

LAPORAN PENELITIAN DOSEN STIS



MODEL FAY-HERRIOT MULTIVARIAT SPASIAL UNTUK PENDUGAAN AREA KECIL

Aplikasi pada Data Pengeluaran Per Kapita
di Jawa Tengah Tahun 2015

Dr. Azka Ubaidillah, SST., M.Si.
Sukim, SST, M.Si



**PUSAT PENELITIAN DAN PENGABDIAN MASYARAKAT (PPPM)
POLITEKNIK STATISTIKA STIS**

2018

Sengaja dikosongkan

LEMBAR PENGESAHAN

Laporan Penelitian Dosen dengan judul:

MODEL FAY-HERRIOT MULTIVARIAT SPASIAL UNTUK PENDUGAAN AREA KECIL

Aplikasi pada Data Pengeluaran Per Kapita
di Jawa Tengah Tahun 2015

Nama Peneliti:

Dr. Azka Ubaidillah, SST., M.Si.
Sukim, SST, M.Si

Dilaksanakan pada Februari 2018 sampai dengan Oktober 2018

Telah disahkan oleh Kepala Pusat Penelitian dan Pengabdian Masyarakat (PPPM)
Politeknik Statistika STIS, pada tanggal **25 Oktober 2018**

Menyetujui,
Kepala PPPM

Ketua Peneliti

Setia Pramana, S.Si., Ph.D.
NIP. 197707222000031002

Dr. Azka Ubaidillah, SST., M.Si.
NIP. 198202072004121001

Mengetahui
Wakil Direktur I,

Dr. Hardius Usman, S.Si., M.Si.
NIP. 196704251989011002

Sengaja dikosongkan

KATA PENGANTAR

Alhamdu lillah, segala puji bagi Alloh SWT. Atas berkat rahmat-Nya kami telah berhasil menyusun laporan hasil penelitian dengan judul “Model Fay-Herriot Multivariat Spasial untuk Pendugaan Area Kecil Aplikasi pada Data Pengeluaran Per Kapita di Jawa Tengah Tahun 2015” ini. Laporan hasil penelitian ini disusun sebagai pertanggungjawaban peneliti kepada Pusat Penelitian dan Pengabdian Masyarakat Sekolah Tinggi Ilmu Statistik (PPPM PS STIS).

Laporan hasil penelitian ini diharapkan mampu memberikan kontribusi pemahaman kepada pembaca (pengguna) dalam memberikan sedikit gambaran tentang pola model Fay-Herriot Multivariat Spasial terutama untuk pendugaan area kecil (*Small Area Estimation*).

Dengan selesainya laporan hasil penelitian ini, tak lupa kami mengucapkan terima kasih kepada PPPM PS STIS dan Unit Pendukung PS STIS (Bagian Umum dan Bagian Administrasi Akademik dan Kemahasiswaan) atas kerja samanya. Ucapan terima kasih kami sampaikan juga kepada Direktur Politeknik Statistika STIS dan para Ketua Unit Akademik dan Unit Pendukung PS STIS.

Akhirnya, kami berharap semoga hasil penelitian ini juga bermanfaat bagi peneliti dalam melakukan penelitian, terutama tentang SAE. Oleh karena itu, saran dan kritik demi perbaikan hasil penelitian ini sangat diharapkan.

Jakarta, Oktober 2018
Tim peneliti

Sengaja dikosongkan

ABSTRAK

MODEL FAY-HERRIOT MULTIVARIAT SPASIAL UNTUK PENDUGAAN AREA KECIL

**Aplikasi pada Data Pengeluaran Per Kapita
di Jawa Tengah Tahun 2015**

*(Spatial Multivariate Fay-Herriot Model for Small Area Estimation
with Application to Household Percapita Expenditure Data
in Central Java, 2015)*

Azka Ubaidillah¹, Sukim²

Politeknik Statistika STIS¹

Politeknik Statistika STIS²

Jalan Otista Raya 64 C Jakarta Timur

E-mail: azka@stis.ac.id¹ sukim@stis.ac.id²

ABSTRAK

Model Fay-Herriot multivariat (FHM) adalah salah satu model dalam pendugaan area kecil yang cukup populer. Model tersebut memanfaatkan korelasi antar beberapa peubah respon yang menjadi amatan. Benavent dan Morales (2016) dan Ubaidillah (2017) membuktikan bahwa model Fay-Herriot multivariat menghasilkan hasil kinerja model yang lebih baik daripada model Fay-Herriot univariat (FHU). Penelitian ini dilakukan untuk mengembangkan model Fay-Herriot multivariat dengan menambahkan efek spasial pada komponen acak area. Kajian analitik dilakukan untuk memperoleh formula penduga *empirical best linear unbiased prediction* (EBLUP) dan penduga *mean squared error* (MSE) dari penduga EBLUP berbasis model Fay-Herriot multivariat spasial (FHMS). Kajian simulasi dilakukan untuk mengevaluasi efisiensi model untuk dibandingkan dengan FHM dan FHU. Sedangkan aplikasi model diterapkan pada data pengeluaran perkapita rumah tangga di Jawa Tengah tahun 2015 yang bersumber dari Badan Pusat Statistik (BPS). Dari kajian simulasi dan aplikasi data pengeluaran perkapita rumah tangga di Jawa Tengah tahun 2015 diperoleh hasil bahwa model FHMS lebih efisien daripada model FHM dimana nilai *root mean squared error* (RMSE) model FHMS yang lebih kecil daripada model FHM.

Kata kunci: model Fay-Herriot multivariat, korelasi spasial, pendugaan area kecil, pengeluaran perkapita rumah tangga

ABSTRACT

The multivariate Fay-Herriot model (FHM) is one of the popular models for small area estimation. The FHM model utilizes the correlation between several variables of interest. Benavent and Morales (2016) and Ubaidillah (2017) showed that the multivariate Fay-Herriot model produces better model performance results than the univariate Fay-Herriot model. This research was conducted to develop a multivariate Fay-Herriot model by adding spatial effects to random area components. Analytical studies were conducted to obtain the formula of empirical best linear unbiased prediction (EBLUP) estimator and the formula of mean of squared error (MSE) of EBLUP estimator based on the spatial multivariate Fay-Herriot model (FHMS). Simulation studies were designed to evaluate the efficiency of the proposed models compared to FHM and FHU models. While the application model is applied to the expenditure data percapita of households in Central Java in 2015 which are sourced from the Statistics Indonesia (BPS). The simulation study and the

application results were found that the FHMS models are more efficient than the FHM models where the root mean squared error (RMSE) values of the FHMS models are smaller than the FHM models.

Keywords: *multivariate Fay-Herriot model, spatial correlation, small area estimation, household percapita expenditure*

DAFTAR ISI

KATA PENGANTAR	i
ABSTRAK	iii
DAFTAR ISI	v
DAFTAR TABEL	vii
DAFTAR GAMBAR	ix
DAFTAR LAMPIRAN	xi
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Identifikasi dan Perumusan Masalah	2
1.3. Tujuan Penelitian	3
BAB II TINJAUAN LITERATUR	5
2.1. Landasan Teori	5
2.2. Kerangka Pikir	10
BAB III METODOLOGI	11
3.1. Ruang Lingkup Penelitian	11
3.2. Metode Analisis	11
BAB IV HASIL DAN PEMBAHASAN	15
4.1. Kajian Simulasi	15
4.2. Aplikasi untuk Pendugaan Pengeluaran Perkapita Rumah tangga	16
BAB V KESIMPULAN DAN SARAN	19
5.1. Kesimpulan	19
5.2. Saran	19
DAFTAR PUSTAKA	21
LAMPIRAN	23

Sengaja dikosongkan

DAFTAR TABEL

Tabel 1. Nilai average RMSE	16
Tabel 2. Nilai rata-rata RMSE dan rata-rata CV ($(CV)^{-}$)	17

Sengaja dikosongkan

DAFTAR GAMBAR

Gambar 1. Kerangka Pikir Penelitian	10
Gambar 2. Coefficient of Variation (CV) dari peubah amatan Y1	18
Gambar 3. Coefficient of Variation (CV) dari peubah amatan Y2	18

Sengaja dikosongkan

DAFTAR LAMPIRAN

Lampiran 1. Data Penelitian	23
Lampiran 2. Program R untuk model FHM dan FHMS	24

Sengaja dikosongkan

BAB I

PENDAHULUAN

1.1 Latar Belakang

Berbagai survei umumnya dirancang untuk menduga parameter populasi untuk area besar, misalnya untuk wilayah nasional atau regional (provinsi, kabupaten/kota) dan pendugaan parameternya didasarkan pada rancangan pendugaan langsung (direct estimation). Untuk pendugaan parameter wilayah yang lebih kecil dimana jumlah contoh (sampel) kurang mencukupi atau bahkan tidak ada sampel sama sekali, maka pendugaan berdasarkan rancangan survei menjadi tidak valid karena kesalahan baku (standard error) hasil dugaannya besar sehingga analisis terhadap pendugaan langsung dari hasil survei untuk area kecil tersebut tidak dapat dipercaya.

Untuk mengatasi persoalan tersebut, saat ini telah dikembangkan metode pendugaan area kecil (Small Area Estimation/SAE) sebagai pendugaan tidak langsung (indirect estimation). Metode SAE selain menggunakan informasi dari peubah yang diamati dari hasil survei (sebagai hasil pendugaan langsung), juga memanfaatkan informasi tambahan dari peubah penyerta (auxiliary variable) yang memiliki hubungan kuat dengan peubah yang sedang diamati tersebut. Peubah penyerta yang digunakan dalam model SAE bisa diperoleh dari hasil sensus, catatan administrasi atau dari survei lain.

Secara umum, model SAE dibedakan dalam dua kelompok, yaitu model level area dan model level unit (Rao dan Molina, 2015). Model level area digunakan apabila peubah respon yang diamati (berupa rata-rata karakteristik area kecil) dan peubah penyerta hanya tersedia pada level area kecil. Sedangkan model level unit digunakan apabila peubah respon yang diamati (berupa karakteristik unit atau elemen dalam area kecil) dan peubah penyerta tersedia pada level unit dalam area kecil. Beberapa metode pendugaan area kecil antara lain Syntetic, Empirical Best Linier Unbiased Prediction (EBLUP), Empirical Bayes (EB) dan Hierarchical Bayes (HB) (Ghosh dan Rao, 1994). Pendugaan dengan metode EBLUP diperoleh dari formula penduga Best Linier Unbiased Prediction (BLUP). Penduga BLUP yang pertama kali dikenalkan oleh Henderson (1953), diperoleh dengan cara meminimalisasi nilai Mean Squared Error (MSE) diantara kelompok penduga tak bias linier. Penduga BLUP tidak tergantung pada sebaran normal dari pengaruh komponen acak (random effect), namun tergantung pada nilai ragam (koragam) dari pengaruh komponen acak. Nilai ragam (koragam) pengaruh komponen acak pada penduga BLUP dapat

diduga dengan metode moment (Henderson, 1953), Maximum likelihood (ML) atau Restricted maximum likelihood (REML) berdasarkan likelihood dari sebaran normal sebagaimana dikemukakan oleh Hartley dan Rao (1967) dan Patterson dan Thompson (1971). Dengan memasukkan nilai dugaan ragam (koragam) pada penduga BLUP maka akan diperoleh penduga Empirical BLUP (EBLUP).

Penerapan metode EBLUP dalam SAE untuk pertama kalinya diterapkan oleh Fay dan Herriot (1979) dengan menghitung nilai dugaan area kecil untuk pendapatan perkapita (percapita income) di Amerika Serikat. Fay dan Herriot (1979) membuktikan bahwa penggunaan metode EBLUP untuk pendugaan area kecil pada pendapatan perkapita di Amerika Serikat tahun 1970 menghasilkan rata-rata kesalahan yang lebih kecil dibandingkan dengan metode proporsi yang sebelumnya pernah dilakukan oleh Biro Statistik Amerika Serikat (The U.S. Census Bureau).

Metode EBLUP level area yang dipopulerkan oleh Fay dan Herriot (1987) kemudian dikembangkan oleh beberapa peneliti, diantaranya adalah Datta dkk (1991) yang mengemukakan model Fay-Herriot multivariat untuk menduga besarnya median pendapatan suatu keluarga di setiap negara bagian (area kecil) di Amerika Serikat. Datta dkk (1991) dan Benavent dan Morales (2016) membuktikan bahwa penggunaan metode EBLUP multivariat menghasilkan nilai absolute relative error (ARE) yang lebih kecil dari pada penggunaan metode pendugaan langsung dan metode EBLUP univariat. Isaki dkk (2000) mengembangkan model EBLUP dengan kesalahan contoh yang berkorelasi (correlated sampling errors) dimana hasil dugaannya menghasilkan nilai efisiensi MSE sebesar 400 persen relatif dibandingkan dengan hasil pendugaan langsung. Selain itu, Pfeiffermann dan Tiller (2005) juga mengembangkan model EBLUP untuk data yang berkorelasi tempat dan waktu (state-space models) dan Marhuenda dkk (2013) menggunakan metode EBLUP berbasis model spatio-temporal dengan mengasumsikan pengaruh area pada model Rao-Yu mengikuti model Spatial Autoregressive (SAR).

Sebagaimana diketahui, data pengeluaran perkapita rumah tangga merupakan data yang penting sebagai salah satu indikator ekonomi yang banyak digunakan untuk mengukur tingkat kemakmuran dan kesejahteraan, khususnya tingkat kemiskinan. Analisis tentang data pengeluaran perkapita rumah tangga menjadi penting terutama bagi pemerintah untuk merencanakan dan mengevaluasi berbagai program dan paket kebijakan dalam rangka pengurangan angka kemiskinan dan meningkatkan kesejahteraan warga negara. Penelitian tentang analisis pengeluaran perkapita rumah menggunakan SAE diantaranya dilakukan oleh Ubaidillah (2017)

yang menerapkan metode Fay-Herriot multivariat untuk pendugaan pengeluaran perkapita rumah tangga di Jawa Tengah menurut kabupaten/kota periode bulan September tahun 2015.

Penelitian ini ditujukan untuk merancang metode baru dalam pendugaan area kecil yaitu metode Fay-Herriot multivariat (FHM) dengan menambahkan pengaruh korelasi spasial pada komponen acak area. Metode yang diusulkan tersebut kemudian disebut sebagai model Fay-Herriot multivariat spasial (FHMS). Formula penduga EBLUP dan penduga MSE nya berbasis FHMS diperoleh dengan penurunan rumus melalui kajian analitik yang cukup intensif. Kajian simulasi juga dilakukan untuk mengevaluasi efisiensi model. Sedangkan kajian aplikasi dilakukan dengan menerapkan model pada data pengeluaran perkapita rumah tangga di Jawa Tengah tahun 2015.

Laporan penelitian ini disusun dalam lima bagian. Bagian pertama yaitu pendahuluan yang berisi kajian latar belakang dan tujuan penelitian. Bagian kedua yaitu metode yang berisi penurunan formula penduga EBLUP dan MSE EBLUP berbasis model FHMS. Bagian ketiga yaitu kajian simulasi yang menampilkan hasil simulasi untuk mengevaluasi kinerja model FHMS dibandingkan dengan model FHM dan FHU. Bagian keempat yaitu aplikasi model yang menampilkan hasil aplikasi model FHMS untuk pendugaan pengeluaran perkapita rumah tangga di Jawa Tengah periode bulan September 2015. Dan terakhir bagian kelima yaitu kesimpulan yang berisi kesimpulan hasil penelitian dan saran untuk penelitian yang akan datang.

1.2 Identifikasi dan Perumusan Masalah

Berdasarkan uraian latar belakang di atas, dapat disusun perumusan masalah sebagai berikut:

1. Bagaimana memperoleh formula penduga EBLUP dan MSE nya berbasis model Fay-Herriot multivariat spasial?
2. Bagaimana kinerja model yang diukur dari tingkat efisiensinya?
3. Bagaimana aplikasi pengembangan model pada data rata-rata pengeluaran perkapita rumah tangga?

1.3 Tujuan Penelitian

Dari perumusan masalah di atas, maka tujuan penelitian ini adalah:

1. Memperoleh formula penduga EBLUP dan MSE nya berbasis model Fay-Herriot multivariat spasial.

2. Mengetahui kinerja model yang diukur dari tingkat efisiensinya.
3. Mengaplikasikan pengembangan model pada data rata-rata pengeluaran perkapita rumah tangga.

BAB II

TINJAUAN LITERATUR

2.1. Landasan Teori

Metode Pendugaan Area Kecil

Metode pendugaan area kecil atau *small area estimation* (SAE) merupakan pendugaan tidak langsung (*indirect estimation*) yaitu pendugaan berbasis model dengan memanfaatkan informasi tambahan dari peubah lain sebagai peubah penyerta (*auxiliary variable*) yang memiliki hubungan dengan peubah yang sedang diamati. Peubah penyerta ini dapat diperoleh dari hasil sensus atau survei yang lain atau dari catatan administrasi. Dengan menambahkan peubah penyerta dalam pemodelan, metode SAE diharapkan dapat meningkatkan efisiensi penduga parameter dalam area kecil jika dibandingkan dengan penduga parameter dari metode pendugaan langsung. Biasanya ukuran yang digunakan untuk mengetahui perbaikan efisiensi pendugaan tidak langsung dari pendugaan langsung adalah ukuran *Mean Squared Error* (MSE), *Average Relative Error* (ARE) atau *Average Squared Error* (ASE). Dalam hal ini, metode pendugaan tidak langsung akan menghasilkan nilai MSE, ARE atau ASE yang lebih kecil dibandingkan metode pendugaan langsung (Ghosh dan Rao, 1994).

Dalam kaitan hubungan antara peubah respon yang diamati dengan peubah penyerta, maka secara umum model SAE dibedakan menjadi dua kelompok yaitu (Rao dan Molina, 2015):

- a. Pada model level area, peubah yang diamati diasumsikan merupakan fungsi dari rata-rata peubah respon, $\theta_i = g(\bar{Y}_i)$ untuk $g(\cdot)$ tertentu yang berkaitan dengan data penyerta $\mathbf{x}_i = (x_{i1}, \dots, x_{iK})^T$ yang tersedia sampai level area kecil dan mengikuti model linier sebagai berikut:

$$\theta_i = \mathbf{x}_i^T \boldsymbol{\beta} + z_i u_i, i = 1, \dots, m \text{ (area kecil)} \quad (1)$$

dengan z_i adalah konstanta bernilai positif yang diketahui dan $\boldsymbol{\beta} = (\beta_1, \dots, \beta_K)^T$ adalah koefisien regresi. Sedangkan u_i adalah pengaruh acak area kecil yang diasumsikan memiliki sebaran (seringkali sebaran normal) yang identik dan saling bebas dengan

$$E(u_i) = 0, \quad \text{Var}(u_i) = \sigma_u^2.$$

Untuk melakukan inferensi tentang rata-rata area kecil $\bar{Y}_i$ pada persamaan (1), dimisalkan suatu penduga langsung $\hat{Y}_i$ diketahui dan

$$\hat{\theta}_i = g(\hat{Y}_i) = \theta_i + e_i, \quad (2)$$

dimana $e_i \sim^{ind}(0, \psi_i)$ dan ψ_i diketahui. Dengan menggabungkan persamaan (1) dan (2) maka diperoleh model

$$\hat{\theta}_i = \mathbf{x}_i^T \boldsymbol{\beta} + z_i u_i + e_i, \quad (3)$$

dengan u_i dan e_i saling bebas. Model (3) merupakan bentuk model campuran linier atau *linear mixed models* (LMM) dan dikenal sebagai model Fay-Herriot.

- b. Pada model level unit, data peubah respon yang diamati dan data peubah penyerta diasumsikan tersedia untuk masing-masing anggota (unit) populasi ke- j pada setiap area kecil ke- i , $\mathbf{x}_{ij} = (x_{ij1}, \dots, x_{ijK})^T$. Selanjutnya peubah respon y_{ij} yang berkaitan dengan $\mathbf{x}_{ij}$ disusun dalam model regresi tersarang sebagai berikut:

$$y_{ij} = \mathbf{x}_{ij}^T \boldsymbol{\beta} + u_i + e_{ij}, \quad i = 1, \dots, m \text{ dan } j = 1, \dots, n_i \quad (4)$$

dalam hal ini pengaruh acak area kecil u_i mempunyai sebaran identik dan saling bebas, sedangkan $e_{ij} = k_{ij} \tilde{e}_{ij}$ dengan konstanta k_{ij} diketahui dan $\tilde{e}_{ij}$ mempunyai sebaran yang identik dan saling bebas dengan

$$E(\tilde{e}_{ij}) = 0, \quad \text{Var}(\tilde{e}_{ij}) = \sigma_e^2.$$

Komponen pengaruh acak u_i dan e_{ij} dalam hal ini juga diasumsikan saling bebas.

Empirical Best Linear unbiased Prediction (EBLUP)

Penduga EBLUP untuk model SAE level area merujuk pada persamaan (3) sebagai berikut (Rao dan Molina, 2015):

$$\hat{\theta}_i = \mathbf{x}_i^T \boldsymbol{\beta} + z_i u_i + e_i, \quad i = 1, \dots, m, \quad (5)$$

dimana $\mathbf{x}_i$ adalah vektor kovariat level area berukuran $K \times 1$, pengaruh area $u_i \sim^{iid}(0, \sigma_u^2)$ diasumsikan bebas dengan galat contoh (*sampling error*) $e_i \sim^{ind}(0, \psi_i)$, $\hat{\theta}_i$ adalah penduga langsung untuk parameter area $\theta_i = g(\bar{Y}_i)$, dan z_i adalah konstanta bernilai positif yang diketahui. Model (5) adalah bentuk khusus dari *general linear mixed models* pada persamaan dengan penyesuaian sebagai berikut:

$$\mathbf{y}_i = \hat{\theta}_i, \quad \mathbf{X}_i = \mathbf{x}_i^T, \quad \mathbf{Z}_i = z_i, \quad \mathbf{u}_i = u_i, \quad \mathbf{e}_i = e_i, \text{ dan } \boldsymbol{\beta} = (\beta_1, \dots, \beta_K)^T.$$

Dengan penyesuaian tersebut, maka matriks koragam dari $\mathbf{u}_i$ dan $\mathbf{e}_i$ menjadi skalar, yaitu

$$\mathbf{G}_i = \sigma_u^2, \quad \mathbf{R}_i = \psi_i,$$

dan matriks ragam koragam dari $\mathbf{y}_i = \hat{\theta}_i$ menjadi:

$$\mathbf{V}_i = \sigma_u^2 + \psi_i.$$

Demikian juga target parameter $\mu_i = \mathbf{1}_i^T \boldsymbol{\beta} + \mathbf{m}_i^T \mathbf{u}_i$ disesuaikan menjadi $\mu_i = \hat{\theta}_i = \mathbf{x}_i^T \boldsymbol{\beta} + z_i u_i$ dengan $\mathbf{1}_i = \mathbf{x}_i$ dan $\mathbf{m}_i = z_i$.

Berdasarkan penyesuaian di atas dan menggunakan cara yang sama untuk memperoleh penduga BLUP untuk μ_i pada persamaan (5), maka penduga BLUP level area untuk $\hat{\theta}_i$ dirumuskan sebagai berikut:

$$\tilde{\theta}_i^H = \mathbf{x}_i^T \tilde{\boldsymbol{\beta}} + \gamma_i (\hat{\theta}_i - \mathbf{x}_i^T \tilde{\boldsymbol{\beta}})$$

$$= \gamma_i \hat{\theta}_i + (1 - \gamma_i) \mathbf{x}_i^T \tilde{\boldsymbol{\beta}}, \quad (6)$$

dimana

$$\tilde{\boldsymbol{\beta}} = \tilde{\boldsymbol{\beta}}(\sigma_u^2) = [\sum_{i=1}^m \mathbf{x}_i \mathbf{x}_i^T / (\sigma_u^2 z_i^2 + \psi_i)]^{-1} [\sum_{i=1}^m \mathbf{x}_i \hat{\theta}_i / (\sigma_u^2 z_i^2 + \psi_i)] \quad (7)$$

adalah penduga BLUE untuk $\boldsymbol{\beta}$, dan

$$\gamma_i = \sigma_u^2 z_i^2 / (\sigma_u^2 z_i^2 + \psi_i). \quad (8)$$

Nilai *Mean Squared Error* (MSE) penduga BLUP $\tilde{\theta}_i^H$ dirumuskan:

$$\text{MSE}(\tilde{\theta}_i^H) = E(\tilde{\theta}_i^H - \theta_i)^2 = g_{1i}(\sigma_u^2) + g_{2i}(\sigma_u^2), \quad (9)$$

dimana

$$g_{1i}(\sigma_u^2) = \frac{\sigma_u^2 z_i^2 \psi_i}{(\sigma_u^2 z_i^2 + \psi_i)} = \gamma_i \psi_i \quad (10)$$

dan

$$g_{2i}(\sigma_u^2) = (1 - \gamma_i)^2 \mathbf{x}_i^T [\sum_{t=1}^m \mathbf{x}_t \mathbf{x}_t^T / (\sigma_u^2 z_t^2 + \psi_t)]^{-1} \mathbf{x}_i. \quad (11)$$

Penduga BLUP (6) mengandung nilai σ_u^2 yang tidak diketahui. Beberapa metode pendugaan σ_u^2 diantaranya metode moment, ML dan REML digunakan untuk memperoleh penduga asimtotik untuk σ_u^2 yaitu $\hat{\sigma}_u^2$ (Cressie, 1992). Dengan mengganti σ_u^2 yang tidak diketahui pada persamaan (7) dengan penduga $\hat{\sigma}_u^2$, maka diperoleh penduga EBLUP $\hat{\theta}_i^H$ sebagai berikut:

$$\hat{\theta}_i^H = \hat{\gamma}_i \hat{\theta}_i + (1 - \hat{\gamma}_i) \mathbf{x}_i^T \hat{\boldsymbol{\beta}}, \quad (12)$$

dimana $\hat{\gamma}_i$ dan $\hat{\boldsymbol{\beta}}$ adalah nilai dari γ_i dan $\tilde{\boldsymbol{\beta}}$ pada persamaan (6) ketika σ_u^2 diganti dengan $\hat{\sigma}_u^2$.

Model Fay-Herriot Spasial

Dalam aplikasi real data, penduga langsung yang menjadi amatan dapat mengandung korelasi spasial. Oleh karena itu Cressie dan Chan (1989) mengembangkan model spasial untuk EBLUP. Penambahan pengaruh spasial dilakukan dengan penyesuaian ragam pengaruh acak area, yaitu

$$\mathbf{G}_i = (\mathbf{I} - \rho \mathbf{W})^{-1} \sigma_u^2, \quad \mathbf{V}_i = (\mathbf{I} - \rho \mathbf{W})^{-1} \sigma_u^2 + \psi_i. \quad (13)$$

dimana ρ adalah koefisien spasial, $\mathbf{W}$ adalah matriks pembobot spasial dan $\mathbf{I}$ adalah matriks identitas.

EBLUP Multivariat

Metode EBLUP multivariat dikembangkan oleh Gonzáles-Manteiga (2008) dari metode EBLUP dengan peubah respon yang lebih dari satu dan diasumsikan saling berkorelasi. Formula penduga metode BLUP multivariat adalah (Rao dan Molina, 2015):

$$\tilde{\boldsymbol{\theta}}_i^H = \boldsymbol{\Sigma}_u (\boldsymbol{\Psi}_i + \boldsymbol{\Sigma}_u)^{-1} \hat{\boldsymbol{\theta}}_i + \boldsymbol{\Psi}_i (\boldsymbol{\Psi}_i + \boldsymbol{\Sigma}_u)^{-1} \mathbf{X}_i \tilde{\boldsymbol{\beta}}, \quad (14)$$

dimana $\tilde{\boldsymbol{\beta}}$ adalah penduga *weighted least squares* (WLS) untuk $\boldsymbol{\beta}$ dengan rumus:

$$\tilde{\boldsymbol{\beta}} = [\sum_{i=1}^m \mathbf{X}_i^T (\boldsymbol{\Psi}_i + \boldsymbol{\Sigma}_u)^{-1} \mathbf{X}_i]^{-1} [\sum_{i=1}^m \mathbf{X}_i^T (\boldsymbol{\Psi}_i + \boldsymbol{\Sigma}_u)^{-1} \hat{\boldsymbol{\theta}}_i]$$

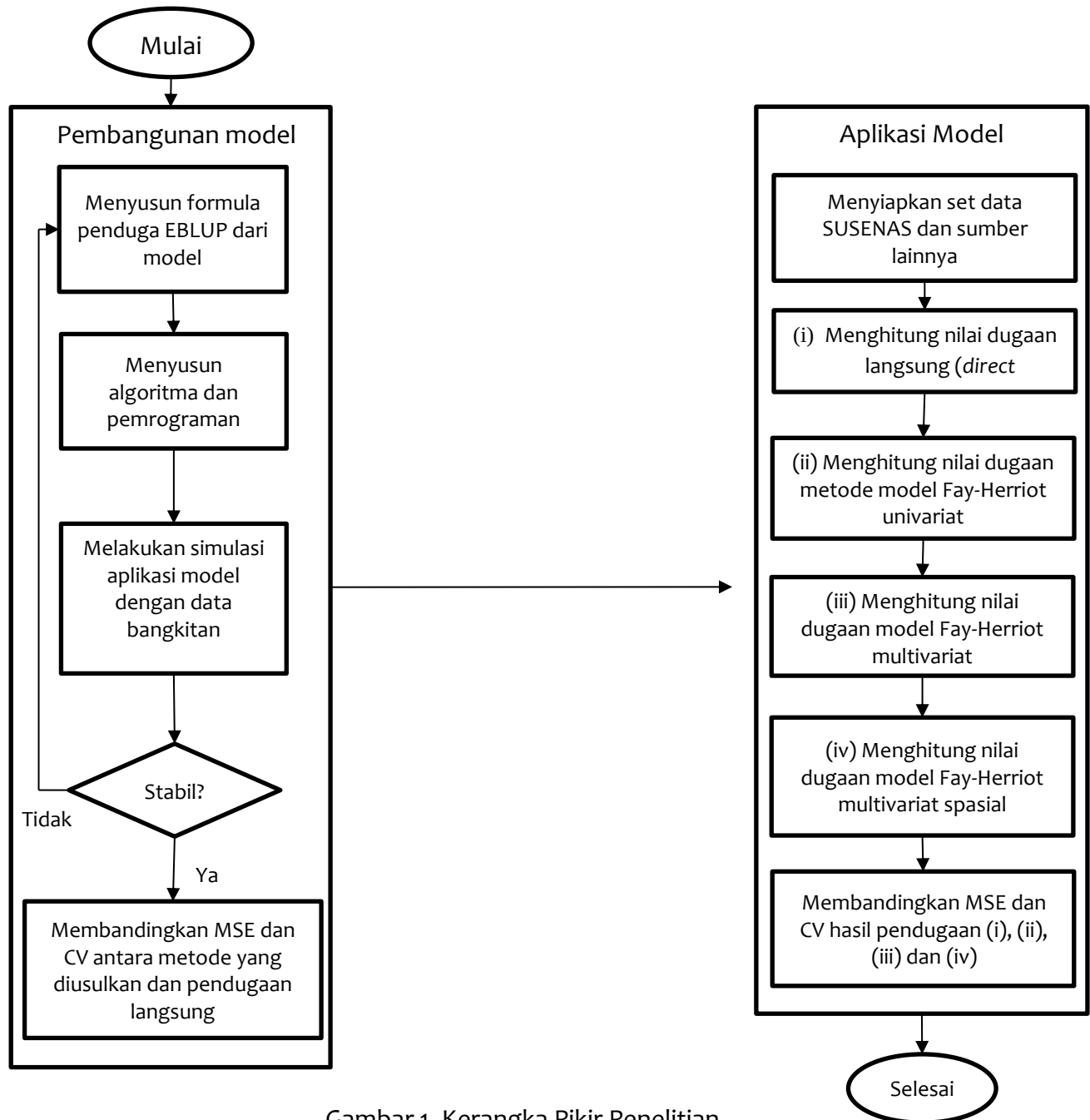
Adapun MSE dari $\tilde{\boldsymbol{\theta}}_i^H$ dirumuskan:

$$\begin{aligned} \text{MSE}(\tilde{\boldsymbol{\theta}}_i^H) &= E(\tilde{\boldsymbol{\theta}}_i^H - \boldsymbol{\theta}_i)(\tilde{\boldsymbol{\theta}}_i^H - \boldsymbol{\theta}_i)^T \\ &= (\boldsymbol{\Psi}_i^{-1} + \boldsymbol{\Sigma}_u^{-1})^{-1} + (\boldsymbol{\Psi}_i^{-1} + \boldsymbol{\Sigma}_u^{-1})^{-1} \boldsymbol{\Sigma}_u^{-1} \mathbf{X}_i \\ &\quad \times [\sum_{i=1}^m \mathbf{X}_i^T (\boldsymbol{\Psi}_i + \boldsymbol{\Sigma}_u)^{-1} \mathbf{X}_i]^{-1} \mathbf{X}_i' \boldsymbol{\Sigma}_u^{-1} (\boldsymbol{\Psi}_i^{-1} + \boldsymbol{\Sigma}_u^{-1})^{-1}. \end{aligned} \quad (15)$$

Penduga BLUP multivariat mengandung struktur matriks ragam koragam $\boldsymbol{\Sigma}_u$ yang tidak diketahui. Matriks ragam koragam $\boldsymbol{\Sigma}_u$ dapat diduga dengan metode REML untuk memperoleh $\hat{\boldsymbol{\Sigma}}_u$. Matriks ragam koragam yang tidak diketahui $\boldsymbol{\Sigma}_u$ pada persamaan (13) diganti dengan $\hat{\boldsymbol{\Sigma}}_u$ untuk diperoleh penduga EBLUP multivariat.

2.2. Kerangka Pikir

Kerangka pikir penelitian ini ditampilkan pada Gambar 1 sebagai berikut:



Gambar 1. Kerangka Pikir Penelitian

BAB III

METODOLOGI

3.1. Ruang Lingkup Penelitian

Metode SAE yang digunakan dalam penelitian ini mencakup tiga metode, yaitu metode Fay-Herriot univariat (FHU), metode Fay-Herriot multivariat (FHM) dan metode Fay-Herriot multivariat spasial (FHMS). Data yang digunakan adalah data bangkitan untuk kajian simulasi, sedangkan untuk kajian aplikasi digunakan data pengeluaran perkapita rumahtangga makanan (Y_1) dan pengeluaran perkapita rumahtangga non makanan (Y_2) sebagai peubah amatan yang bersumber dari Survei Sosial Ekonomi Nasional (SUSENAS) periode bulan September 2015. Untuk peubah penyerta digunakan lima peubah yaitu banyaknya keluarga pra-sejahtera (X_1), upah minimum kabupaten/kota (X_2), produk domestik regional bruto menurut harga konstan tahun 2010 (X_3), luas panen padi sawah (X_4) dan rasio banyaknya sekolah menengah atas terhadap 10000 penduduk (X_5).

3.2. Metode Analisis

Model Fay-Herriot Multivariat Spasial

Misalkan untuk setiap d -parameter yang diamati pada m -area diasumsikan merupakan fungsi dari rata-rata peubah respon, $\boldsymbol{\mu}_d = (\mu_1, \dots, \mu_m)_d^T = g(\bar{\mathbf{y}}_d)$ untuk $g(\cdot)$ tertentu yang berkaitan dengan peubah penyerta $\mathbf{X}_d = (\mathbf{x}_1, \dots, \mathbf{x}_{p_d})_d^T$ tersedia sampai level area kecil dan mengikuti model linier sebagai berikut:

$$\boldsymbol{\mu}_d = \mathbf{X}_d \boldsymbol{\beta}_d + \mathbf{u}_d, \quad d = 1, \dots, D \text{ (peubah)} \quad (16)$$

dengan $\boldsymbol{\beta}_d = (\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_{p_d})_d^T$ adalah koefisien regresi. Sedangkan $\mathbf{u}_d$ adalah pengaruh acak area kecil yang diasumsikan memiliki sebaran (seringkali sebaran normal) yang identik dan saling bebas dengan $E(\mathbf{u}_d) = 0$, $\text{Var}(\mathbf{u}_d) = \sigma_{u_d}^2$.

Untuk melakukan inferensi tentang rata-rata area kecil $\boldsymbol{\mu}_d$ pada persamaan (16), dimisalkan suatu penduga langsung $\mathbf{y}$ diketahui dan

$$\mathbf{y}_d = \boldsymbol{\mu}_d + \mathbf{e}_d, \quad (17)$$

dimana $\mathbf{e}_d \sim^{ind}(\mathbf{0}, \boldsymbol{\Psi}_d)$ merupakan vektor galat contoh pada peubah amatan ke- d dan $\boldsymbol{\Psi}_d = \text{diag}_{1 \leq i \leq m}(\psi_i)$ adalah matriks kovarian galat contoh yang berupa matriks diagonal dan diasumsikan $\psi_i = \sigma_{e_{ii}}$ diketahui, $i = 1, \dots, m$ (area kecil). Dengan menggabungkan persamaan (16) dan (17) maka diperoleh model

$$\mathbf{y}_d = \mathbf{X}_d \boldsymbol{\beta}_d + \mathbf{u}_d + \mathbf{e}_d, \quad (18)$$

dengan $\mathbf{u}_d$ dan $\mathbf{e}_d$ saling bebas. Model (18) merupakan bentuk model campuran linier dan dikenal sebagai model Fay-Herriot. Dengan menyusun model (18) menjadi bentuk *stack*, maka model (18) dapat ditulis sebagai model Fay-Herriot multivariat (FHM) sebagai berikut:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} + \mathbf{e}, \quad (19)$$

dimana $\mathbf{y} = \text{col}_{1 \leq d \leq D}(\mathbf{y}_d)$, $\mathbf{X} = \text{diag}_{1 \leq d \leq D}(\mathbf{X}_d)$, $\boldsymbol{\beta} = \text{col}_{1 \leq d \leq D}(\boldsymbol{\beta}_d)$, $\mathbf{u} = \text{col}_{1 \leq d \leq D}(\mathbf{u}_d)$, dan $\mathbf{e} = \text{col}_{1 \leq d \leq D}(\mathbf{e}_d)$.

Asumsi yang digunakan pada model (4) adalah $\mathbf{u} \sim N_D(\mathbf{0}, \mathbf{V}_u)$, $\mathbf{e} \sim N_D(\mathbf{0}, \mathbf{V}_e)$, dan $\text{cov}(\mathbf{u}, \mathbf{e}) = \mathbf{0}$. Matriks $\mathbf{V}_u = \text{diag}_{1 \leq d \leq D}(\sigma_{u_d}^2) \otimes \mathbf{I}_m$ adalah matriks kovarian dari efek acak area berukuran $mD \times mD$, $\mathbf{I}_m$ adalah matriks identitas berukuran $m \times m$ dan notasi $\otimes$ adalah operasi *cronecker product matrix*. Matriks $\mathbf{V}_e = \text{col}_{1 \leq d \leq D} \left(\text{row}_{1 \leq d \leq D} \left(\text{diag}_{1 \leq i \leq m}(\sigma_{e_{id}}) \right) \right)$ adalah matriks kovarian galat contoh berukuran $mD \times mD$ dan diasumsikan diketahui. Dengan demikian, matriks kovarian $\mathbf{y}$ dapat ditulis $\text{var}(\mathbf{y}) = \mathbf{V}_u + \mathbf{V}_e = \mathbf{V}$. Model (19) akan menjadi model FHU jika matriks $\mathbf{V}_e$ dan matriks $\mathbf{V}$ berupa matriks diagonal.

Pada kenyataannya, penduga langsung dari beberapa area kecil kemungkinan mempunyai hubungan secara spasial. Pengembangan metode SAE dengan menambahkan korelasi spasial dilakukan pada komponen area acak area (Rao dan Molina, 2015). Dari sisi model FHM, model (19) dapat dimodifikasi menjadi model FHMS sebagai berikut:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{v} + \mathbf{e}, \quad (20)$$

dimana $\mathbf{v} = \mathbf{I}_D \otimes (\mathbf{I}_m - \rho \mathbf{W})\mathbf{v} + \mathbf{u}$, dengan ρ adalah koefisien korelasi spasial, $\mathbf{W}$ adalah matriks pembobot spasial berukuran $m \times m$, $\mathbf{I}_m$ adalah matriks identitas berukuran $m \times m$ dan $\mathbf{u} \sim N_D(\mathbf{0}, \mathbf{V}_u)$.

Selanjutnya, komponen acak area $\mathbf{v}$ dapat ditulis $\mathbf{v} = (\mathbf{I}_D \otimes (\mathbf{I}_m - \rho \mathbf{W}))^{-1} \mathbf{u}$ sehingga matriks ragamnya $\text{var}(\mathbf{v}) = (\mathbf{I}_D \otimes (\mathbf{I}_m - \rho \mathbf{W}))^{-1} \mathbf{V}_u (\mathbf{I}_D \otimes (\mathbf{I}_m - \rho \mathbf{W}))^{-1} = \text{diag}_{1 \leq d \leq D}(\sigma_{u_d}^2) \otimes (\mathbf{I}_m - \rho \mathbf{W})^{-1} (\mathbf{I}_m - \rho \mathbf{W})^{-1} = \mathbf{G}$. Dengan demikian, matriks kovarian $\mathbf{y}$ dapat ditulis $\text{var}(\mathbf{y}) = \mathbf{G} + \mathbf{V}_e = \mathbf{\Sigma}$. Jika $\rho = 0$, model (20) yaitu model FHMS, akan menjadi model FHM sebagaimana model (19).

Empirical Best Linear Unbiased Prediction (EBLUP)

Setelah melalui proses penurunan rumus (proses detailnya tidak ditampilkan dalam paper ini), penduga EBLUP dari persamaan (20) adalah:

$$\hat{\boldsymbol{\mu}}_{FHMS} = \mathbf{X} \hat{\boldsymbol{\beta}}_{FHMS} + \hat{\mathbf{G}} \hat{\boldsymbol{\Sigma}}^{-1} (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}}_{FHMS}), \quad (21)$$

dimana $\hat{\mathbf{G}} = \text{diag}_{1 \leq d \leq D}(\hat{\sigma}_{u_d}^2) \otimes (\mathbf{I}_m - \hat{\rho} \mathbf{W})^{-1} (\mathbf{I}_m - \hat{\rho} \mathbf{W})^{-1}$ dan $\hat{\boldsymbol{\Sigma}} = \hat{\mathbf{G}} + \mathbf{V}_e$. Penduga komponen ragam area $\hat{\sigma}_{u_d}^2$ dan korelasi spasial $\hat{\rho}$ diperoleh dengan metode *restricted maximum likelihood* (REML). Penduga $\hat{\boldsymbol{\beta}}_{FHMS}$ diperoleh dengan metode *weighted least square* (WLS) sebagai berikut:

$$\hat{\boldsymbol{\beta}}_{FHMS} = (\mathbf{X}^T \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{y},$$

dengan $\text{cov}(\hat{\boldsymbol{\beta}}_{FHMS}) = (\mathbf{X}^T \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{X})^{-1}$.

Penduga *mean squared error* (MSE) dari persamaan (21) adalah sebagai berikut:

$$\text{mse}(\hat{\boldsymbol{\mu}}_{FHMS}) = g_1(\hat{\boldsymbol{\delta}}) + g_2(\hat{\boldsymbol{\delta}}) + 2g_3(\hat{\boldsymbol{\delta}}), \quad (22)$$

dimana

$$\begin{aligned} g_1(\hat{\boldsymbol{\delta}}) &= \text{diag}(\hat{\mathbf{F}} \mathbf{V}_e), \\ g_2(\hat{\boldsymbol{\delta}}) &= \text{diag} \left[(\mathbf{I}_{mD} - \hat{\mathbf{F}}) \mathbf{X} (\mathbf{X}^T \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{I}_{mD} - \hat{\mathbf{F}})^T \right], \text{ dan} \\ g_3(\hat{\boldsymbol{\delta}}) &= \text{diag} \left[\sum_{k=1}^q \sum_{l=1}^q \text{cov}(\hat{\delta}_k, \hat{\delta}_l) \hat{\mathbf{F}}_{(k)} \hat{\boldsymbol{\Sigma}} \hat{\mathbf{F}}_{(l)} \right], k, l = 1, \dots, q, \end{aligned}$$

dimana $\boldsymbol{\delta} = (\sigma_{u_1}^2, \dots, \sigma_{u_D}^2, \rho)^T$ adalah vektor komponen ragam berukuran $q \times 1$ yang diduga menggunakan metode REML. Matriks $\hat{\mathbf{F}} = \hat{\mathbf{G}} \hat{\boldsymbol{\Sigma}}^{-1}$ adalah rasio ragam acak area dengan ragam

total. Matriks $\mathbf{\Gamma}_{(k)} = \frac{\partial \mathbf{\Gamma}}{\partial \delta_k}$ adalah matriks turunan parsial $\mathbf{\Gamma}$ terhadap δ_k , dan $\mathbf{I}_{mD}$ adalah matriks identitas berukuran $m \times D$.

BAB IV

HASIL DAN PEMBAHASAN

3.1. Kajian Simulasi

Tujuan dilakukan kajian simulasi adalah untuk mengevaluasi kinerja model yang diukur berdasarkan nilai *root mean squared error* (RMSE). Semakin kecil nilai RMSE maka kinerja model semakin baik. Pada kajian ini, nilai RMSE dibandingkan untuk 3 model, yaitu M_1 : model FHMS, M_2 : model FHM dan M_3 : FHU.

Tahapan simulasi yang dilakukan adalah sebagai berikut:

- a. Dengan observasi sebanyak $m = 35$, bangkitkan $X_1 \sim N(10, 4)$, $X_2 \sim Unif(5, 10)$, $\beta_1 = (1, 2)^T$, $\beta_2 = (1, 3)^T$, $\mathbf{V}_u = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$, $\mathbf{V}_e = \begin{bmatrix} 25 & \sigma_{e_{12}} \\ \sigma_{e_{12}} & 25 \end{bmatrix}$, dimana $\sigma_{e_{12}} = \rho_e \sqrt{25 \times 25}$.

Gunakan $\mathbf{W}_{35 \times 35}$ matriks pembobot spasial Provinsi Jawa Tengah dengan pendekatan *Queen*.

- b. Untuk masing-masing nilai korelasi spasial $\rho = 0.1, 0.5$ dan 0.9 , lakukan lagi untuk masing-masing nilai korelasi galat contoh $\rho_e = 0.1, 0.5$ dan 0.9 , ulangi sebanyak $B=1000$ kali untuk:
 - i. Bangkitkan matriks $\mathbf{U}^{(b)} \sim N_2(\mathbf{0}, \mathbf{V}_u)$ dan matriks $\mathbf{E}^{(b)} \sim N_2(\mathbf{0}, \mathbf{V}_e)$, $b = 1, \dots, B$, dimana N_2 menunjukkan sebaran normal multivariat dengan 2 peubah. Kemudian lakukan stack terhadap matriks $\mathbf{U}^{(b)}$ dan matriks $\mathbf{E}^{(b)}$ sehingga diperoleh vektor $\mathbf{u}^{(b)}$ dan vektor $\mathbf{e}^{(b)}$.
 - ii. Hitung vektor $\mathbf{v}^{(b)} = (\mathbf{I}_D \otimes \rho \mathbf{W})^{-1} \mathbf{u}^{(b)}$, selanjutnya hitung $\mathbf{y}$ berdasarkan persamaan (5).
 - iii. Hitung nilai parameter $\mu^{(b)} = \mathbf{X}\beta + \mathbf{v}^{(b)}$.
 - iv. Hitung penduga EBLUP untuk tiga model, $\hat{\mu}^{(bM)}$, $M = 1, 2, 3$.
 - v. Hitung RMSE untuk tiga model sebagai berikut:

$$\text{RMSE}(\hat{\mu}^{(bM)}) = \sqrt{\frac{1}{m} \sum_{i=1}^m (\hat{\mu}^{(bM)} - \mu^{(b)})^2}$$

- c. Hitung *average* RMSE untuk tiga model sebagai berikut:

$$\overline{\text{RMSE}}(\hat{\mu}^{(M)}) = \frac{1}{B} \sum_{b=1}^B \text{RMSE}(\hat{\mu}^{(bM)})$$

Hasil simulasi yang ditampilkan pada Tabel 1 menunjukkan bahwa model FHMS menghasilkan rata-rata RMSE yang lebih kecil dibanding model FHM dan FHU. Secara umum terlihat juga bahwa untuk nilai koefisien korelasi spasial dan koefisien korelasi galat contoh yang semakin kuat, nilai rata-rata RMSE menunjukkan nilai yang semakin kecil. Tabel 1 juga memperlihatkan bahwa model FHM lebih dominan dibanding model FHU untuk semua kondisi nilai koefisien korelasi.

Tabel 1. Nilai *average* RMSE $\left(\overline{\text{RMSE}}(\hat{\mu}^{(M)})\right)$, dari model FHU, FHM dan FHMS menurut berbagai nilai koefisien korelasi spasial (ρ) dan koefisien korelasi galat contoh (ρ_e).

ρ	Variable	$\rho_e = 0.1$			$\rho_e = 0.5$			$\rho_e = 0.9$		
		FHU	FHM	FHMS	FHU	FHM	FHMS	FHU	FHM	FHMS
0.1	Y1	1.181	1.182	1.200	1.181	1.153	1.168	1.182	1.076	1.086
	Y2	1.187	1.188	1.205	1.180	1.152	1.166	1.193	1.078	1.088
0.5	Y1	1.215	1.215	1.194	1.224	1.195	1.173	1.218	1.114	1.088
	Y2	1.218	1.218	1.197	1.221	1.193	1.169	1.212	1.113	1.086
0.9	Y1	1.328	1.328	1.177	1.325	1.310	1.159	1.324	1.260	1.089
	Y2	1.326	1.326	1.170	1.332	1.314	1.163	1.320	1.259	1.089

Dengan demikian, dari kajian simulasi tersebut dapat dikatakan bahwa model FHMS memiliki kinerja model yang lebih baik dibandingkan dengan model FHM dan model FHU. Model FHM juga terlihat memiliki kinerja model yang lebih baik daripada model FHU. Hal tersebut selaras dengan penelitian sebelumnya yang dilakukan oleh Datta dkk (1991), Benavent dan Morales (2016) dan Ubaidillah (2017).

3.2. Aplikasi untuk Pendugaan Pengeluaran Perkapita Rumahtangga

Data pengeluaran perkapita rumahtangga di Indonesia dikumpulkan secara rutin oleh Badan Pusat Statistik (BPS) melalui Survei Sosial Ekonomi Nasional (Susenas). Sejak tahun 2015, Susenas diselenggarakan dua kali selama satu tahun, yaitu di bulan Maret untuk estimasi level kabupaten/kota dan bulan September untuk estimasi level provinsi. Penelitian ini mengambil lokus di Provinsi Jawa Tengah untuk Susenas bulan September 2015.

Sebelum menerapkan metode SAE untuk menduga tingkat pengeluaran perkapita rumahtangga, terlebih dahulu dilakukan penghitungan langsung (*direct estimation*). Penduga langsung untuk rata-rata area kecil dan penduga ragamnya menggunakan formula yang diusulkan oleh Benavent dan Morales (2015).

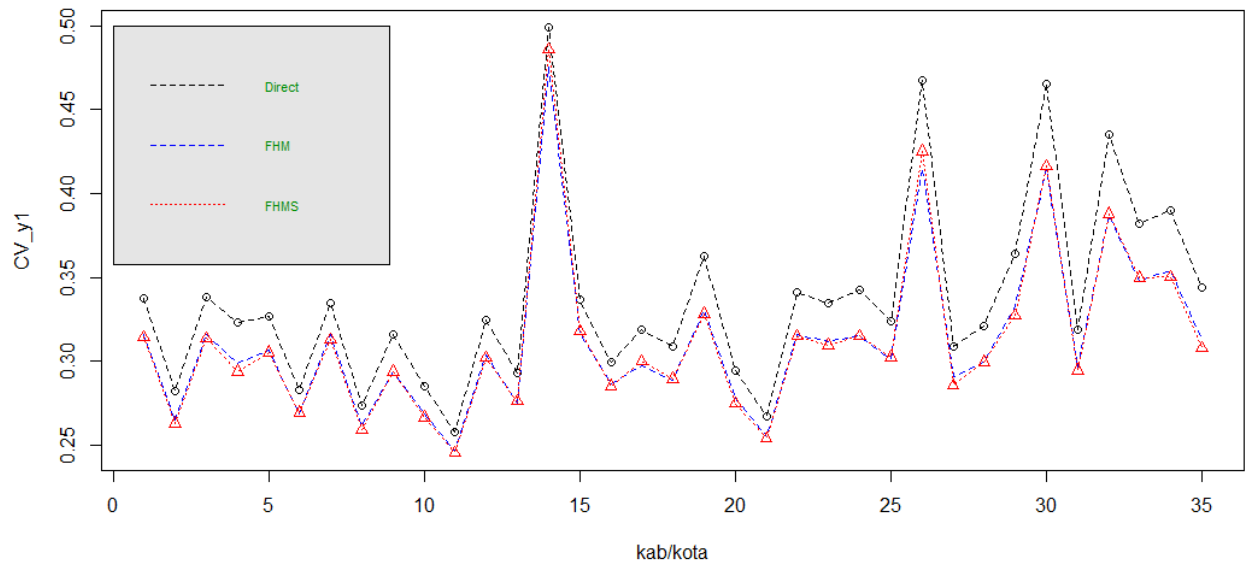
Peubah amatan yang digunakan yaitu data pengeluaran perkapita rumahtangga makanan (Y_1) dan pengeluaran perkapita rumahtangga non makanan (Y_2) yang bersumber dari Survei Sosial Ekonomi Nasional (SUSENAS) periode bulan September 2015. Untuk peubah penyerta digunakan lima peubah yaitu banyaknya keluarga pra-sejahtera (X_1), upah minimum kabupaten/kota (X_2), produk domestik regional bruto menurut harga konstan tahun 2010 (X_3), luas panen padi sawah (X_4) dan rasio banyaknya sekolah menengah atas terhadap 10000 penduduk (X_5). Kelima peubah penyerta tersebut bersumber dari publikasi Provinsi Jawa Tengah dalam Angka 2016.

Hasil penerapan model terhadap data pengeluaran perkapita rumahtangga ditampilkan pada Tabel 2 dan Gambar 1 dan Gambar 2. Dari Tabel 1 terlihat bahwa rata-rata RMSE ($\overline{RMSE}$) dan rata-rata *coefficient of variation* ($\overline{CV}$) yang dihasilkan oleh model FHMS lebih kecil dibandingkan $\overline{RMSE}$ dan $\overline{CV}$ yang dihasilkan oleh model pendugaan langsung, model FHU dan model FHM baik untuk peubah Y_1 dan Y_2 .

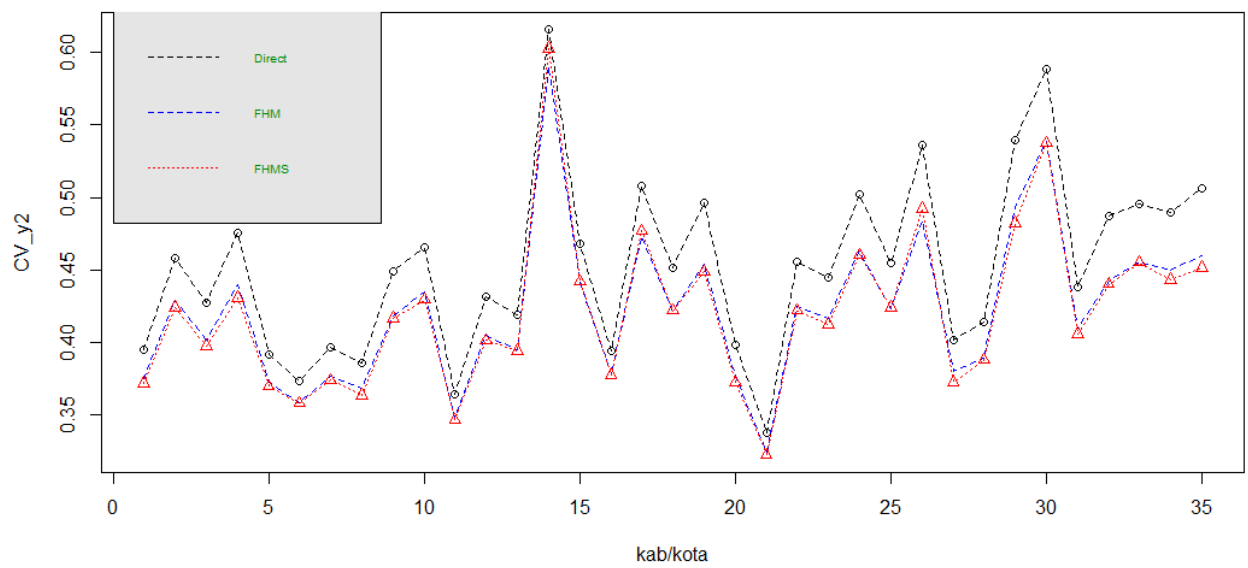
Tabel 2. Nilai rata-rata RMSE ($\overline{RMSE}$) dan rata-rata CV ($\overline{CV}$) dari model Pendugaan Langsung, model FHU, model FHM dan model FHMS untuk Pendugaan Rata-rata Pengeluaran Perkapita Rumahtangga.

Kriteria	Peubah	Pendugaan langsung	FHU	FHM	FHMS
$\overline{RMSE}$	Y_1	0.0434	0.0412	0.0403	0.0402
	Y_2	0.0567	0.0544	0.0530	0.0527
$\overline{CV}$	Y_1	0.3372	0.3208	0.3131	0.3125
	Y_2	0.4531	0.4350	0.4232	0.4208

Gambar 1 dan Gambar 2 menampilkan nilai CV dari model pendugaan langsung, model FHM dan model FHMS menurut kabupaten/kota. CV dari model FHU yang nilainya lebih besar dari CV model FHM tidak ditampilkan dalam gambar. Gambar 1 menunjukkan bahwa CV untuk EBLUP dari peubah amatan Y_1 yang dihasilkan oleh model FHMS sedikit lebih rendah daripada CV untuk EBLUP dari peubah amatan Y_1 yang dihasilkan oleh model FHM. Jika dibandingkan dengan CV dari model pendugaan langsung, CV model FHM dan FHMS menunjukkan nilai yang lebih kecil. Seiring dengan itu, Gambar 2 juga menunjukkan pola yang sama dengan Gambar 1.



Gambar 2. *Coefficient of Variation (CV)* dari peubah amatan Y_1 yang dihasilkan oleh model pendugaan langsung, model FHM dan model FHMS menurut kabupaten/kota.



Gambar 3. *Coefficient of Variation (CV)* dari peubah amatan Y_2 yang dihasilkan oleh model pendugaan langsung, model FHM dan model FHMS menurut kabupaten/kota.

BAB V

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Dari kajian analitik diperoleh formula penduga EBLUP dan formula penduga MSE berbasis model FHMS sebagaimana persamaan (21) dan persamaan (22). Dari kajian simulasi dan kajian aplikasi data pengeluaran perkapita rumahtangga dapat disimpulkan bahwa model FHMS menunjukkan kinerja yang lebih baik daripada model FHM dan model FHU. Penambahan pengaruh korelasi spasial pada komponen pengaruh acak area cukup signifikan mempengaruhi perbaikan hasil pendugaan metode SAE multivariat.

5.2. Saran

Berdasarkan hasil dan kesimpulan yang telah diperoleh, maka peneliti dapat memberikan saran sebagai berikut:

1. Sebelum menerapkan model FHMS, peneliti sebaiknya melakukan uji korelasi spasial antar peubah respon yang diperoleh dari penduga langsung. Korelasi spasial yang signifikan mengindikasikan bahwa model FHMS akan dapat bekerja secara efektif.
2. Untuk kasus peubah respon yang tidak mengikuti sebaran normal, peneliti dapat mengembangkan model FHMS dengan menggunakan metode *empirical bayes* atau *hierarchical bayes*.
3. Model FHMS belum dapat mengatasi masalah pencilan (*outlier*), untuk itu pada penelitian berikutnya peneliti dapat mengembangkan model FHMS yang kekar (*robust*).

Sengaja dikosongkan

DAFTAR PUSTAKA

- Benavent R, Morales D. (2016). Multivariate Fay-Herriot models for small area estimation. *Computational Statistics and Data Analysis*, **94**:372-390.
- Cressie, N. (1992). REML Estimation in Empirical Bayes Smoothing of Census Undercount. *Survey Methodology*, **18**:75-94.
- Datta, G.S., Fay, R.E., and Ghosh, M. (1991). Hierarchical and Empirical Bayes Multivariate Analysis in Small Area Estimation. *Proceedings of Bureau of The Census 1991 Annual Research Conference*. U.S. Bureau of The Census, Washington, DC. pp. 63-67.
- Datta, G.S., Ghosh, M., Smith, D.D., dan Lahiri, P. (2002). On The Asymtotic Theory of Conditional and Unconditional Coverage Probabilities of Empirical Bayes Confidence Intervals. *Scandinavian Journal of Statistics*, **29**:139-152.
- Fay, R.E., dan Herriot, R.A. (1979). Estimates of Income for Small Places: an Application of James-Stein Procedures to Census Data. *Journal of the American Statistical Association*, **74**:269-277.
- Ghosh, M dan Rao, J.N.K. (1994). Small Area Estimation: An Appraisal. *Statistical Science*. Vol. 9, No. 1, p. 55-76.
- Gonzáles-Manteiga, W., Lombardia, M.J., Molina, L., Morales, D., and Santamaria, L. (2008). Analytic and Bootstrap Approximations of Prediction Errors Under a Multivariate FayHerriot Model. *Computational Statistics and Data Analysis*, **52**:5242-5252.
- Henderson, C.R. (1953). Estimation of Variance and Covariance Components. *Biometrics*, **9**:226-252.
- Isaki, C.T., Huang, E.T., dan Tsay, J.H. (1991). Smoothing Adjustment Factors from the 1990 Post Enumeration Survey, *Proceedings of the Social Statistics Section*, Washington, DC: American Statistical Association, pp. 338–343.
- Marhuenda, Y., Molina, I., and Morales, D. (2013). Small Area Estimation with Spatio-Temporal Fay-Herriot Models. *Computational Statistics and Data Analysis*, **58**:308-325.
- Molina, I., dan Morales, D. 2009. Small Area Estimation on Poverty Indicators, *Boletin de Estadistica e Investigacion Operativa*, **25**:218-225.
- Patterson, H.D. dan Thompson, R. (1971). Recovery of Inter-Block Information When Block Sizes are Unequal. *Biometrika*, **58**:545-554.
- Pfeffermann, D. dan Tiller, R.B. (2005). Bootstrap Approximation to Prediction MSE for State-Space Models with Estimated Parameters. *Journal of Time Series Analysis*, **26**:893-916.
- Rao, J.N.K., dan Molina, I. (2015). *Small Area Estimation Second Edition*. John Wiley and Sons, Inc., New York.
- Rao, J.N.K., dan Yu, M. (1994). Small Area Estimation by Combining Time Series and Cross-Sectional Data. *Canadian Journal of Statistics*, **22**:511-528.
- Ubaidillah, A. (2017). Simultaneous Equation Models for Small Area Estimation [Disertasi]. Bogor: IPB.

Sengaja dikosongkan

LAMPIRAN 1. Data Penelitian

Kabupaten/Kota	y1	y2	x1	x2	x3	x4	x5	Ukuran sampel
Banjarnegara	12,53404	12,07523	10,52549	14,05058	9,35733	10,05655	0,57661	188
Batang	12,70237	12,11166	10,51992	14,19907	9,36652	10,61142	0,67287	190
Blora	12,5695	12,13464	12,01708	14,09956	9,41142	11,30068	0,99752	185
Boyolali	12,78903	12,51021	11,33499	14,15448	9,74599	10,73474	0,86127	194
Brebes	12,79363	12,2098	11,53692	14,08554	10,13029	11,51157	0,65679	234
Cilacap	12,86825	12,51667	11,3788	14,23882	11,3359	11,79616	0,77298	247
Demak	13,00459	12,53446	11,27148	14,37227	9,5522	11,46034	1,33284	208
Grobogan	12,62306	12,05735	12,49512	14,08171	9,61938	11,72356	0,92495	234
Jepara	12,79415	12,36241	11,36835	14,11562	9,70057	10,63489	1,10242	205
Karanganyar	12,68334	12,50774	10,22369	14,16617	9,91388	10,78168	0,63069	182
Kebumen	12,78085	12,37091	10,78095	14,09662	9,6275	11,21171	0,95368	209
Kendal	12,87801	12,36427	11,31049	14,30996	10,06318	10,67563	0,85962	195
Klaten	12,86977	12,51247	10,87216	14,15198	9,97076	11,10454	0,7853	207
Kota Magelang	13,19389	13,21875	7,979	14,10893	8,51467	6,2634	3,06317	113
Kota Pekalongan	12,92577	12,42562	8,6625	14,22098	8,65787	7,50549	0,94467	149
Kota Semarang	13,14104	13,17248	10,39882	14,46209	11,53763	8,82717	1,09928	225
Kota Tegal	12,97982	12,98956	8,89385	14,14121	9,04465	6,45205	1,34081	137
Kudus	12,67189	12,41018	10,30384	14,29063	11,04457	10,17877	0,92626	184
Magelang	12,5698	12,25888	11,45591	14,1591	9,79344	10,98672	0,78683	195
Pati	12,87924	12,47697	11,67025	14,08554	10,05737	11,57166	1,03821	206
Pekalongan	12,75702	12,33259	10,7374	14,196	9,4436	10,64956	0,73227	189
Pemalang	12,81092	12,30757	11,32438	14,09692	9,53918	11,31482	0,5898	206
Purbalingga	12,70545	12,29308	10,7084	14,13578	9,51446	10,59368	0,58995	187
Purworejo	12,79286	12,25912	10,65108	14,07787	9,24426	10,94738	0,87276	172
Rembang	12,87222	12,32482	10,96961	14,07787	9,23817	10,54821	0,90444	174
Kota Salatiga	13,35648	13,06854	8,40782	14,18773	8,89872	7,21082	1,57763	126
Semarang	12,96088	12,72644	11,12221	14,29174	10,22321	10,62816	0,72935	196
Sragen	12,88349	12,53491	11,36339	14,07787	9,912	11,45099	0,88734	197
Sukoharjo	13,03563	12,90082	10,26458	14,14912	9,92444	10,81505	0,6827	178
Kota Surakarta	13,06739	13,16383	9,06705	14,16476	10,20193	5,273	1,81559	157
Tegal	12,93436	12,44402	11,0618	14,13251	9,84986	11,04089	0,70181	215
Temanggung	12,67052	12,4028	10,63871	14,08783	9,39165	10,2325	0,72403	187
Wonogiri	12,86103	12,59223	10,67119	14,07248	9,68718	10,9702	0,72707	195
Wonosobo	12,80581	12,39837	10,63166	14,09768	9,29587	10,302	0,57906	198
Banyumas	12,78853	12,50913	11,16335	14,11562	10,27844	11,07526	0,73354	231

LAMPIRAN 2. Program R untuk model FHM dan FHMS

```
direct <- read.csv2("D:/STIS/UPPM/Riset 1/direct.csv")
Penimbang <- read.delim("D:/STIS/UPPM/Riset 1/Penimbang.txt", header=FALSE)
W=Penimbang
W=as.matrix(W)

y1=direct$y1
y2=direct$y2
n=length(y1)
I=diag(1,n)

var1=direct$y1^2
var2=direct$y2^2
var12=direct$y12^2
size=direct$size

x1=direct$x1
x2=direct$x2
x3=direct$x3
x4=direct$x4
x5=direct$x5
R0=matrix(0,n,n)

# EBLUP
y1_e=mseFH(y1~x1+x2+x5,var1)
y2_e=mseFH(y2~x3+x4+x5,var2)
y1_e$est$fit$estcoef
y2_e$est$fit$estcoef
y1_e$est$fit$refvar
y2_e$est$fit$refvar
```



```
PRECISION = 1e-10
```

```
MAXITER=100
```

```
##Model FHMS
```

```
Id <- diag(rep(1,2))
```

```
R1=diag(var1)
```

```
R2=diag(var2)
```

```
R12=diag(var12)
```

```
R_12=rbind(cbind(R1,R12),cbind(R12,R2))
```

```
#R_1_2=rbind(cbind(R1,R0),cbind(R0,R2))
```

```
d1omega=matrix(c(1,0,0,0),2,2)
```

```
d2omega=matrix(c(0,0,0,1),2,2)
```

```
d1Omega=kronecker(d1omega,I)
```

```
d2Omega=kronecker(d2omega,I)
```

```
X1=cbind(rep(1,n),x1,x2,x5,rep(0,n),rep(0,n),rep(0,n),rep(0,n))
```

```
X2=cbind(rep(0,n),rep(0,n),rep(0,n),rep(0,n),rep(1,n),x3,x4,x5)
```

```
X=rbind(X1,X2)
```

```
y=as.vector(c(y1,y2))
```

```
sigmau=as.matrix(c(median(var1),median(var2)),2,1)
```

```
rho=-0.9
```

```
param=c(sigmau,rho)
```

```
Param=c()
```

```
k <- 0
```

```
diff <- PRECISION + 1
```

```
while (!all(diff< PRECISION) & (k < MAXITER)) {
```

```
  k <- k + 1
```

```
  param1=param
```

```
  sigmau <- param[-3]
```

```
  rho= param[3]
```

```
  G_12=matrix(c(sigmau[1],0,0,sigmau[2]),2,2)
```

```

GI=kronecker(G_12,I)
rho_W=solve(I-rho*W)
rho_Wt=t(rho_W)
V1=kronecker(d1omega,rho_W%*%rho_Wt)
V2=kronecker(d2omega,rho_W%*%rho_Wt)
V3=kronecker(G_12,rho_W%*%(W%*%rho_W%*%rho_Wt+rho_Wt%*%W%*%rho_Wt))
Omega1=kronecker(G_12,rho_W%*%rho_Wt) + R_12
Omega <- solve(Omega1)
Xto <- t(Omega %*% X)
Q <- solve(Xto %*% X)
P <- Omega - t(Xto) %*% Q %*% Xto
Py <- P %*% y
s1 <- (-0.5) * tr(P%*%V1) + 0.5 * (t(Py)%*%V1%*% Py)
s2 <- (-0.5) * tr(P%*%V2) + 0.5 * (t(Py)%*%V2%*% Py)
s3 <- (-0.5) * tr(P%*%V3) + 0.5 * (t(Py)%*%V3%*% Py)
f11 <- 0.5 * tr(P%*%V1%*%P%*%V1)
f12 <- 0.5 * tr(P%*%V1%*%P%*%V2)
f13 <- 0.5 * tr(P%*%V1%*%P%*%V3)
f22 <- 0.5 * tr(P%*%V2%*%P%*%V2)
f23 <- 0.5 * tr(P%*%V2%*%P%*%V3)
f33 <- 0.5 * tr(P%*%V3%*%P%*%V3)
s <- matrix(c(s1,s2,s3),3,1)
F <- matrix(c(f11,f12,f13,f12,f22,f23,f13,f23,f33),3,3)
param=param1 + solve(F)%*%s
diff <- abs((param - param1)/param1)
Param=c(Param,param[3])
if(abs(param[3])>1){
  param[3]=param1[3]
}
}

```

```

G_12=matrix(c(param[1],0,0,param[2]),2,2)

```

```

GI=kronecker(G_12,I)
sigmau <- param[-3]
rho= param[3]
rho_W=solve(I-rho*W)
rho_Wt=t(rho_W)
V1=kronecker(d1omega,rho_W%*%rho_Wt)
V2=kronecker(d2omega,rho_W%*%rho_Wt)
V3=kronecker(G_12,rho_W%*%(W%*%rho_W%*%rho_Wt+rho_Wt%*%W%*%rho_Wt))
Omega1=kronecker(G_12,rho_W%*%rho_Wt) + R_12
Omega <- solve(Omega1)
Xto <- t(Omega %*% X)
Qh <- solve(Xto %*% X)
P <- Omega - t(Xto) %*% Qh %*% Xto
Py <- P %*% y
s1 <- (-0.5) * tr(P%*%V1) + 0.5 * (t(Py)%*%V1%*% Py)
s2 <- (-0.5) * tr(P%*%V2) + 0.5 * (t(Py)%*%V2%*% Py)
s3 <- (-0.5) * tr(P%*%V3) + 0.5 * (t(Py)%*%V3%*% Py)
f11 <- 0.5 * tr(P%*%V1%*%P%*%V1)
f12 <- 0.5 * tr(P%*%V1%*%P%*%V2)
f13 <- 0.5 * tr(P%*%V1%*%P%*%V3)
f22 <- 0.5 * tr(P%*%V2%*%P%*%V2)
f23 <- 0.5 * tr(P%*%V2%*%P%*%V3)
f33 <- 0.5 * tr(P%*%V3%*%P%*%V3)
s <- matrix(c(s1,s2,s3),3,1)
F <- matrix(c(f11,f12,f13,f12,f22,f23,f13,f23,f33),3,3)
beta_12=Qh%*%Xto%*%y
res=y-X%*%beta_12

y_12=X%*%beta_12+kronecker(Id,rho_W)%*%GI%*%kronecker(Id,rho_Wt)%*%O
mega%*%res
se.b_1_2=sqrt(diag(Qh))
t_1_2=beta_12/se.b_1_2

```

```

FI=solve(F)
t_sigmau_=(sigmau[1]-sigmau[2])/sqrt(FI[1,1]+FI[2,2]-2*FI[1,2])
t_sig1=sigmau[1]/sqrt(FI[1,1])
t_sig2=sigmau[2]/sqrt(FI[2,2])
t_rho=rho/sqrt(FI[3,3])
pv <- 2 * pnorm(abs(t_1_2), lower.tail = FALSE)
prho <- 2 * pnorm(abs(t_rho), lower.tail = FALSE)
pv_sig <- pnorm(abs(t_sigmau_), lower.tail = FALSE)
pv_sig1 <- pnorm(abs(t_sig1), lower.tail = FALSE)
pv_sig2 <- pnorm(abs(t_sig2), lower.tail = FALSE)
d=kronecker(Id,I)-kronecker(G_12,rho_W%*%rho_Wt)%*%Omega
gg1=diag(kronecker(G_12,rho_W%*%rho_Wt)%*%Omega%*%R_12)
gg2=diag(d%*%X%*%Qh%*%t(X)%*%t(d))
dg1=V1%*%Omega-kronecker(G_12,rho_W%*%rho_Wt)%*%Omega%*%V1%*%Omega
dg2=V2%*%Omega-kronecker(G_12,rho_W%*%rho_Wt)%*%Omega%*%V2%*%Omega
dg3=V3%*%Omega-kronecker(G_12,rho_W%*%rho_Wt)%*%Omega%*%V3%*%Omega

g3_1=FI[1,1]*(dg1%*%Omega1%*%t(dg1))
g3_12=FI[1,2]*(dg1%*%Omega1%*%t(dg2))
g3_13=FI[1,3]*(dg1%*%Omega1%*%t(dg3))
g3_2=FI[2,2]*(dg2%*%Omega1%*%t(dg2))
g3_23=FI[2,3]*(dg2%*%Omega1%*%t(dg3))
g3_3=FI[3,3]*(dg3%*%Omega1%*%t(dg3))

gg3=diag(g3_1+g3_2+g3_3+2*g3_12+2*g3_13+2*g3_23)
mse1_2=gg1+gg2+2*gg3
MSE_y_1=mse1_2[1:n]
MSE_y_2=mse1_2[(n+1):(2*n)]
cbind(beta_12,se.b_1_2,t_1_2,pv)

```

```

# MFH

sigmau=as.matrix(c(median(var1),median(var2)),2,1)
k <- 0
diff <- PRECISION + 1
while (!all(diff< PRECISION) & (k < MAXITER)) {
  k <- k + 1
  sigmau1=sigmau
  G=matrix(c(sigmau[1],0,0,sigmau[2]),2,2)
  GI=kronecker(G,I)
  Omega <- solve(GI + R_12)
  Xto <- t(Omega %*% X)
  Q <- solve(Xto %*% X)
  P <- Omega - t(Xto) %*% Q %*% Xto
  Py <- P %*% y
  s1 <- (-0.5) * tr(P%*%d1Omega) + 0.5 * (t(Py)%*%d1Omega %*% Py)
  s2 <- (-0.5) * tr(P%*%d2Omega) + 0.5 * (t(Py)%*%d2Omega %*% Py)
  f11 <- 0.5 * tr(P%*%d1Omega%*%P%*%d1Omega)
  f12 <- 0.5 * tr(P%*%d1Omega%*%P%*%d2Omega)
  f22 <- 0.5 * tr(P%*%d2Omega%*%P%*%d2Omega)
  s <- matrix(c(s1,s2),2,1)
  F <- matrix(c(f11,f12,f12,f22),2,2)
  sigmau <- sigmau1 + solve(F)%*%s
  diff <- abs((sigmau - sigmau1)/sigmau1)
}

G_M=matrix(c(sigmau[1],0,0,sigmau[2]),2,2)
GI=kronecker(G_M,I)
Omega <- solve(GI + R_12)
Xto <- t(Omega %*% X)
Qh <- solve(Xto %*% X)
P <- Omega - t(Xto) %*% Qh %*% Xto

```

```

Py <- P %*% y
s1 <- (-0.5) * tr(P%*%d1Omega) + 0.5 * (t(Py)%*%d1Omega %*% Py)
s2 <- (-0.5) * tr(P%*%d2Omega) + 0.5 * (t(Py)%*%d2Omega %*% Py)
f11 <- 0.5 * tr(P%*%d1Omega%*%P%*%d1Omega)
f12 <- 0.5 * tr(P%*%d1Omega%*%P%*%d2Omega)
f22 <- 0.5 * tr(P%*%d2Omega%*%P%*%d2Omega)
s <- matrix(c(s1,s2),2,1)
F_M <- matrix(c(f11,f12,f12,f22),2,2)
beta_M=Qh%*%Xto%*%y
res_M=y-X%*%beta_M
y_M=X%*%beta_M+GI%*%Omega%*%res_M
se.b_M=sqrt(diag(Qh))
t_M=beta_M/se.b_M
FI_M=solve(F_M)
t_sigma_M=(sigma[1]-sigma[2])/sqrt(FI_M[1,1]+FI_M[2,2]-2*FI_M[1,2])
t_sig1_M=sigma[1]/sqrt(FI_M[1,1])
t_sig2_M=sigma[2]/sqrt(FI_M[2,2])
pv_M <- 2 * pnorm(abs(t_M), lower.tail = FALSE)
pv_sig_M <- pnorm(abs(t_sigma_M), lower.tail = FALSE)
pv_sig1_M <- pnorm(abs(t_sig1_M), lower.tail = FALSE)
pv_sig2_M <- pnorm(abs(t_sig2_M), lower.tail = FALSE)
d=kronecker(Id,I)-GI%*%Omega
gg1=diag(GI%*%Omega%*%R_12)
gg2=diag(d%*%X%*%Qh%*%t(X)%*%t(d))
dg1=d1Omega%*%Omega-GI%*%Omega%*%d1Omega%*%Omega
dg2=d2Omega%*%Omega-GI%*%Omega%*%d2Omega%*%Omega
g3_1=FI_M[1,1]*(dg1%*%(GI+R_12)%*%t(dg1))
g3_12=FI_M[1,2]*(dg1%*%(GI+R_12)%*%t(dg2))
g3_21=FI_M[2,1]*(dg2%*%(GI+R_12)%*%t(dg1))
g3_2=FI_M[2,2]*(dg2%*%(GI+R_12)%*%t(dg2))
gg3=diag(g3_1+g3_2+2*g3_12)
mse_M=gg1+gg2+2*gg3

```

```
MSE_y_1_M=mse_M[1:n]
```

```
MSE_y_2_M=mse_M[(n+1):(2*n)]
```