Pengembangan Metode *Geoadditive Small Area Model* Pada *Software R*

Ketut Karang Pradnyadika (221709776, 4SD1)

Dosen Pembimbing: Ika Yuni Wulansari, SST., M.Stat

***Ringkasan*— Metode pengumpulan data yang dilakukan BPS masih berkutat pada sensus dan survei. Namun kenyataannya data yang dikumpulkan memiliki keterbatasan dalam estimasi. Untuk melakukan estimasi level area kecil, solusinya dengan mengestimasi tidak langsung dengan metode *Small Area Estimation* (SAE). Metode SAE memerlukan variabel penyerta yang tersedia hingga level yang hendak diestimasi. Model SAE merupakan model campuran dengan varians dalam subpopulasi dapat dijelaskan dengan efek tetap dan acak. Asumsi yang harus terpenuhi dalam estimasi model *mixed linear* adalah normalitas. Namun banyak data yang ditemukan tidak berdistribusi normal, sehingga diperlukan metode alternatif yang dapat digunakan sebagai solusi dalam estimasi statistik area kecil, yaitu *Geoadditive Model*. Dalam penerapannya, metode ini belum dikembangkan secara maksimal, sehingga peneliti ingin mengembangkan *Geoadditive Small Area Model* pada *software* R dalam melakukan estimasi pada area kecil. Pengembangan metode ini menghasilkan suatu *package R* dengan nama ‘*geoSAE*’, yang dapat digunakan sebagai analisis SAE dengan metode *Geoadditive*, sehingga dapat digunakan oleh pengguna pada berbagai data yang memiliki keterbatasan asumsi statistik.**

***Kata Kunci*— *Small Area Estimation, Geoadditive model, Software R***

1. Latar Belakang

Selama ini, metode pengumpulan data yang dilakukan BPS masih berkutat pada sensus dan survei. Sensus mengkaji seluruh populasi sehingga data yang disajikan dalam sensus sampai satuan wilayah terkecil. Namun, pengumpulan data dengan sensus memiliki keterbatasan pada variabelnya, dan juga membutuhkan waktu, tenaga, dan biaya yang besar. Di sisi lain, kekurangan tersebut dapat diatasi dengan survei, dimana survei dapat mencakup banyak variabel serta lebih efisien dari segi waktu, tenaga, dan biaya. Meskipun demikian, survei memiliki keterbatasan dalam estimasi yang disebabkan permasalahan kecukupan sampel. Apabila melakukan estimasi untuk tingkat administrasi di bawah provinsi (area kecil) akan memiliki nilai standard error yang besar sehingga statistik yang dihasilkan tidak dapat diandalkan. Maka dari itu, dibutuhkan suatu pendugaan dengan metode tidak langsung yang dapat memberikan hasil estimasi parameter pada level unit kecil dengan standar error yang kecil dan jumlah sampel yang terbatas. Salah satu metode tidak langsung dapat dijadikan solusi adalah *Small Area Estimation* (SAE), yaitu metode pendugaan tidak langsung yang “meminjam kekuatan” dari informasi wilayah yang bersesuaian atau informasi variabel penyerta yang mempunyai hubungan kuat dengan variabel amatan [1].

Model SAE merupakan model *mixed linear* dengan keragaman dalam subpopulasi dapat dijelaskan dengan efek tetap dan acak [2]. Efek tetap diterangkan oleh keragaman dari peubah penyerta, sedangkan efek acak diterangkan oleh keragaman khusus yang tidak dapat diterangkan oleh peubah penyerta. Peneliti pertama yang mengembangkan statistik area kecil berdasarkan model campuran linear dengan *Empirical Best Linear Unbiased Prediction* (EBLUP) sebagai metode solusi adalah Fay dan Herriot (1979) [3]. Metode EBLUP membutuhkan asumsi normalitas, namun pada kenyataannya banyak ditemukan bahwa data tidak berdistribusi normal.

Dengan terlanggarnya asumsi tersebut, maka dibutuhkan suatu pendekatan semi ataupun nonparametrik. Salah satu metode alternatif yang dapat digunakan sebagai solusi dalam estimasi statistik area kecil dengan pendekatan semiparametrik adalah *Geoadditive Model*. Metode ini merupakan metode penggabungan model *kriging* dan *additive* yang disajikan dalam bentuk model *mixed linear*. Model ini memiliki fleksibilitas dalam menentukan bentuk hubungan antara respon dengan informasi spasial, karena merupakan fungsi pemulusan. Karena teknik pendugaan area kecil “meminjam kekuatan” di area tetangga sehingga SAE dikembangkan dengan mengintegrasikan informasi spasial ke dalam model, yang dikenal sebagai SAE spasial (spasial Fay-Herriot). Baik model Fay-Herriot maupun model spasial Fay-Herriot memerlukan pemenuhan asumsi linearitas kovariat serta normalitas distribusi respons yang terkadang dilanggar, maka *Geoadditive Model* menawarkan penanganan tersebut dengan menggunakan fungsi *smoothing* [2].

Dengan perkembangan teknologi saat ini, terdapat beberapa perangkat lunak *(software)* yang dapat digunakan untuk analisis SAE, salah satunya adalah R. Pada *software* ini, belum terdapat *package* dengan metode *Geoadditive Small Area*, sehingga pengguna harus mengetikkan kode *script* dalam melakukan estimasi. Hal ini akan berdampak pada ketidak efisiensinya pengolahan data dan pada akhirnya akan menguras banyak waktu dan tenaga.

Berdasarkan permasalahan di atas, pada penelitian ini akan dilakukan pengembangan metode *Geoadditive Small Area Model* pada *software R*.

1. Tujuan Penelitian

Berdasarkan latar belakang yang telah dijelaskan, maka tujuan dari penelitian ini adalah sebagai berikut.

1. Melakukan simulasi data dengan *script R* *Geoadditive Small Area Model* yang dibangun.
2. Membangun sebuah *package* dengan metode *Geoadditive Small Area Model* pada *software* *R.*
3. Menguji performa dari pembangunan *package R* dengan metode *Geoadditive Small Area Model* yang telah dibangun menggunakan data publikasi dari BPS sebagai studi kasus.
4. Penelitian Terkait

Kammann dan Wand (2003) merupakan peneliti pertama yang mengemukakan metode *geoadditive* [5], dimana studi kasus yang dilakukan tentang kesehatan reproduksi di Amerika Serikat. Hasilnya menunjukkan hubungan non-linier dari variabel respon sehingga menggabungkan kriging dengan model aditif. penggabungan model tersebut disebut dengan *Geoadditive Model*. Terdapat beberapa peneliti yang sudah melakukan penelitian dengan metode *Geoadditive*, seperti Kandala, et.al (2009) untuk memodelkan kekurangan gizi anak-anak [8], dan Nkurunziza, et. al (2011) untuk memodelkan penyakit malaria [9].

Djuraidah dan Aunuddin (2006) melakukan penelitian menggunakan metode *kriging* dan *spline*-2 pada data pencemaran ozon [10]. *Kriging* dan *spline*-2 dinyatakan sebagai kombinasi linear dari fungsi basis radial, sehingga kedua model tersebut dapat didekati dengan model linear campuran. Hasilnya menunjukkan prediksi spasial dengan *kriging* menghasilkan kurva permukaan yang kasar, sedangkan *spline*-2 menghasilkan kurva permukaan yang halus.

Salvati, et. al (2008) melakukan estimasi area kecil menggunakan *direct estimator* berbasis model nonparametrik [11]. Konsep estimasi langsung berbasis model digunakan untuk mengembangkan pendekatan nonparametrik alternatif untuk estimasi rata-rata area kecil. Kurnia (2009) menggunakan pendekatan SAE dengan nonparametrik, dimana menggabungkan konsep *spline* dan model SAE karena dapat dipandang sebagai model dengan efek acak [4].

Kemudian pada tahun 2010, Bocci melakukan penelitian tentang konsumsi pengeluaran di Albania dengan menggunakan *Geoadditive Small Area Model*, yang hasilnya menunjukkan bahwa model inidapat diterapkan untuk memperkirakan rata-rata tingkat distrik dari pengeluaran konsumsi per kapita log rumah tangga untuk Republik Albania [6].

Di Indonesia, terdapat beberapa permasalahan yang dapat diukur dengan model *Geoadditive Small Area Model*, seperti Ardiansyah, et. al. (2018) pernah menggunakan model ini untuk mengestimasi produktivitas padi di Kabupaten Seruyan [7], dan Pusponegoro, et. al. (2019) menggunakan metode ini untuk menghitung kemiskinan di Provinsi Bangka Belitung [2].

TABEL I

TABEL LITERATUR

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *No* | *Judul* | *Penulis, Publikasi* | *Tertulis* | *Komentar* |
| 1 | *Small Area Estimation, 2nd Edition* | J.N.K. Rao and Isabel Molina, New Jersey: John Wiley and Sons Inc. | *In the context of small area estimation, direct estimators lead to unacceptably large standard errors for areas with unduly small sample sizes; in fact, no sample*  *units may be selected from some small domains,* | Melakukan estimasi langsung dengan keterbatasan sampel akan menyebabkan besarnya nilai standar error |
| *Therefore, it is often necessary to use indirect estimates that “borrow strength” by using values of the variables of interest from related areas, thus increasing the “effective” sample size,* | Metode pendugaan tidak langsung yang “meminjam kekuatan” dari informasi wilayah yang bersesuaian atau informasi variabel penyerta yang mempunyai hubungan kuat dengan variabel amatan |
| 2 | *Small Area Estimation Using A Nonparametric Model-Based Direct Estimator* | Salvati, N., Chandra, H., Ranalli, M. G. and Chambers, R. (2008). Submitted for publication. | *Nonparametric regression is widely used as a method of characterizing a non-linear relationship between a variable of interest and a set of covariates.* | Pendekatan nonparametrik dapat digunakan dapat melakukan pendugaan. |
| 3 | *Geoadditive Models in Small Area Estimation of Poverty* | Novi Hidayat Pusponegoro, Anik Djuraidah, Anwar Fitrianto, I Made Sumertajaya, Journal of Data Science and Its Applications, VOL. 2, No. 1, PP.11-18, January 2019 | *Since, the Fay-Herriot model and the spatial Fay-Herriot model require the fulfillment of covariate linearity assumption as well as the normality of the response distribution that is sometimes violated and the Geoadditive model offers those handling using the smoothing function,* | Banyak kasus yang melanggar asumsi pada model Fay-Herriot dan model spasial Fay-Herriot, maka *Geoadditive model* menawarkan penanganan tersebut dengan menggunakan fungsi *smoothing* |
| 4 | *The Best Predictions for Empirical Logarithm Transformation Models in Small Area Estimation with Application on SUSENAS Data* | Kurnia, A. (2009), Doctoral dissertation-Ph. D dissertation, Bogor Agricultural University | *The lognormal empirical best prediction (EBP) model which are suggested in this research give better results than standard SAE models in terms of the smallest relative root mean square error (RRMSE).* | Penemuan metode baru yaitu dengan transformasi logaritma apabila asumsi normalitas tidak terpenuhi pada estimasi *mixed linear model* |
| 5 | *Geoadditive Models for Data with Spatial Information* | Bocci, C. (2010), Doctoral dissertation-Ph. D dissertation, Italy: University of Florence | *As both the SAE models and the geoadditive models are formulated as linear mixed models, it seems an obvious choice to merge the two models in a geoadditive SAE model to exploit the spatial information and produce estimates at small area level.* | Model SAE dan model *geoaditive* diformulasikan sebagai model campuran linier, tampaknya merupakan pilihan yang jelas untuk menggabungkan kedua model |

1. Metode Penelitian
   1. Metode Pengumpulan Data

Data yang digunakan dalam penelitian terdiri dari data simulasi berupa data bangkitan, dan data studi kasus yang bersumber dari referensi penelitian Ardiansyah, et. al. (2018) [7] untuk menghitung produktivitas padi.

Data simulasi yang digunakan merupakan pembangkitan data yang mengacu kepada penelitian Benavent dan Morales (2016) [12], Ardiansyah, et. al. (2021) [13]. Berikut adalah tahapan analisis data simulasi yang digunakan pada penelitian ini.

1. Menentukan jumlah unit dan area kecil
2. Membangkitkan yang merupakan variabel penyerta dengan mengikuti distribusi binomial dan uniform, dengan rincian:
3. Membangkitkan dengan distribusi normal dengan rata-rata 0, dan *varians random effect*
4. Membangkitkan dengan distribusi normal dengan rata-rata 0, dan *varians sampling error*
5. Menentukan nilai koefisien
6. Menghitung nilai y sebagai model based yaitu .
7. Menentukan jumlah *knot* yang digunakan dalam beberapa skenario
8. Menghitung matriks Z dari spline-2 berdasarkan fungsi basis radial, dengan formula seperti berikut:

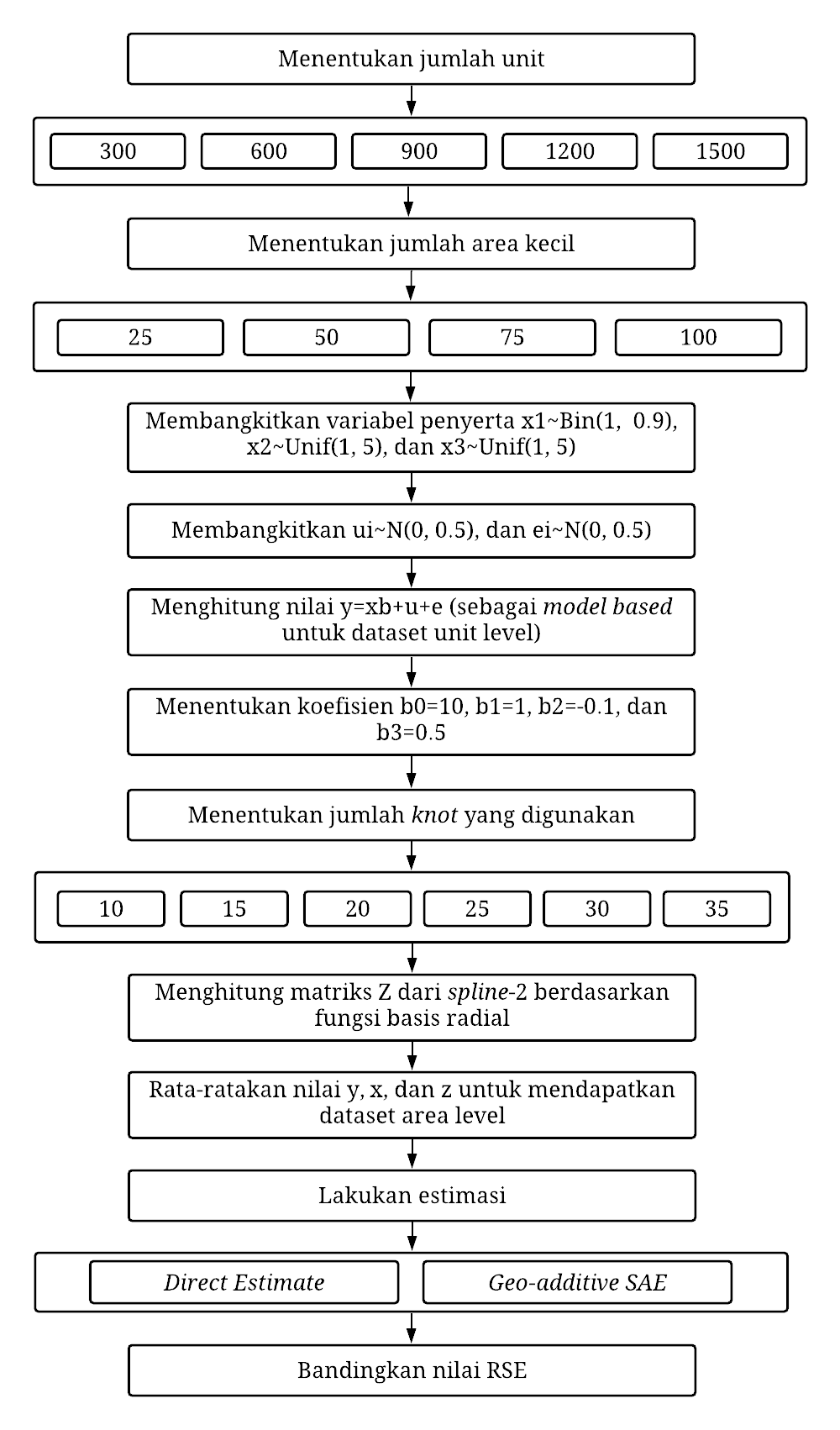
dengan

1. Merata-ratakan nilai dan yang diperoleh
2. Menyatakan bentuk area kecil dengan mengikuti suatu fungsi *Geoadditive Small Area Model* , kemudian lakukan pendugaan nilai parameter
3. Menghitung nilai RSE untuk setiap area ke-i menggunakan metode *bootstrap* parametrik, dengan langkah seperti berikut:

* Bangkitkan efek acak dari
* Bangkitkan efek acak dari
* Bangkitkan efek acak dari
* Bandingkan antara target parameter *bootstrap* dengan contoh *bootstrap* yang dihitung berdasarkan *Geoadditive Small Area Model*
* Hitung nilai RMSE *bootstrap* dengan formula seperti berikut:

1. Bandingkan nilai estimasi dan RSE yang diperoleh dengan metode *Geoadditive Small Area Model,* dan estimasi langsung.

Dari proses tahapan tersebut, maka dapat dibuat suatu diagram *flowchart* yang menggambarkan alur pengolahan data dengan simulasi data diatas menggunakan metode *Geoadditive Small Area Model* sebagai berikut.



Gambar 1. *Flowchart* tahapan simulasi data

* 1. Metode Analisis
     1. *Geoadditive Small Area Model*

*Geo-additive Small Area Model* merupakan penggabungan antara model *Geo-additive* dan model SAE, yaitu dengan menambahkan informasi geospasial ke dalam model SAE. Berikut persamaan gabungan antara *Geo-additive Model* dan SAE *(Geo-additive Small Area Model)* dengan dua komponen efek acak [7].

|  |  |  |
| --- | --- | --- |
|  |  | ( 1 ) |

dengan , , ,, dan

Matriks peubah penyerta menjadi , adalah vektor koefisien efek tetap berukuran *(p+1)x1*, adalah matriks berukuran *nxK* dari spline-2 berdasarkan fungsi dan . Efek acak adalah vektor koefisien spline-2 (efek acak kedua) berukuran *Kx1*. Efek acak adalah vektor efek acak area kecil berukuran *mx1*, adalah galat acak level unit, dan matriks dengan jika pengamatan berada dalam area kecil ke dan untuk lainnya.

* + 1. *Empirical Best Linear Unbiased Prediction (EBLUP)*

Metode *Empirical Best Linear Unbiased Prediction* (EBLUP) merupakan salah satu pendekatan yang sering digunakan pada pendugaan *small area estimation.* EBLUP dilatarbelakangi oleh ketidakmampuan BLUP (*Best Linear Unbiased Predi*ctor) dalam melakukan pendugaan komponen varians yang tidak diketahui. Pada EBLUP, komponen varians tidak diketahui, oleh karena itu diperlukan estimasi varians dengan menggunakan metode *maximum likelihood* (ML), metode *restricted maximum likelihood* (REML), atau metode *Prasad Rao*. Persamaan EBLUP adalah sebagai berikut.

|  |  |  |
| --- | --- | --- |
|  |  | ( 2 ) |

dimana:

komponen penyeimbang dari estimasi EBLUP;

= estimasi koefisien regresi;

= estimasi langsung area ke-*i* yang dihasilkan dari survei;

= vektor variabel penyerta area ke-*i*

* + 1. MSE *Bootstrap* Parametrik

Metode *bootstrap* parametrik mengasumsikan bahwa ketiga efek acak menyebar normal dengan rataan dan ragamnya adalah , dan dengan dan . Prosedur *bootstrap* parametrik dilakukan dengan membangkitkan data contoh *bootstrap* dan sebagai . Efek acak dibangkitkan dari , dibangkitkan dari , dan dibangkitkan dari . Misalkan menjadi versi *bootstrap* dari target parameter kemudian versi *bootstrap* dengan  **,** dan dihitung dengan cara yang sama dengan **,** dan tetapi dengan menggunakan data contoh *bootstrap*.

𝑀𝑆𝐸𝑏𝑜𝑜𝑡𝑠𝑡𝑟𝑎𝑝 diperkirakan dengan simulasi *Monte Carlo* seperti langkah yang telah disebutkan sebelumnya sebanyak B kali sehingga diperoleh yang merupakan nilai bootstrap sebenarnya dengan dan juga diperoleh dari dugaan *Geo-Additive Small Area Model* EBLUP dengan dan adalah ulangan ke b dari . Secara teori, . RMSE bootstrap untuk setiap area ke- 𝑖 diperoleh berdasarkan rumus sebagai berikut:

|  |  |  |
| --- | --- | --- |
|  |  | ( 3 ) |

dengan .

* 1. Tahapan Penelitian

Tahapan pengembangan metode Geoadditive *Small Area Model* pada penelitian ini menggunakan metode *design science research*. Metode ini terdiri dari lima tahapan, yaitu sebagai berikut.

1. *Awareness of problem*

Tahap ini dilakukan untuk mengidentifikasi permasalahan terkait *Geoadditive Small Area Model.*

1. *Suggestion*

Tahap ini mendefinisikan solusi dari permasalahan yang telah diidentifikasi pada tahap sebelumnya. Dalam pencarian solusi terhadap masalah yang dihadapi pada penelitian ini, dilakukan dengan menggunakan studi literatur.

1. *Development*

Tahapan ini merupakan proses perancangan dan pembangunan program untuk menjawab permasalahan berdasarkan solusi dan saran yang telah diusulkan di tahapan sebelumnya.

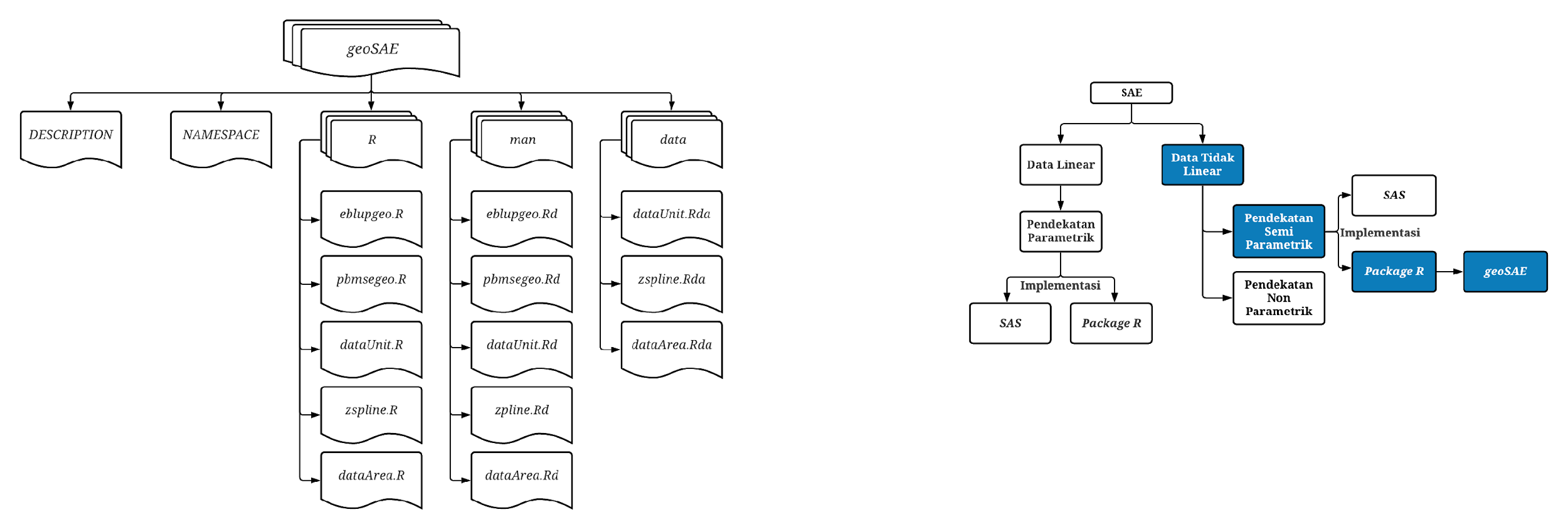
1. *Evaluation*

Dalam tahap evaluasi ini dilakukan pengujian terkait dengan ketepatan dari program yang telah dibangun. Evaluasi yang diterapkan berupa uji validitas yang digunakan untuk menilai kebenaran dari performa algoritma dan program yang telah dibuat.

1. *Conclusion*

Pada tahap ini dilakukan penarikan kesimpulan serta saran dari seluruh tahapan penelitian yang telah dilaksanakan.

1. Kerangka Pikir



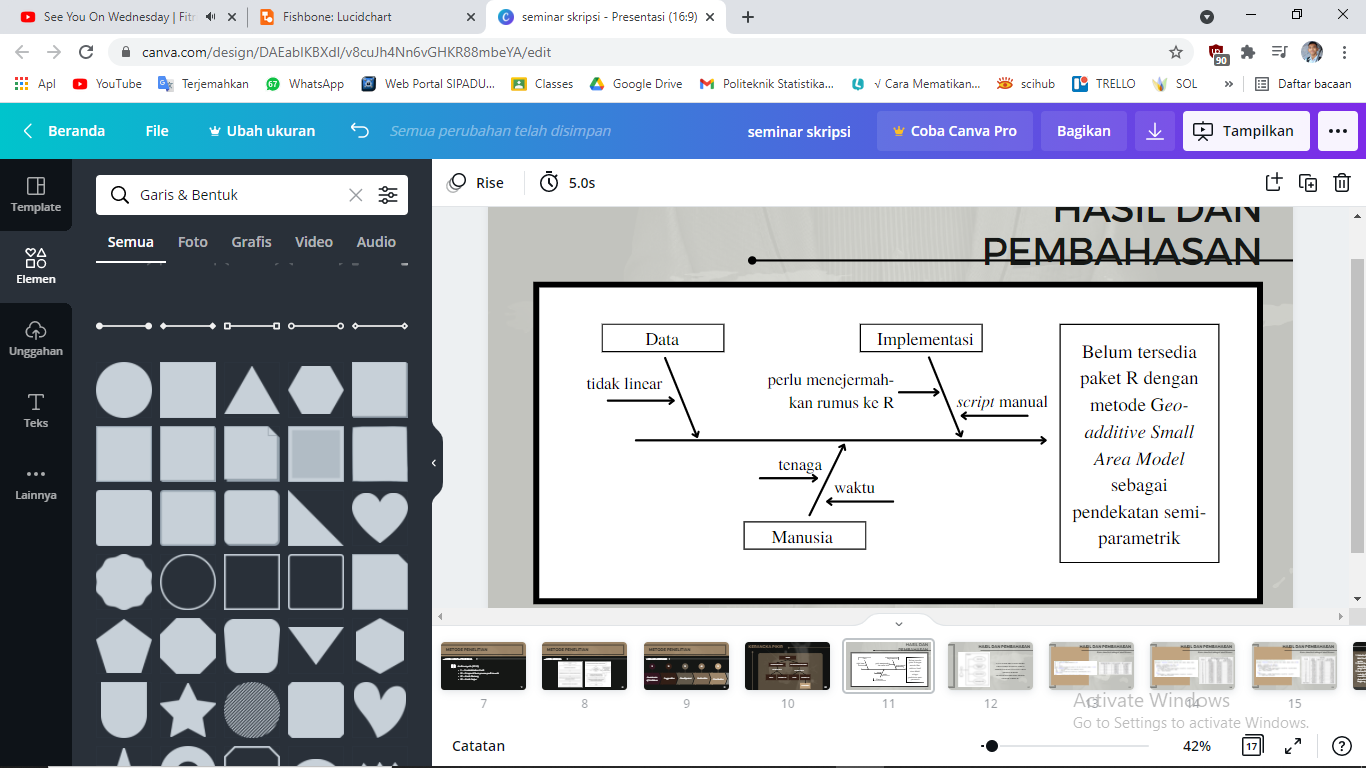
Gambar 2. Kerangka pikir penelitian

Telah banyak peneliti yang menggunakan analisis SAE untuk berbagai studi kasus dalam estimasi parameter pada area kecil seiring dengan perkembangan teknologi. *Software* yang sering digunakan peneliti untuk analisis SAE adalah R.

Secara implementasinya, pendekatan parametrik sudah diimplementasikan dalam bentuk SAS maupun *package* R, sedangkan pendekatan semi ataupun nonparametrik belum dikembangkan secara maksimal. Ketika pengguna melakukan pengolahan SAE pada R dengan *Geoadditive Small Area Model*, pengguna harus menuliskan kode *script* secara manual. Sehingga hal ini menyebabkan ketidakefisienan dari segi waktu dan tenaga. Maka dari itu, peneliti ingin mengembangkan metode SAE dengan pendekatan semiparametrik, yaitu *Geoadditive Small Area Model* menjadi dalam bentuk *package* *R*, yang nantinya akan diberi nama ‘*geoSAE*’

1. Hasil dan Pembahasan
   1. *Awareness of problems*

Dalam penelitian ini, proses mengidentifikasi masalah dilakukan dengan studi literatur dan wawancara dengan dosen yang terkait dengan SAE. Hasil yang diperoleh adalah model *Geoadditive Small Area Model* diketahui belum dikembangkan secara maksimal, perlu dilakukan suatu pengembangan agar mendapatkan hasil estimasi dengan akurasi yang tinggi pada area kecil.



Gambar 3. Diagram *fishbone* kebutuhan paket R dengan *Geoadditive Small Area Model*

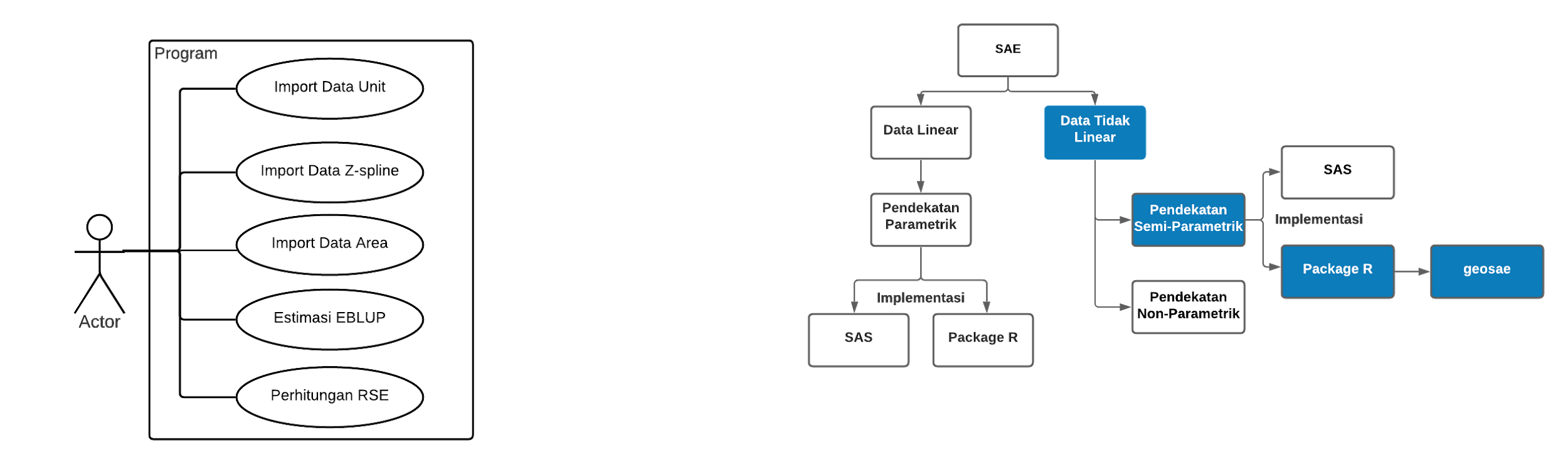
Saat ini, dalam melakukan analisis dengan *Geoadditive Small Area Model* belum adanya *tools* yang mampu memudahkan pengolahan data. Salah satu perangkat lunak atau *software* yang sering digunakan dalam analisis SAE adalah R. Untuk melakukan pengolahan melalui R, pengguna harus menulis kode *script* dengan manual. Hal ini menyebabkan ketidakefisienan dari segi waktu dan tenaga karena harus menulis *script* secara manual pada *software* R dalam melakukan pengolahan. Peneliti juga mengalami kesulitan dalam penyebarluasan hasil *script*.

* 1. *Suggestion*

Solusi yang ditawarkan peneliti untuk permasalahan tersebut pada penelitian ini adalah dengan membangun suatu *package* (paket) pada *software* R dengan metode *Geoadditive Small Area Model* untuk memfasilitasi serta memudahkan pengguna dalam melakukan analisis SAE dengan metode ini. Pada dasarnya, pengembangan metode *Geoadditive Small Area Model* dengan *software* Rpada penelitian ini merujuk kepada *package* “*sae*” dari Molina untuk menghitung nilai EBLUP dan nilai MSE.

* 1. *Development*
     1. Pemodelan

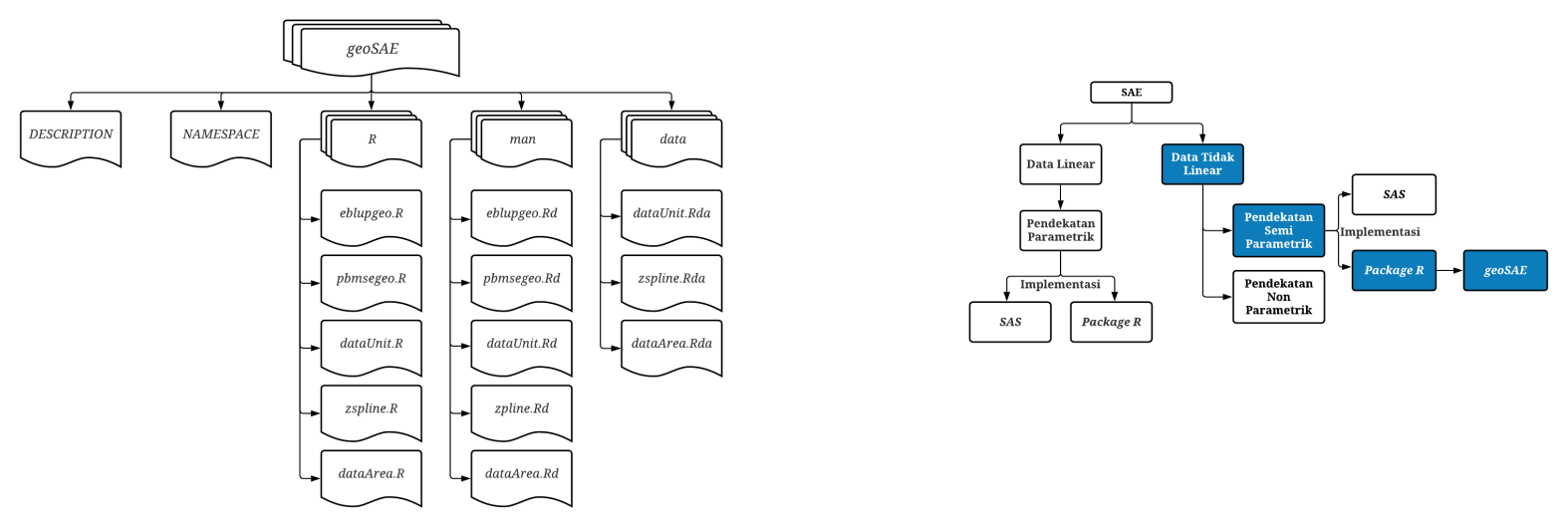
Pada pengembangan metode *Geoadditive Small Area Model*, pemodelan yang dibangun dengan membuat diagram *use case* yang mana hanya terdapat satu aktor yang berinteraksi dengan sistem. Gambar 4 dibawah menjelaskan bagaimana interaksi antara aktor sebagai pengguna dengan sistem, dan program apa saja yang dapat dilakukan sistem.



Gambar 4. Diagram *use case* pengembangan *Geoadditive Small Area Model*

* + 1. Perancangan dan Implementasi

Algoritma dan pemodelan yang telah disusun kemudian diimplementasikan dalam bentuk *package ‘geoSAE’*. Penjabaran mengenai komponen yang terdapat dalam *package geoSAE* adalah sebagai berikut.



Gambar 5. Struktur *package geoSAE*

* + - 1. *File DESCRIPTION*

Merupakan informasi umum tentang *package* yang dibangun. Informasi yang tercakup di dalamnya antara lain: judul *package*, nama penulis dan pengembang, versi, deskripsi, lisensi, dependensi *package* yang digunakan, dan beberapa informasi lainnya.

* + - 1. *File NAMESPACE*

File ini dibuat secara otomatis menggunakan *package* *roxygen2*. NAMESPACE menggambarkan interaksi *package* yang dibuat dengan *package* lain dan pengguna.

File ini memastikan bahwa *package* yang dibangun dan *package* lain tidak saling mengganggu. Interaksi dengan *package* lain digambarkan pada item *‘import’*, sedangkan interaksi dengan pengguna digambarkan pada item *‘export'* yang berisi beberapa *function* yang dapat diakses oleh pengguna.

* + - 1. *Folder R*

Berisi dokumentasi *script* R untuk menyusun *package*, dapat berupa *function*, ataupun dataset.

* + - 1. *Folder man*

File ini di *generate* secara otomatis dari komentar berformat khusus pada *script* R. File ini berisi dokumentasi objek dalam *package* dalam *format.Rd.*

* + - 1. *Folder data*

Merupakan file yang berisikan data yang tersedia pada *package*. Data yang tersedia dalam *package* ini merupakan data yang berhubungan dengan metode yang digunakan.

* 1. *Evaluation*
     1. Data Simulasi

Pada penelitian ini, simulasi data dilakukan dengan membangkitkan variabel penyerta, *random effect*, dan *sampling error* dari distribusi tertentu. Tabel 2 dibawah ini menunjukkan perbandingan rata-rata nilai RSE yang dihasilkan dari data bangkitan dengan *Geoadditive Small Area Model.*

TABEL II

RATA-RATA RSE(%) HASIL SIMULASI DATA BANGKITAN

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Unit*** | ***Area*** | ***Knot*** | ***Mean RSE (%)*** | |
| ***Direct Estimate*** | ***geoSAE*** |
| 300 | 25 | 10 | 7,502849 | 1,609813 |
| 600 | 25 | 10 | 7,961569 | 1,164057 |
| 15 | 8,182034 | 1,206794 |
| 20 | 7,777237 | 1,253490 |
| 50 | 10 | 7,784023 | 1,650245 |
| 900 | 25 | 10 | 7,988857 | 0,979371 |
| 15 | 8,208081 | 0,961374 |
| 20 | 7,918791 | 0,965207 |
| 25 | 8,196916 | 0,988756 |
| 30 | 7,903386 | 0,952520 |
| 35 | 7,869620 | 0,945899 |
| 50 | 10 | 7,999528 | 1,288753 |
| 15 | 7,865969 | 1,305281 |
| 75 | 10 | 7,808762 | 1,594652 |
| 1200 | 25 | 10 | 7,898577 | 0,857674 |
| 15 | 7,993157 | 0,863562 |
| 20 | 7,994682 | 0,850408 |
| 25 | 7,857139 | 0,843406 |
| 30 | 8,119159 | 0,842816 |
| 35 | 8,231651 | 0,850419 |
| 50 | 10 | 8,057111 | 1,221510 |
| 15 | 7,835967 | 1,170801 |
| 20 | 7,966072 | 1,199354 |
| 75 | 10 | 7,598859 | 1,450714 |
| 15 | 7,777063 | 1,457837 |
| 100 | 10 | 7,750185 | 1,694684 |
| 1500 | 25 | 10 | 7,831821 | 0,732824 |
| 15 | 8,044614 | 0,764948 |
| 20 | 8,106639 | 0,765445 |
| 25 | 8,122636 | 0,767368 |
| 30 | 7,968192 | 0,745775 |
| 35 | 8,038391 | 0,754399 |
| 50 | 10 | 7,860669 | 1,067440 |
| 15 | 8,088350 | 1,086302 |
| 20 | 8,047208 | 1,075494 |
| 25 | 7,661133 | 1,050261 |
| 75 | 10 | 7,854994 | 1,322685 |
| 15 | 7,878814 | 1,321720 |
| 100 | 10 | 7,923388 | 1,518002 |
| 125 | 10 | 7,889565 | 1,653933 |

Dari skenario di atas, terlihat perbedaan yang dihasilkan dimana rata rata nilai RSE *Geoadditive Small Area Model* selalu memiliki nilai yang lebih kecil dibandingkan dengan pendugaan langsung.

Rata-rata nilai RSE yang dihasilkan dengan pendugaan langsung bersifat fluktuatif, karena tidak dipengaruhi oleh *knot* maupun jumlah area.

Lain halnya dengan pendugaan langsung, rata-rata nilai RSE dengan *Geoadditive Small Area Model* dipengaruhi oleh *knot* dan jumlah area. Pada area dan unit yang sama, penambahan jumlah knot maka akan menurunkan nilai RSE. Sedangkan penambahan jumlah area kecil dengan jumlah unit dan jumlah *knot* yang sama, akan menghasilkan RSE yang semakin besar.

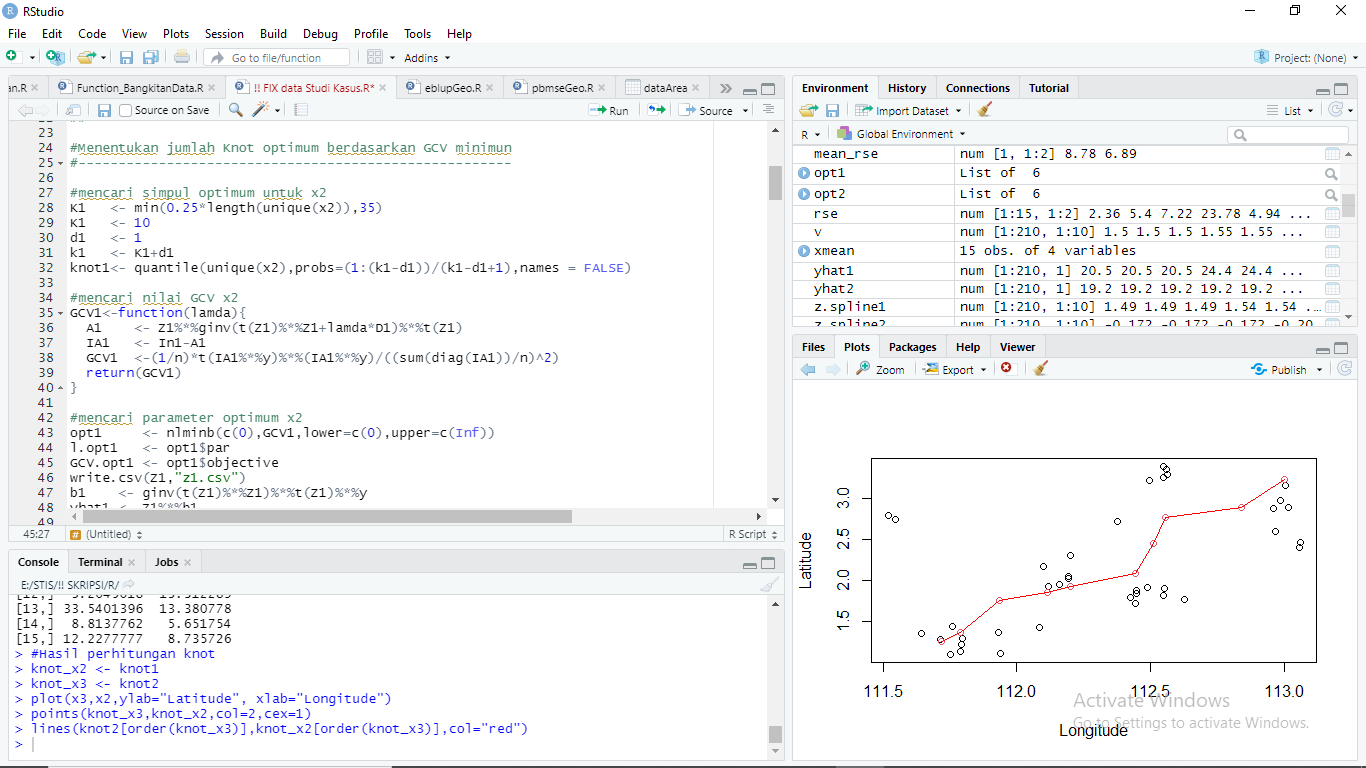
* + 1. Data Studi Kasus

Data studi kasus yang digunakan pada penelitian ini bersumber dari Ardiansyah, et. al. (2018) [7] dengan menambahkan lokasi penelitian yang terdekat dengan Kabupaten Seruyan, yaitu Kotawaringin Barat dan Kotawaringin Timur. Berikut rincian data yang digunakan:

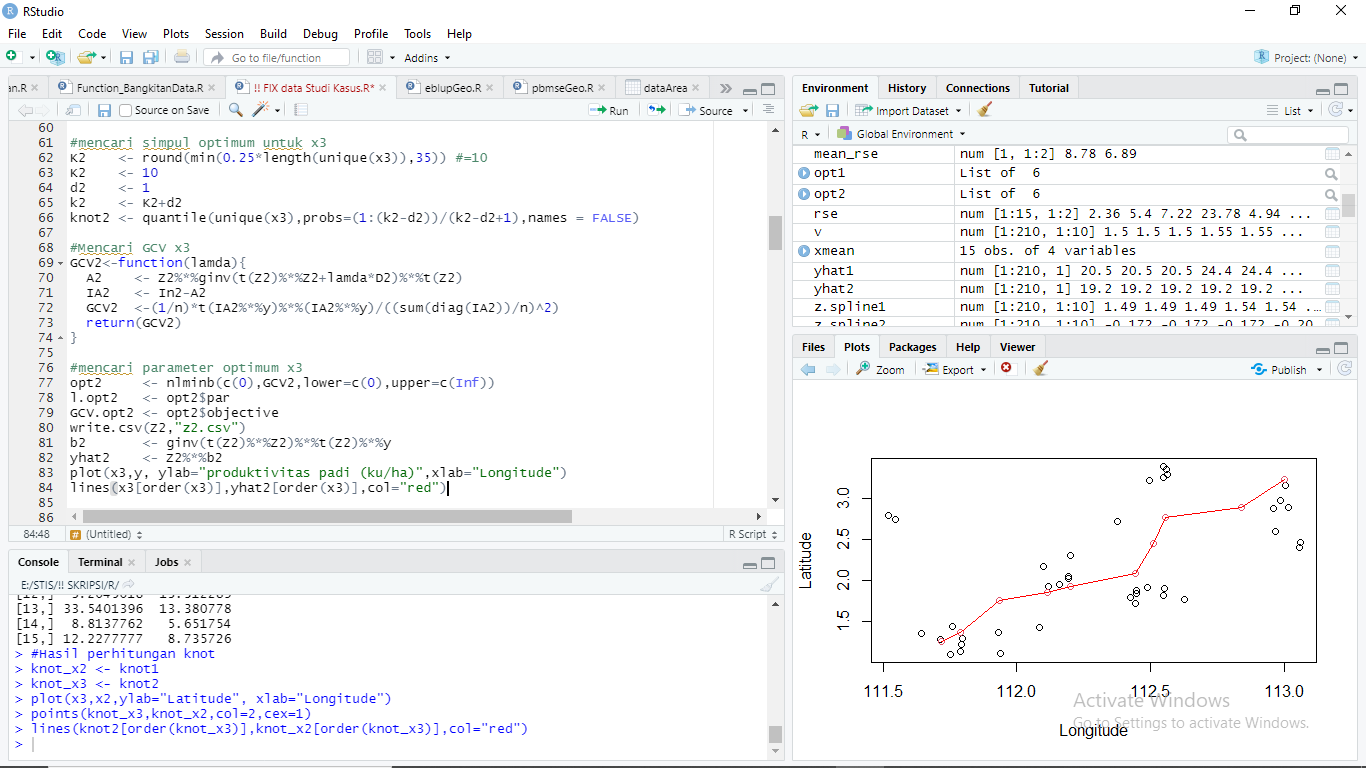
1. Data produktivitas tanaman padi level unit sebagai peubah respon dan digunakan sebagai *direct estimate* diperoleh dari Survei Ubinan Tahun 2016 di 15 kecamatan pada Kabupaten Seruyan, Kotawaringin Barat, dan Kotawaringin Timur yang bersumber dari BPS.
2. Data peubah penyerta tingkat kecamatan diperoleh dari Potensi Desa (PODES) Tahun 2014, dan Dinas Pertanian Kabupaten Seruyan, Kotawaringin Barat, dan Kotawaringin Timur, dengan rincian sebagai berikut:

* X1 = Proporsi luas panen padi sawah terhadap total luas panen (Podes untuk data level unit; Dinas Pertanian untuk data populasi tingkat kecamatan)
* X2 = Garis lintang atau *latitude* (Podes)
* X3 = Garis bujur atau *longitude* (Podes)

1. Penentuan Titik *Knot* Optimum



Gambar 6. Potongan *syntax* penentuan titik *knot* optimum pada X2



Gambar 7. Potongan *syntax* penentuan titik *knot* optimum pada X3

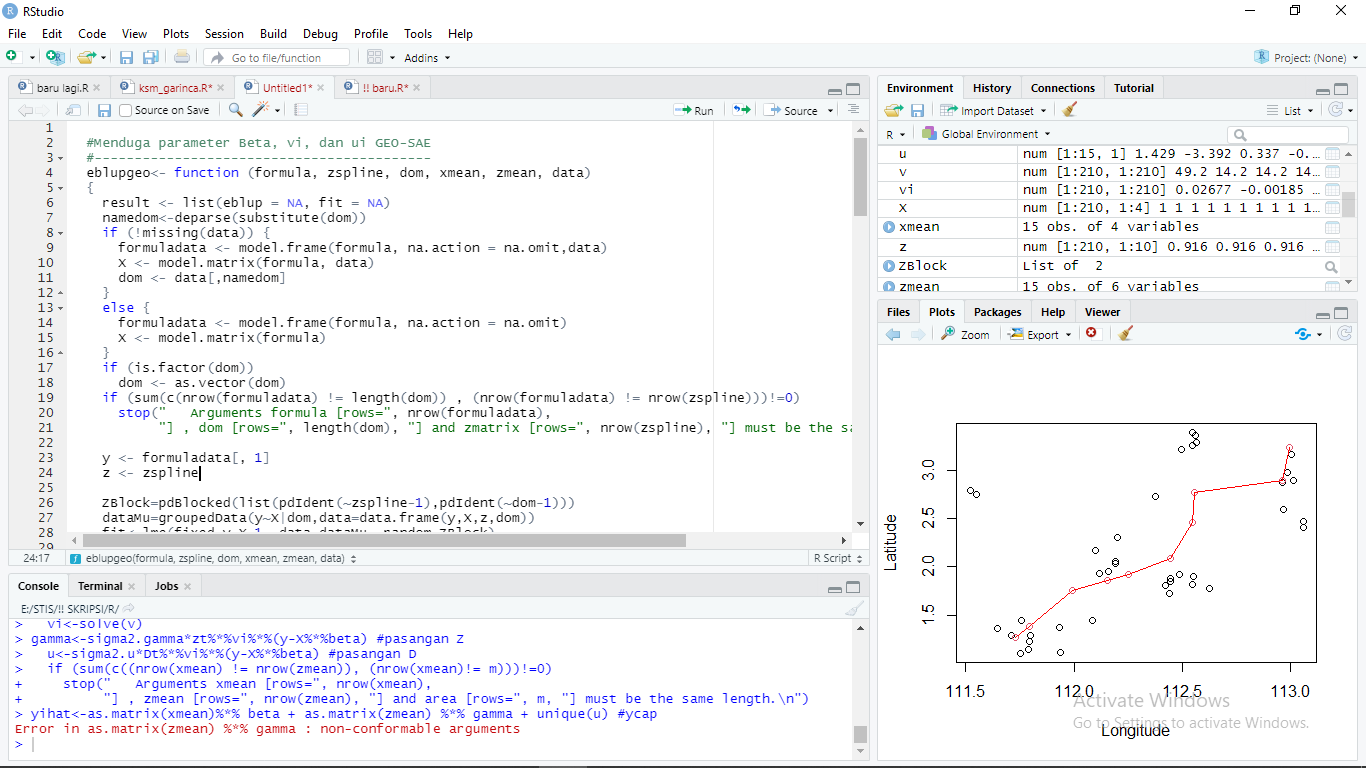
TABEL III

JUMLAH TITIK *KNOT* DAN GCV

|  |  |  |
| --- | --- | --- |
| ***knot*** | ***GCV Latitude*** | ***GCV Longitude*** |
| 1 | 75,86093 | 63,43841 |
| 2 | 75,86093 | 60,37617 |
| 3 | 75,86093 | 58,17688 |
| 4 | 75,86093 | 57,58701 |
| 5 | 71,41754 | 57,25100 |
| 6 | 57,69196 | 56,46000 |
| 7 | 66,04990 | 56,50157 |
| 8 | 67,02682 | 56,25743 |
| 9 | 59,50842 | 55,97979 |
| 10 | 56,97506 | 55,93138 |

Sebelum melakukan estimasi area kecil, maka langkah pertama adalah menentukan jumlah knot optimum berdasarkan nilai GCV minimum. Kemulusan kurva permukaan pendugaan spasial dipengaruhi oleh jumlah simpul yang digunakan pada fungsi basis radialnya [7]. Dari tabel III, *output* *syntax* yang dijalankan menunjukkan bahwa model aditif spasial dengan jumlah knot 10 memberikan GCV terendah, yaitu sebesar 56,975 untuk *latitude* dan 55,931 untuk *longitude*

1. Estimasi Produktivitas Padi



Gambar 8. Potongan *syntax* *function eblupgeo*

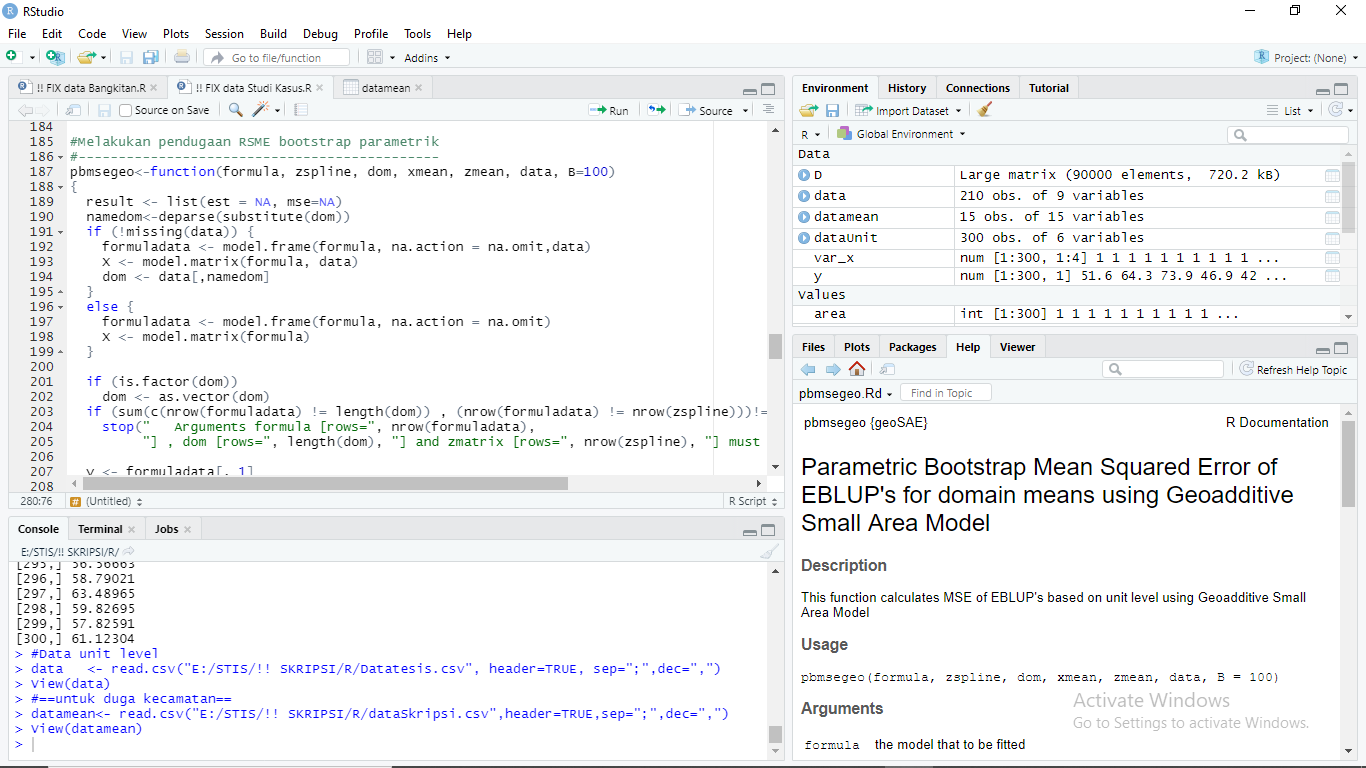
TABEL IV

ESTIMASI PRODUKTIVITAS PADI

|  |  |  |
| --- | --- | --- |
| ***Kecamatan*** | ***Hasil Estimasi Produktivitas Padi*** | |
| ***Direct Estimate*** | ***geoSAE*** |
| Arus Selatan | 18,512 | 29,103 |
| Seruyan Hilir | 30,020 | 34,811 |
| Seruyan Hilir Timur | 34,420 | 35,715 |
| Danau Sembuluh | 32,820 | 34,006 |
| Seruyan Tengah | 26,640 | 23,523 |
| Batu Ampar | 31,400 | 22,929 |
| Seruyan Hulu | 19,860 | 21,023 |
| Suling Tambun | 20,570 | 20,409 |
| Mentaya Hilir Selatan | 37,896 | 34,460 |
| Teluh Sampit | 42,636 | 39,796 |
| Pulau Hanaut | 29,686 | 35,127 |
| Seranau | 21,740 | 29,401 |
| Mentaya Hulu | 25,385 | 32,318 |
| Bukit Santuai | 24,711 | 24,321 |
| Telaga Antang | 34,593 | 31,213 |

Hasil estimasi ketiga pendekatan memberikan nilai yang hampir sama untuk 15 kecamatan pada Kabupaten Seruyan, Kotawaringin Barat, dan Kotawaringin Timur, hal ini berarti kecamatan yang memiliki nilai estimasi tinggi pada satu pendekatan, maka nilai estimasi pada pendekatan lainnya akan tinggi pula.

1. Perbandingan RSE



Gambar 9. Potongan *syntax* *function pbmsegeo*

TABEL V

PERBANDINGAN NILAI RSE

|  |  |  |
| --- | --- | --- |
| ***Kecamatan*** | ***RSE (%)*** | |
| ***Direct Estimate*** | ***geoSAE*** |
| Arus Selatan | 2,362 | 1,979 |
| Seruyan Hilir | 5,397 | 2,386 |
| Seruyan Hilir Timur | 7,217 | 1,934 |
| Danau Sembuluh | 23,777 | 9,414 |
| Seruyan Tengah | 4,942 | 1,569 |
| Batu Ampar | 6,236 | 3,656 |
| Seruyan Hulu | 13,242 | 3,102 |
| Suling Tambun | 5,611 | 1,907 |
| Mentaya Hilir Selatan | 1,180 | 1,510 |
| Teluh Sampit | 0,890 | 0,618 |
| Pulau Hanaut | 3,019 | 2,354 |
| Seranau | 3,205 | 15,936 |
| Mentaya Hulu | 33,540 | 17,073 |
| Bukit Santuai | 8,814 | 2,737 |
| Telaga Antang | 12,228 | 5,695 |
| ***Rata-Rata*** | ***8,777*** | ***4,791*** |

RSE merupakan ukuran presisi suatu estimasi relatif terhadap estimasinya. Nilai RSE bersifat stabil, sehingga dapat digunakan untuk membandingkan standard error hasil estimasi setiap karakteristik. Pada tabel V, terlihat bahwa rata-rata nilai RSE dengan pendekatan *Geoadditive Small Area Model* memiliki nilai lebih kecil yaitu sebesar 4,79% dibandingkan dengan *Direct Estimate*. Hal ini menunjukkan bahwa *Geoadditive Small Area Model* mampu menghasilkan hasil estimasi yang lebih presisi.

1. Penutup

Dari hasil dan pembahasan yang telah dijelaskan, maka dapat ditarik kesimpulan sebagai berikut.

1. *Script R* dengan metode *Geoadditive Small Area Model* telah dijalankan pada data simulasi. Hasil yang diperoleh dengan data simulasi menunjukkan bahwa metode *Geoadditive Small Area Model* memiliki rata-rata nilai RSE yang lebih kecil daripada pendugaan langsung.
2. Metode *Geoadditive Small Area Model* telah berhasil dikembangkan dengan pembangunan *package R* dengan nama *‘geoSAE’,* yang dapat diakses pada <https://cran.r-project.org/web/packages/geoSAE/index.html>
3. *Package R* yang telah dibangun dapat diimplementasikan dengan data simulasi dan data studi kasus. Hasil yang diperoleh dengan data studi kasus menunjukkan bahwa metode *Geoadditive Small Area Model* mampu menghasilkan hasil estimasi yang lebih presisi dibandingkan dengan pendugaan langsung.

Daftar Pustaka

1. Rao, J., & Molina, I. (2015). Small Area Estimation 2nd Edition. New Jersey: John Wiley and Sons Inc.
2. Pusponegoro, N. H., Djuraidah, A., Fitrianto, A., & Sumertajaya, I. M. (2019). Geoadditive Models in Small Area Estimation of Poverty. Journal of Data Science and Its Applications, 2(1), 11-18.
3. Fay, R. E., & Herriot, R. A. (1979). Estimates of Income for Small Places: An Application of James-Sten Procedures to Census Data. Journal of the American Statistical Association, 268-277.
4. Kurnia, A. (2009). The Best Predictions for Empirical Logarithm Transformation Models in Small Area Estimation with Application on SUSENAS Data (Doctoral dissertation, Ph. D dissertation, Department of Statistics. Bogor Agricultural University, Bogor, Indonesia).
5. E. E. Kammann and M. P. Wand, 2003. Geoadditive models, Applied Statistics, 52(1), pp.1–18
6. Bocci, C. (2010). Geoadditive small area model for the estimation of consumption expenditure in albania. Italy: University of Florence.
7. Ardiansyah, M., Djuraidah, A., & Kurnia, A. (2018). Pendugaan Area Kecil Data Produktivitas Tanaman Padi Dengan Geoadditive Small Area Model. Jurnal Penelitian Pertanian Tanaman Pangan, 2(2), 101-110.
8. Kandala, N. B., Fahrmeir, L., Klasen, S., & Priebe, J. (2009). Geo‐additive models of childhood undernutrition in three sub‐Saharan African countries. *Population, Space and Place*, *15*(5), 461-473.
9. Nkurunziza, H., Gebhardt, A., & Pilz, J. (2011). Geoadditive modelling of malaria in Burundi. *Malaria Journal*, *10*(1), 1-7.
10. Djuraidah, Anik, & Aunuddin. (2006). Kriging pada Thin-Plate Spline dengan Pendekatan Model Campuran. Jurnal Matematika Integratif, 5(2), 1-12
11. Salvati, N., Chandra, H., Ranalli, M. G. and Chambers, R. (2008), ‘Small Area Estimation Using a Nonparametric Model Based Direct Estimator’. Submitted for publication.
12. Benavent, Roberto & Morales, Domingo. (2015). Multivariate Fay-Herriot models for small area estimation. Computational Statistics and Data Analysis 94 2016 372-390.
13. Ardiansyah, M., Kurnia, A., Sadik, K., Djuraidah, A., & Wijayanto, H. (2021, March). Numerical Prediction of paddy weight of Crop Cutting Survey using Generalized Geoadditive Linear Mixed Model. In *Journal of Physics: Conference Series* (Vol. 1863, No. 1, p. 012024). IOP Publishing.