phi_0$  are the latitude and longitude from the center of the region.

The data from the repeat stations was cut by the diurnal variation from BMKG observatories. The data was then checked and refined by applying a secular variation. After the data has the same time epoch (e.g., epoch 2015.0), geomagnetic data (excluding the 8 validation locations) were used to estimate the geomagnetic values of the validation locations for the CC and KED methods.

For the OK, SCHA, and polynomial methods, the geomagnetic data for each component  $X$ ,  $Y$ , and  $Z$  were decremented with the DGRF data to calculate the geomagnetic field anomaly. The short spectrum of the SCHA method was not suitable to model the main geomagnetic field<sup>38</sup>. Mathematically, the procedure was formulated in Eqs. (6), (7), and (8):

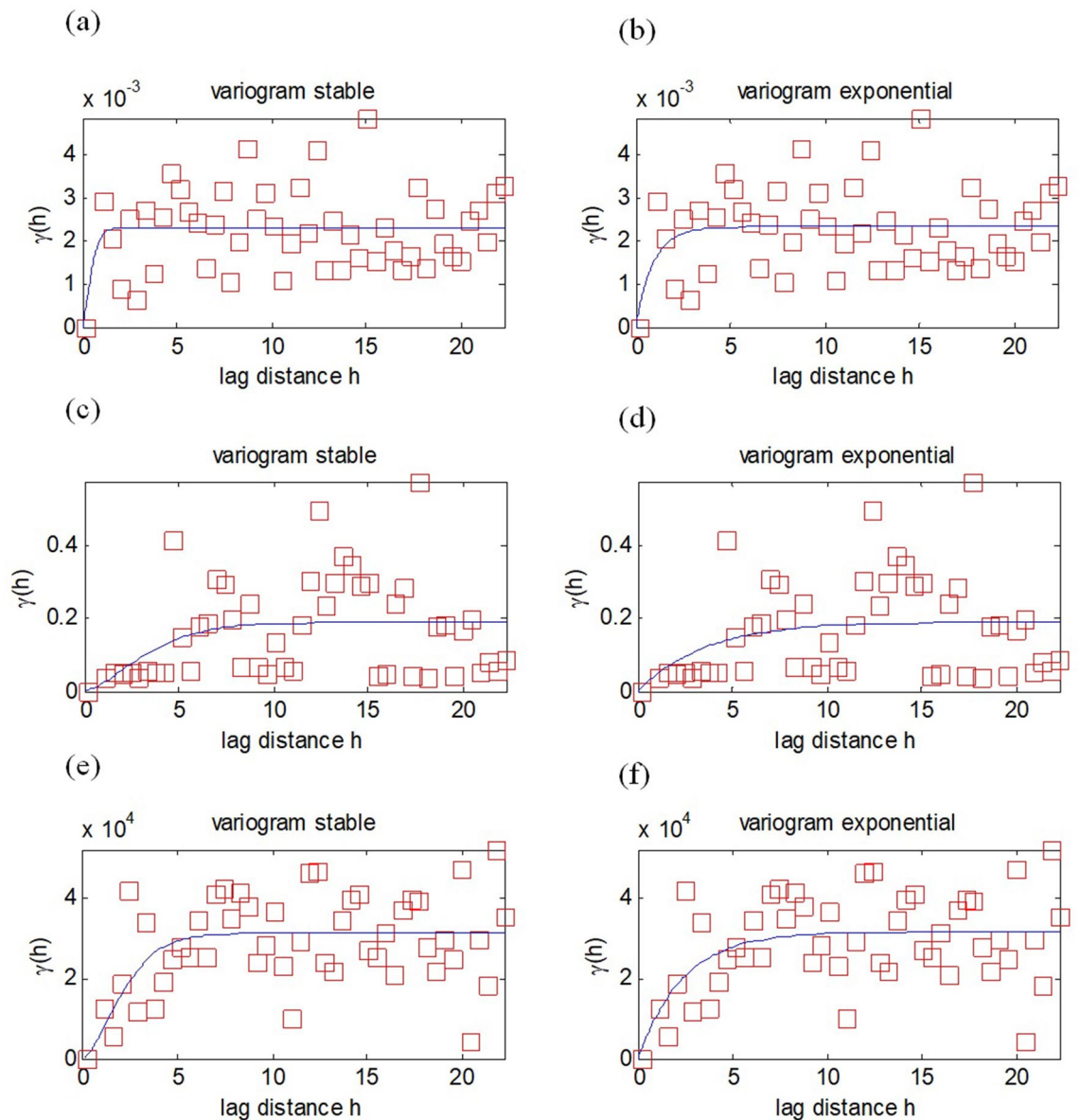
$$\Delta X = X - \text{DGRF}_X \quad (6)$$

$$\Delta Y = Y - \text{DGRF}_Y \quad (7)$$

$$\Delta Z = Z - \text{DGRF}_Z \quad (8)$$

where  $\Delta X$ ,  $\Delta Y$ , and  $\Delta Z$  are the geomagnetic field anomaly of each component and  $\text{DGRF}_X$ ,  $\text{DGRF}_Y$ , and  $\text{DGRF}_Z$  are the main geomagnetic definitive data obtained from the IGRF-13 model for each component.

The number of sample data sets and locations used for each epoch's calculation is shown in Table 1. Equation (1) was used for the OK method, Eq. (2) was used for the CC method, Eq. (3) was used for the KED method, Eq. (4) was used for the SCHA method, and Eq. (5) was used for the polynomial method with  $\theta_0 = 2^\circ$  and  $\phi_0 = 117.5^\circ$ . The OK method used stable and exponential variograms. The variogram parameters used in the OK model are range 3.41, sill 0.0023, and nugget 0.0001 for the declination component; range 10.2, sill 0.19, and nugget 0.01 for the inclination component; and range 5.95, sill 31,216, and nugget 1 for total intensity (Fig. 5). The selection of the variogram affects the estimation results of the kriging method<sup>26</sup>. The CC and KED used a spherical variogram. For the SCHA method, an 8th-order truncation was used for the geomagnetic anomaly, because the RMSE of the  $\Delta X$ ,  $\Delta Y$ , and  $\Delta Z$  components for the 8th order was smaller and stable. For the polynomial method, the 5th order was used because it produced a small RMSE value. We then analyzed the estimation data to determine the root mean square error (RMSE). We then used RMSE to compare the accuracy of the geostatistical method in modeling the regional geomagnetic field with the SCHA and polynomial methods.



**Figure 5.** Semivariogram of geomagnetic data from epoch 2010 for the OK method. Clustered overlapping experimental variograms as characteristic of the clustered distributed data pattern, resulting in identical theoretical variogram parameters between stable and exponential. **(a)** stable variogram of declination comp., **(b)** exponential variogram of declination comp., **(c)** stable variogram of inclination comp., **(d)** exponential variogram of inclination comp., **(e)** stable variogram of total intensity comp., and **(f)** exponential variogram of total intensity comp.

### Data availability

The geomagnetic data for 1985–2015 used in this study was extracted from BMKG. The Definitive Geomagnetic Reference Field (DGRF) data was generated from IGRF-13 (<https://www.ngdc.noaa.gov/AGA/vmod/igrf.html>). The analysis software tool used was Matlab (License ID: 1,106,171). Nearest-neighbor analysis to determine the type of data sample distribution was calculated using the Z score (<http://ceadserv1.nku.edu/longa/geomed/ppa/doc/NNA/NNA.htm>). The geomagnetic regional model epoch 2015 is available at <http://doi.org/10.5281/zenodo.5094914>.

Received: 26 September 2021; Accepted: 30 March 2022

Published online: 06 May 2022

### References

1. Elsasser, W. M. Hydromagnetic dynamo theory. *Rev. Mod. Phys.* **28**, 135–163 (1956).
2. Brandenburg, A. Hydromagnetic dynamo theory. *Scholarpedia* **2**, 2309 (2007).
3. Cook, A. Edmond halley and the magnetic field of the earth. *Notes Rec. R. Soc. Lond.* **55**(3), 473–490 (2001).



4. Garland, G. D. The contributions of Carl Friedrich Gauss to geomagnetism. *Hist. Math.* **6**, 5–29 (1979).
5. Heppner, J. P. The world magnetic survey. *Space Sci. Rev.* **2**, 315–354 (1963).
6. Zmuda, A. J. International geomagnetic reference field 1965.0. *J. Geomagn. Geoelectr.* **21**, 569–571 (1969).
7. Macmillan, S. & Finlay, C. The international geomagnetic reference field. *Geomagn. Observ. Models* [https://doi.org/10.1007/978-90-481-9858-0\\_10](https://doi.org/10.1007/978-90-481-9858-0_10) (2011).
8. Manda, M. & Macmillan, S. International geomagnetic reference field the eighth generation. *Earth Planets Sp.* **52**, 1119–1124 (2000).
9. Maus, S. *et al.* The 10th generation international geomagnetic reference field. *Phys. Earth Planet. Inter.* **151**, 320–322 (2005).
10. Golovkov, V. P., Bondar, T. N. & Burdelnaya, I. A. Spatial-temporal modeling of the geomagnetic field for 1980–2000 period and a candidate IGRF secular-variation model for 2000–2005. *Earth Planets Sp.* **52**, 1125–1135 (2000).
11. Korte, M., Korte, M. & Lesur, V. Repeat station data compared to a global geomagnetic field model. *Ann. Geophys.* **55**, 1101–1111 (2012).
12. Wang, T. The analysis of the IGRF error in the China continent. *Chin. J. Geophys.* **46**, 236–241 (2003).
13. Macmillan, S. Earth's magnetic field. *Geophys. Monogr. Ser.* **145**, 299–320 (2004).
14. Manda, M. & Purucker, M. Observing, modeling, and interpreting magnetic fields of the solid earth. *Surv. Geophys.* **26**, 415–459 (2005).
15. Newitt, L. R., Barton, C. E. & Bitterly, J. *Guide for magnetic repeat station surveys*. (International Association of Geomagnetism and Aeronomy, 1996).
16. Barraclough, D. R. & Santis, A. D. Repeat station activities. in *Geomagnetic Observations and Models* (ed. M. Manda, M. K.) 45–56 (Springer Netherlands, 2011). doi:<https://doi.org/10.1007/978-90-481-9858-0>
17. Newitt, L. R., Barton, C. E., Bitterly, J. & International Association of Geomagnetism and Aeronomy. Working Group V-8: analysis of the global and regional geomagnetic field and its secular variation. *Guide Magn. Rep. Stat. Surv.* **112** (1996).
18. Düzgüç, Z., Baydemir, N. & Malin, S. R. C. Rectangular polynomial analysis of the regional geomagnetic field. *Geophys. J. Int.* **128**, 737–743 (1997).
19. Allredge, L. R. Rectangular harmonic analysis applied to the geomagnetic field. *J. Geophys. Res.* **86**, 3021–3026 (1981).
20. Hall, C. A. & Meyer, W. W. Optimal error bounds for cubic spline interpolation. *J. Approx. Theory* **16**, 105–122 (1976).
21. Wahba, G. *Spline Models for Observational Data*. (Society for industrial and applied mathematics, 1990). doi:<https://doi.org/10.1137/1.9781611970128>
22. Haines, G. V. Spherical cap harmonic analysis. *J. Geophys. Res.* **90**, 2583–2591 (1985).
23. Thébaud, E., Schott, J. J., Manda, M. & Hoffbeck, J. P. A new proposal for spherical cap harmonic modelling. *Geophys. J. Int.* **159**, 83–103 (2004).
24. Thébaud, E. A proposal for regional modelling at the Earth's surface, R-SCHA2D. *Geophys. J. Int.* **174**, 118–134 (2008).
25. Torta, J. M. Modelling by spherical cap harmonic analysis: a literature review. *Surv. Geophys.* **41**, 201–247 (2020).
26. Webster, R. & Oliver, M. A. *Geostatistics for environmental scientists*. (Wiley, 2008). doi:<https://doi.org/10.1002/9780470517277>
27. Krige, D. G. A statistical approach to some basic mine valuation problems on the Witwatersrand. *J. Chem. Metall. Min. Soc. S. Afr.* **52**, 119–139 (1951).
28. Goovaerts, P. *Geostatistics for natural resources and evaluation*. (Oxford University Press, 1997).
29. Huysmans, M. & Dassargues, A. Application of multiple-point geostatistics on modelling groundwater flow and transport in a cross-bedded aquifer (Belgium). *Hydrogeol. J.* **17**, 1901–1911 (2009).
30. Poon, D. C., McCormack, M. & Thimm, H. F. The application of fractal geostatistics to oil estimates. *J. Can. Pet. Technol.* **32**(10), 24–27 (1993).
31. Xi, Z. & Morgan, E. Combining decline-curve analysis and geostatistics to forecast gas production in the Marcellus shale. *SPE Reserv. Eval. Eng.* **22**, 1562–1574 (2019).
32. Isaaks, E. H. & Srivastava, R. M. *Applied geostatistics*. (Oxford University Press, 1989).
33. Goovaerts, P. Geostatistics for natural resource evaluation. in *Technometrics* **42**, (1997).
34. Wackernagel, H. Ordinary Kriging. in *Multivariate geostatistic* 74–81 (Springer, 1995). doi:[https://doi.org/10.1007/978-3-662-03098-1\\_11](https://doi.org/10.1007/978-3-662-03098-1_11)
35. Montero, J.-M., Gema, F.-A. & Mateu, J. *Spatial and Spatio-Temporal Geostatistical Modeling and Kriging*. (Wiley, 2015).
36. Hengl, T., Heuvelink, G. & Stein, A. Comparison of kriging with external drift and regression-kriging. *ITC Techn. Note Enschede Netherlands* [https://doi.org/10.1016/S0016-7061\(00\)00042-2](https://doi.org/10.1016/S0016-7061(00)00042-2) (2003).
37. Rivoirard, J. Which models for collocated cokriging?. *Math. Geol.* **33**, 117–128 (2001).
38. Thébaud, E. & Gaya-Piqué, L. Applied comparisons between SCHA and R-SCHA regional modeling techniques. *Geochem. Geophys. Geosyst.* **9**, Q07005. <https://doi.org/10.1029/2008GC001953> (2008).
39. Schott, J.-J. & Thébaud, E. Modelling the earth's magnetic field from global to regional scales. in *Geomagnetic Observations and Models* 229–264 (Springer Netherlands, 2011). doi:[https://doi.org/10.1007/978-90-481-9858-0\\_9](https://doi.org/10.1007/978-90-481-9858-0_9)
40. Chen, D.-X., Liu, D.-Z., Zeng, X.-N., Meng, L. & Yang, X.-J. Application and improvement of spatial temporal Kriging in geomagnetic field interpolation. *Acta Geophys. Sin.* **59**, 1743–1752 (2016).
41. Bonito, A., DeVore, R., Guignard, D., Jantsch, P. & Petrova, G. Polynomial approximation of anisotropic analytic functions of several variables. *Constr. Approx.* **53**, 319–348 (2021).
42. Korte, M. & Thébaud, E. Geomagnetic repeat station crustal biases and vectorial anomaly maps for Germany. *Geophys. J. Int.* **170**, 81–92 (2007).
43. Zhao, S., Zhou, Y., Wang, M., Xin, X. & Chen, F. Thickness, porosity, and permeability prediction: comparative studies and application of the geostatistical modeling in an Oil field. *Environ. Syst. Res.* **3**, 7 (2014).
44. Han, F., Zhang, H., Guo, Q., Wei, K. & Shang, Z. An integrated method for seismic velocity modeling based on collocated cokriging. *J. Geophys. Eng.* **15**, 1389–1398 (2018).
45. Gu, Z. *et al.* Geomagnetic survey and geomagnetic model research in China. *Earth Planets Sp.* **58**, 741–750 (2006).
46. Ryan, T. P. *Modern Regression Methods*. (Wiley, 2008).
47. Abedi, M., Asghari, O. & Norouzi, G.-H. Collocated cokriging of iron deposit based on a model of magnetic susceptibility: a case study in Morvarid mine, Iran. *Arab. J. Geosci.* **8**, 2179–2189 (2015).
48. Madani, N. Multi-collocated cokriging: an application to grade estimation in the mining industry. in *39th International symposium on Application of Computers and Operations Research in the Mineral Industry, APCOM 2019* (eds. C., M. *et al.*) 158–167 (CRC Press/Balkema, 2019). doi:<https://doi.org/10.1201/9780429320774-18>
49. Verbanac, G. On regional modeling of the main geomagnetic field. *Geofizika* **24**, 1–27 (2007).
50. Korté, P. B. & Korte, M. Morphology of the southern African geomagnetic field derived from observatory and repeat station survey observations: 2005–2014. *Earth Planets Sp.* **68**, 23 (2016).
51. Rasson, J. L. & Delipetrov, T. *Geomagnetics for aeronautical safety: a case study in and around the Balkans*. NATO Security through Science Series C: Environmental Security (Springer, 2006).
52. Loubser, L. & Newitt, L. *Guide for Calibrating a compass swing base*. (IAGA, 2009).
53. Talarn, À., Pavón-Carrasco, F. J., Torta, J. M. & Catalán, M. Evaluation of using R-SCHA to simultaneously model main field and secular variation multilevel geomagnetic data for the North Atlantic. *Phys. Earth Planet. Inter.* **263**, 55–68 (2017).

54. Căţeanu, M. & Ciubotaru, A. Accuracy of ground surface interpolation from airborne laser scanning (ALS) data in dense forest cover. *ISPRS Int. J. Geo-Inform.* **9**, 224 (2020).
55. Massimi, L., Ristorini, M., Astolfi, M. L., Perrino, C. & Canepari, S. High resolution spatial mapping of element concentrations in PM10: A powerful tool for localization of emission sources. *Atmos. Res.* **244**, 105060 (2020).
56. Nistor, M. M. *et al.* Investigation of groundwater table distribution using borehole piezometer data interpolation: case study of Singapore. *Eng. Geol.* **271**, 105590 (2020).
57. Sunkari, E. D., Abu, M., Zango, M. S. & Lomoro Wani, A. M. Hydrogeochemical characterization and assessment of groundwater quality in the Kwahu-Bombouaka Group of the Voltaian Supergroup. *Ghana. J. Afr. Earth Sci.* **169**, 103899 (2020).
58. Geese, A., Korte, M., Kotze, P. B. & Lesur, V. Southern African geomagnetic secular variation from 2005 to 2009. *S. Afr. J. Geol.* **114**, 515–524 (2011).

### Acknowledgements

We thank the Meteorological, Climatological, and Geophysical Agency of Indonesia (BMKG) for supporting and granting access to the repeat station data used in this research. We also thank the Indonesian Higher Education Directorate for providing partial funding through the PDD schema.

### Author contributions

M.S. performed the analysis of geomagnetic regional model using SCHA, Taylor polynomial, and geostatistical methods, E.H. the research framework formulation and editing the manuscript, and S.A. contributed to the preparation of the manuscript.

### Funding

This research was supported and funded by BMKG and the Indonesian Higher Education Directorate through the PDD schema with contact number: 2246/UN1/DITLIT/DIT-LIT/PT/2021.

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to E.H.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022