

ESTIMASI PRODUKTIVITAS PADI LEVEL KECAMATAN DI KABUPATEN TULUNGAGUNG MENGGUNAKAN GEOADDITIVE SAE

Estimation of Paddy Productivity at Sub-District Level in Tulungagung Regency Using Geoadditive SAE

Garinca Firgiana Santoso¹, Siti Muchlisoh²

^{1,2} Politeknik Statistika STIS
Jalan Otto Iskandardinata No.64C, Jakarta Timur (13330)
e-mail: 211709712@stis.ac.id

Abstrak

Data produktivitas tanaman padi merupakan salah satu data dasar sebagai tolok ukur pemerintah untuk menilai keberhasilan program swasembada pangan daerah. Data produktivitas padi digunakan untuk menghitung produksi padi di suatu wilayah. Pemerintah kabupaten/kota membutuhkan data produksi padi hingga level kecamatan untuk mengidentifikasi kondisi pemenuhan pangan daerah. Namun, pendugaan data produksi padi level kecamatan masih terkendala karena tidak tersedianya data produktivitas padi level kecamatan. BPS hanya menyajikan data produktivitas padi hingga level kabupaten/kota saja. Penelitian ini bertujuan untuk menduga produktivitas padi level kecamatan di Kabupaten Tulungagung tahun 2019 menggunakan *geoadditive small area estimation*, mengevaluasi kebaikan pendugaan menggunakan *Root Mean Square Error* (RMSE) dan *Relative Standard Error* (RSE), dan mengidentifikasi surplus-defisit beras menurut kecamatan. Metode analisis yang digunakan adalah analisis inferensia menggunakan estimasi tidak langsung dengan pendekatan *geoadditive SAE*. Hasil pendugaan menggunakan *geoadditive SAE* menunjukkan bahwa produktivitas padi tertinggi berada di Kecamatan Pucanglaban (8,8648 ton/ha), sedangkan produktivitas padi terendah di Kecamatan Pagerwojo (3,6576 ton/ha). Penggunaan model *geoadditive SAE* memberikan tingkat presisi lebih baik karena menghasilkan RMSE dan RSE lebih kecil daripada estimasi langsung. Hasil pendugaan ini juga menunjukkan bahwa sebagian besar kecamatan di Kabupaten Tulungagung mengalami surplus beras tahun 2019, tetapi ada enam kecamatan yang masih mengalami defisit beras.

Kata kunci: *geoadditive small area estimation*, produktivitas padi, surplus-defisit beras

Abstract

Paddy productivity data is one of the benchmarks for the government to audit the success of local food self-sufficiency program. Paddy productivity data is used to calculate paddy production in a region. Local government needs paddy production data at sub-district level to identify local food supply for the population. However, the estimation of paddy production data at sub-district level is constrained by the absence of paddy productivity data at sub-district level. BPS presents the data at regency level only. This research aims to estimate paddy productivity at sub-district level in Tulungagung Regency in 2019 using *geoadditive small area estimation*, evaluate the accuracy of the estimation using *Root Mean Square Error* (RMSE) and *Relative Standard Error* (RSE), and identify the rice surplus-deficit at sub-district level. Analysis method being used was inferential analysis using indirect estimation by *geoadditive SAE*. The estimation showed that the highest paddy productivity was in Pucanglaban Sub-district (8,8648 ton/ha), while the lowest paddy productivity in Pagerwojo Sub-district (3,6576 ton/ha). The use of *geoadditive SAE* gave more precision to the estimation because it produced smaller RMSE and RSE than direct estimation method. The estimation also showed that major sub-districts of Tulungagung Regency experienced surplus in rice during 2019, but there were also six sub-districts which suffered deficit in rice.

Keywords: *geoadditive small area estimation*, paddy productivity, rice surplus-deficit

PENDAHULUAN

Pasal 12 Ayat 2 UU No. 18 Tahun 2012 tentang Pangan menyebutkan bahwa pemerintah daerah bertanggung jawab atas ketersediaan pangan dan pengembangan produksi pangan lokal di daerah. Data produktivitas tanaman padi merupakan salah satu data dasar sebagai tolok ukur pemerintah untuk menilai keberhasilan program swasembada pangan daerah. Data produktivitas padi digunakan untuk menghitung produksi padi di suatu wilayah. Pada era otonomi daerah seperti saat ini, setiap daerah memiliki kewenangan untuk mengurus sendiri daerahnya sehingga pemerintah kabupaten/kota membutuhkan dukungan data produksi padi hingga level kecamatan untuk melihat kondisi pemenuhan pangan daerah. Namun, data produksi padi pada level kecamatan belum tersedia karena tidak tersedianya data produktivitas padi level kecamatan. BPS hanya menyajikan data produktivitas padi hingga level kabupaten/kota saja karena ukuran sampel Survei Ubinan tidak representatif untuk mengestimasi langsung hingga level kecamatan dan tidak semua kecamatan memiliki sampel. Jika estimasi langsung menggunakan ukuran sampel yang tidak representatif tetap dilakukan, penduga yang dihasilkan memiliki *standard error* besar dan tidak *reliable* (Ghosh & Rao, 1994).

Tidak tersedianya data produktivitas padi di level kecamatan menghambat evaluasi pemerintah kabupaten/kota terkait produksi beras di wilayahnya dan menyulitkan identifikasi kecamatan yang mengalami surplus-defisit beras. Ada dua solusi untuk menyelesaikan permasalahan tidak tersedianya data produktivitas padi level kecamatan. Pertama, estimasi langsung dengan cara menambah ukuran sampel hingga representatif untuk menyajikan data produktivitas padi di level kecamatan. Kedua, estimasi tidak langsung melalui pemodelan dengan melibatkan variabel penyerta yang dapat memberikan informasi terkait produktivitas padi. Namun, metode estimasi langsung tidak disarankan karena menambah biaya survei

dan beban kerja pengumpul data. Dengan demikian, kebutuhan data area kecil dapat diatasi dengan estimasi tidak langsung menggunakan *small area estimation* (SAE).

SAE adalah suatu teknik statistika untuk menduga parameter subpopulasi pada variabel yang menjadi perhatian dengan ukuran sampel kecil dan tidak memiliki presisi yang memadai (Rao & Molina, 2015). SAE memanfaatkan informasi tambahan secara statistik yang memiliki sifat meminjam kekuatan informasi dari hubungan antara variabel yang menjadi perhatian dan variabel penyerta. Kelebihan utama SAE yaitu meningkatkan efisiensi ukuran sampel yang terbatas dengan meminjam kekuatan yang diperoleh dari dalam area maupun luar area (Chand & Alexander, 1995).

Penelitian tentang pendugaan data produktivitas padi pada area kecil telah dilakukan oleh Ardiansyah (2018) menggunakan *geoadditive* SAE berbasis level unit. Penelitian tersebut bertujuan untuk menduga produktivitas padi per kecamatan di Kabupaten Seruyan Provinsi Kalimantan Tengah tahun 2016. Penelitian tersebut menggunakan metode *geoadditive* SAE karena terdapat hubungan nonlinier antara variabel penyerta dan variabel respon sehingga memerlukan pendekatan nonparametrik. Rata-rata RMSE hasil estimasi *geoadditive* SAE di seluruh kecamatan lebih kecil daripada RMSE estimasi langsung. Merujuk pada penelitian Ardiansyah (2018), penelitian ini dilakukan di Kabupaten Tulungagung Provinsi Jawa Timur.

Kabupaten Tulungagung merupakan salah satu kabupaten andalan di Provinsi Jawa Timur karena menghasilkan produksi padi cukup tinggi yaitu mencapai 196.430,58 ton pada 2019 (BPS Jawa Timur, 2020). Namun, produksi padi di Kabupaten Tulungagung tahun 2015-2019 cenderung menurun, sejalan dengan produksi padi di Jawa Timur. Selain itu, Kabupaten Tulungagung memiliki potensi cukup besar yang dapat dikembangkan untuk pertanian. Hal ini terindikasi dari 243 hektare lahan sementara tidak diusahakan pada 2019 (BPS Tulungagung, 2021).

Ketersediaan data produktivitas padi level kecamatan di Kabupaten Tulungagung dapat dijadikan salah satu pertimbangan dalam program pembukaan lahan pertanian, perencanaan, monitoring, dan evaluasi kebijakan pertanian agar mencapai target produksi padi.

Berdasarkan uraian di atas, penelitian ini bertujuan untuk menduga produktivitas padi level kecamatan di Kabupaten Tulungagung tahun 2019 menggunakan metode *geoadditive* SAE, mengevaluasi kebaikan pendugaan hasil estimasi langsung dan *geoadditive* SAE menggunakan RMSE dan RSE, serta mengidentifikasi surplus-defisit beras menurut kecamatan di Kabupaten Tulungagung tahun 2019.

METODE

Geoadditive SAE

Geoadditive SAE diperoleh dari gabungan model *geoadditive* dan SAE klasik dalam bentuk *linear mixed model*. *Geoadditive* SAE memiliki dua komponen pengaruh acak yaitu pengaruh khusus area (*area specific effect*) dan pengaruh spasial dari pemulus *spline-2* (Pusponegoro, Djuraidah, Fitrianto, & Sumertajaya, 2019).

Model *geoadditive* merupakan metode penggabungan dari model *kriging* dan model *additive* yang disajikan dalam bentuk *linear mixed model* (Kammann & Wand, 2003). Model *geoadditive* mampu menangani masalah pelanggaran asumsi normalitas dan linieritas. Model *geoadditive* mampu menganalisis data yang tidak berdistribusi normal sambil memperhitungkan kemungkinan hubungan variabel penyerta linier maupun nonlinier. Model ini memiliki fleksibilitas dalam menentukan bentuk hubungan antara variabel respon dan variabel penyerta karena model ini merupakan bentuk fungsi pemulusan (Pusponegoro, Djuraidah, Fitrianto, & Sumertajaya, 2019). Model *geoadditive* dapat dituliskan dalam bentuk *linear mixed model* (LMM) sebagai berikut:

$$y = X\beta + Zv + e \quad (1)$$

dimana:

$X = [\mathbf{1}, \mathbf{r}_i, \mathbf{t}_i, \mathbf{s}_i^T]_{1 \leq i \leq n}$ dimana r_i dan t_i merupakan prediktor kontinu dari y_i pada lokasi spasial s_i , $\beta = [\beta_0, \beta_r, \beta_t, \beta_s^T]$, $v = [v_1^s, \dots, v_{K_r}^s, v_1^t, \dots, v_{K_t}^t, v_1^s, \dots, v_{K_s}^s]$, β dan v merupakan vektor parameter, Z berisi matriks gabungan dari fungsi basis *spline*, $K_1^s, \dots, K_{K_s}^s$ adalah knot atau simpul lokasi spasial (Bocchi, 2009).

Knot dapat diartikan sebagai suatu titik fokus dalam *spline* sehingga kurva yang terbentuk tersegmentasi pada titik tersebut. Default pada sejumlah pengamatan unik yaitu ada 4-5 observasi di setiap knot. Kumpulan data yang besar dapat menyebabkan jumlah knot yang besar sehingga jumlah maksimum knot yang disarankan antara 20-40 knot (Ruppert, Wand, & Carroll, 2003). Salah satu metode pemilihan knot optimal yaitu menggunakan *Generalized Cross Validation* (GCV). Definisi GCV dapat dinotasikan sebagai berikut:

$$GCV(K) = \frac{MSE(K)}{[n^{-1} \text{trace}(I - A_{(K)})]^2} \quad (2)$$

dimana: $MSE(K) = n^{-1} \mathbf{y}^T (\mathbf{I} - \mathbf{A}_{(K)})^T (\mathbf{I} - \mathbf{A}_{(K)}) \mathbf{y}$, $K = (K_1^s, \dots, K_{K_s}^s)$ adalah titik knot, $\mathbf{A}_{(K)} = \mathbf{C}(\mathbf{C}^T \mathbf{C} + \lambda \mathbf{D})^{-1} \mathbf{C}^T$, $\mathbf{C} = [\mathbf{X} \ \mathbf{Z}]$, $\mathbf{D} = \text{diag}(\mathbf{0}_{d+1}, \mathbf{1}_K)$, $\mathbf{A}_{(K)}$ disebut sebagai matriks *smoothing* (Ardiansyah, 2018). Penelitian ini menggunakan metode *fixed selection method* untuk menentukan jumlah titik knot maksimum. Tujuan utama metode pemilihan knot yaitu memastikan bahwa jumlah knot (K) cukup besar agar lebih fleksibel ketika mengontrol kemulusan kurva yang diestimasi dengan *smoothing* parameter. Tujuan lainnya yaitu memilih jumlah knot (K) yang tidak terlalu besar agar waktu perhitungan tidak terlalu lama dan MSE kecil (Apriani, 2017). Rumus *fixed selection method* didefinisikan sebagai berikut (Ruppert, Wand, & Carroll, 2003):

$$K = \min \left(\frac{1}{4} \times \text{banyaknya } x_i \text{ yang unik}, 35 \right) \quad (3)$$

Sementara itu, penentuan lokasi knot menggunakan kuantil ke- K_k dari x_i yang unik. Rumus penentuan lokasi knot yaitu sebagai berikut (Ruppert, Wand, & Carroll, 2003):

$$K_k = \left(\frac{k+1}{K+2} \right) \quad (4)$$

Misalkan ada m area kecil yang akan diduga dan y_{ij} menunjukkan variabel respon dimana $i = 1, 2, \dots, m$ dan $j = 1, 2, \dots, n_i$ yang menunjukkan bahwa m adalah jumlah area kecil dalam suatu wilayah dan n_i adalah ukuran sampel pada masing-masing area kecil. Misalkan x_{ij} adalah vektor variabel penyerta yang bersesuaian dengan unit respon, maka model klasik SAE yaitu:

$$y_{ij} = x_{ij}^T \beta + u_i + e_{ij} \quad (5)$$

dimana u_i adalah pengaruh acak area (*random effect area*) yang berdistribusi normal $u_i \stackrel{iid}{\sim} N(0, \sigma_u^2)$, e_{ij} adalah error pada sampel (*individual level random error*) di area kecil- i dan unit sampel- j yang berdistribusi $e_{ij} \stackrel{iid}{\sim} N(0, \sigma_e^2)$, serta u_i dan e_{ij} saling independen baik antar individu maupun antar wilayah (Rao & Molina, 2015). Jika didefinisikan matriks $D = [d_{ij}]$, dimana:

$$d_{ij} = \begin{cases} 1 & ; \text{jika sampel } j \text{ berada di} \\ & \text{dalam area kecil ke } - i \\ 0 & ; \text{untuk lainnya} \end{cases} \quad (6)$$

dengan $y = [y_{ij}]$, $X = [x_{ij}^T]$, dan $e = [e_{ij}]$, maka bentuk matriks dari persamaan (5) yaitu:

$$y = X\beta + Du + e \quad (7)$$

dimana $E \begin{bmatrix} u \\ e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $Cov \begin{bmatrix} u \\ e \end{bmatrix} = \begin{bmatrix} \sigma_u^2 I_m & 0 \\ 0 & \sigma_e^2 I_n \end{bmatrix}$. Dengan demikian, matriks ragam-peragam dari y adalah $\Omega = \sigma_u^2 DD^T + \sigma_e^2 I_n$ (Bocchi, 2009).

Model dengan *spline* maupun model SAE merupakan model yang memiliki pengaruh acak, maka sangat

memungkinkan untuk menggabungkan kedua konsep tersebut dalam *nonparametric small area estimation framework* berdasarkan bentuk *linear mixed model* (Kurnia, 2009). Model *geoaddivitive SAE* (Geo-SAE) memiliki dua pengaruh acak yaitu pengaruh acak spasial dari pemulus *spline-2* (*spatial random effect*) dan pengaruh acak area kecil (*area specific random effect*) yang didefinisikan sebagai berikut:

$$y = X\beta + Zv + Du + e \quad (8)$$

dimana :

- $X = [1, x_{ij}^T, s_{ij}^T]_{1 \leq i \leq m; 1 \leq j \leq n_i}$ adalah matriks pengaruh tetap berukuran $n \times (p + 1)$, p adalah jumlah variabel penyerta atau pengaruh tetap;
- β yaitu vektor koefisien pengaruh tetap (*fixed effect*) berukuran $(p + 1) \times 1$;
- Z adalah matriks dari *spline-2* (*thin-plate spline*) berukuran $n \times K$ menggunakan basis fungsi radial yaitu $Z = [C(s_{ij} - K_k)]_{1 \leq i \leq m; 1 \leq j \leq n_i; 1 \leq k \leq K}$ dengan $C(s) = \|s\|^2 \log \|s\|$, dimana K_k merupakan knot, s_{ij} merupakan koordinat lokasi spasial, $\|\cdot\|$ adalah notasi dari *Euclidean norm*;
- v adalah vektor koefisien *spline-2* dari variabel spasial sebagai pengaruh acak spasial (*spatial random effect*) berukuran $K \times 1$;
- D adalah matriks dari pengaruh acak area kecil berukuran $n \times m$ yang didefinisikan pada persamaan (6);
- u adalah vektor koefisien pengaruh acak area kecil (*area specific random effect*) berukuran $m \times 1$;
- e adalah error acak level unit (*individual level random error*);
- $E \begin{bmatrix} v \\ u \\ e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ dan $Cov \begin{bmatrix} v \\ u \\ e \end{bmatrix} = \begin{bmatrix} \sigma_v^2 I_K & 0 & 0 \\ 0 & \sigma_u^2 I_m & 0 \\ 0 & 0 & \sigma_e^2 I_n \end{bmatrix}$ (Bocchi, Petrucci, & Rocco, 2016).

Komponen ragam atau varians yang tidak diketahui dapat diduga menggunakan

REML sehingga diperoleh $\hat{\sigma}_v^2$, $\hat{\sigma}_u^2$, dan $\hat{\sigma}_e^2$ (Bocci, 2009). Dugaan matriks ragam-peragam dari $\mathbf{y}$ yaitu:

$$\hat{\Omega} = \hat{\sigma}_v^2 \mathbf{Z}\mathbf{Z}^T + \hat{\sigma}_u^2 \mathbf{D}\mathbf{D}^T + \hat{\sigma}_e^2 \mathbf{I}_n \quad (9)$$

Kemudian, penduga EBLUP bagi β , v , dan u adalah:

$$\hat{\beta} = (\mathbf{X}^T \hat{\Omega}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \hat{\Omega}^{-1} \mathbf{y} \quad (10)$$

$$\hat{v} = \hat{\sigma}_v^2 \mathbf{Z}^T \hat{\Omega}^{-1} (\mathbf{y} - \mathbf{X}\hat{\beta}) \quad (11)$$

$$\hat{u} = \hat{\sigma}_u^2 \mathbf{D}^T \hat{\Omega}^{-1} (\mathbf{y} - \mathbf{X}\hat{\beta}) \quad (12)$$

Misalkan terdapat area kecil i , pendugaan rata-rata dari y dapat menggunakan formula berikut $\bar{y}_i = \bar{\mathbf{X}}_i \beta + \bar{\mathbf{z}}_i v + u_i$ dimana $\bar{\mathbf{X}}_i$ dan $\bar{\mathbf{z}}_i$ adalah nilai parameter pada level area kecil ke- i dan diasumsikan diketahui. Penduga EBLUP untuk rata-rata dari variabel respon yaitu:

$$\hat{y}_i = \bar{\mathbf{X}}_i \hat{\beta} + \bar{\mathbf{z}}_i \hat{v} + d_i \hat{u} \quad (13)$$

dimana d_i vektor bernilai 1 jika di dalam area- i dan bernilai 0 jika di luar area- i (Petrucci & Pratesi, 2014).

MSE Parametric Bootstrap

Salah satu cara untuk melihat presisi dari estimator adalah dengan menghitung *mean squared error*. MSE juga digunakan sebagai alat untuk membandingkan antar metode agar terpilih metode terbaik dan paling akurat (Biemer & Lyberg, 2003). Mengikuti penelitian Opsomer et al. (2008), penelitian ini menggunakan *parametric bootstrap* untuk mengestimasi MSE dari penduga rata-rata area kecil. Pendekatan *bootstrap* merupakan penghitungan MSE dengan memanfaatkan teknik *resampling*. Metode *parametric bootstrap* mengasumsikan bahwa pengaruh acak dan komponen error pada *geoadditive SAE* berdistribusi normal $u_i \stackrel{iid}{\sim} N(0, \sigma_u^2)$, $v_i \stackrel{iid}{\sim} N(0, \sigma_v^2)$, dan $e_i \stackrel{iid}{\sim} N(0, \sigma_e^2)$ dimana $i = 1, 2, \dots, m$ dan $j = 1, 2, \dots, n_i$ (Ardiansyah, 2018).

Prosedur *parametric bootstrap* dilakukan dengan cara membangkitkan data sampel *bootstrap* yaitu $\{(y_{ij}^*, x_{ij}); i = 1, 2, \dots, m \text{ dan } j = 1, 2, \dots, n_i\}$ (Rao &

Molina, 2015). Data sampel bangkitan dari *bootstrap* didefinisikan sebagai berikut:

$$y_{ij}^* = x_{ij}^T \hat{\beta} + \bar{\mathbf{z}}_i v_i^* + u_i^* + e_{ij}^* \quad (14)$$

dimana pengaruh acak v_i^* dibangkitkan dari $N(0, \sigma_v^2)$, u_i^* dibangkitkan dari $N(0, \sigma_u^2)$, dan e_{ij}^* dibangkitkan dari $N(0, \sigma_e^2)$ (Ardiansyah, 2018). Misalkan $\mu_i^* = \bar{\mathbf{X}}_i^T \hat{\beta}^* + \bar{\mathbf{z}}_i v_i^* + u_i^*$ menjadi versi *bootstrap* dari target parameter $\mu_i = \bar{\mathbf{X}}_i^T \hat{\beta} + \bar{\mathbf{z}}_i v_i + u_i$. $\bar{\mathbf{X}}_i$ merupakan nilai parameter dari variabel penyerta pada level area kecil ke- i , sedangkan $\bar{\mathbf{z}}_i$ merupakan nilai *spline-2* pada level area kecil ke- i . Dengan menggunakan data *bootstrap*, versi *bootstrap* dari estimator *geoadditive SAE* ($\hat{\mu}_i^G$) yaitu:

$$\hat{\mu}_i^{*G} = \bar{\mathbf{X}}_i^T \hat{\beta}^* + \bar{\mathbf{z}}_i \hat{v}_i^* + \hat{u}_i^* \quad (15)$$

dimana $\hat{\beta}^*$, $\hat{v}_i^*$, dan $\hat{u}_i^*$ dihitung seperti penghitungan $\hat{\beta}$, $\hat{v}_i$, dan $\hat{u}_i$ pada persamaan (10), (11), (12) tetapi menggunakan data sampel bangkitan *bootstrap*. Secara teori, MSE *bootstrap* dari estimator *geoadditive SAE* yang diiterasi sebanyak B didefinisikan sebagai berikut:

$$MSE_B(\hat{\mu}_i^G) = E_*(\hat{\mu}_i^{*G} - \mu_i^*)^2 = \frac{1}{B} \sum_{b=1}^B (\hat{\mu}_i^{*G}(b) - \mu_i^*(b))^2 \quad (16)$$

dimana $MSE_B(\hat{\mu}_i^G) \approx g_{1i}(\hat{\sigma}_v^2, \hat{\sigma}_u^2) + g_{2i}(\hat{\sigma}_v^2, \hat{\sigma}_u^2) + g_{3i}(\hat{\sigma}_v^2, \hat{\sigma}_u^2)$ (Ardiansyah, 2018). Root Mean Square Error (RMSE) dan Relative Standard Error (RSE) untuk setiap area kecil ke- i dihitung menggunakan rumus berikut:

$$RMSE(\hat{\mu}_i^G) = \sqrt{MSE_B(\hat{\mu}_i^G)} \quad (17)$$

$$RSE(\hat{\mu}_i^G) = \frac{RMSE(\hat{\mu}_i^G)}{\hat{\mu}_i^G} \quad (18)$$

Produktivitas Tanaman Padi (GKP)

Produktivitas tanaman padi merupakan suatu nilai yang menunjukkan rata-rata hasil produksi tanaman padi per satuan luas panen pada periode satu laporan. Misalkan berat hasil ubinan dalam satuan kilogram (kg), luas plot ubinan yang

Tabel 1. Variabel Penelitian dan Sumber Data

Variabel	Nama Variabel	Sumber Data
Y	Produktivitas tanaman padi	<i>Raw data</i> Survei Ubinan 2019
X_1	Proporsi luas panen padi sawah terhadap total luas panen padi	Data level unit : <i>Raw data</i> Survei Ubinan 2019 Data populasi tingkat kecamatan : Dinas Pertanian Kabupaten Tulungagung tahun 2019
X_2	<i>Latitude</i>	Data level unit : <i>Raw data</i> Survei Ubinan 2019 Data populasi tingkat kecamatan : <i>Google maps</i>
X_3	<i>Longitude</i>	

digunakan 6,25 m², maka produktivitas padi dijabarkan sebagai berikut:

$$\text{Produktivitas Padi} = \text{berat hasil ubinan} \times \frac{1 \text{ hektare}}{\text{luas ubinan}} \quad (19)$$

Ruang Lingkup dan Sumber Data

Penelitian ini mencakup seluruh kecamatan di Kabupaten Tulungagung Provinsi Jawa Timur tahun 2019. Ada dua sumber data yang digunakan yaitu *raw data* Survei Ubinan 2019 dari BPS dan data dari Dinas Pertanian Kabupaten Tulungagung tahun 2019. Rincian data penelitian disajikan pada tabel 1.

Tahapan Analisis

Analisis dalam penelitian ini berfokus pada estimasi produktivitas padi level kecamatan menggunakan estimasi tidak langsung pendekatan model *geoadditive* SAE berbasis level unit. Pengolahan data pada penelitian ini menggunakan *software* R Studio versi 4.0.2 dan Microsoft Excel 2013. Berikut tahapan analisis:

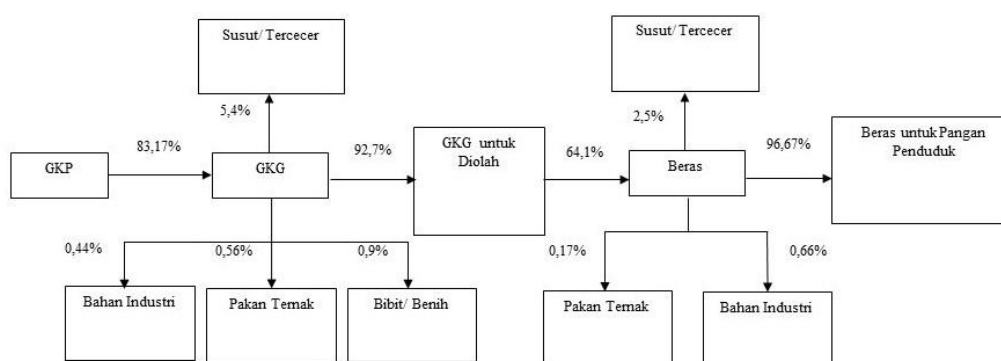
1. Tahap persiapan data
 - a. Membentuk variabel produktivitas padi pada *raw data* Survei Ubinan 2019 menggunakan formula (19).
2. Tahap estimasi tidak langsung menggunakan model *geoadditive* SAE
 - a. Menentukan variabel penyerta yang memiliki hubungan nonlinier terhadap variabel respon menggunakan *scatterplot*.
 - b. Menentukan jumlah knot maksimum menggunakan *fixed selection method* pada formula (3).
 - c. Menghitung nilai GCV untuk masing-masing knot berdasarkan rumus pada formula (2).
 - d. Menentukan jumlah knot optimum yang digunakan berdasarkan GCV minimum.
 - e. Menghitung matriks Z dari *spline-2* yang tertera pada keterangan formula (8).
 - f. Menyatakan bentuk area kecil dengan mengikuti suatu fungsi model *geoadditive* SAE yaitu:
 $y = X\beta + Zv + Du + e$.
 - g. Mengestimasi nilai β, v , dan u menggunakan formula (10), (11), dan (12).
 - h. Pada kecamatan nirsampel, pengaruh acak v_i dan u_i dari area nirsampel disubstitusi menggunakan v_i dan u_i dari area- i yang memiliki kemiripan terdekat berdasarkan variabel penyerta.
 - i. Melakukan estimasi rata-rata produktivitas padi tingkat kecamatan di Kabupaten Tulungagung menggunakan model *geoadditive* SAE yaitu $\hat{y}_i = \bar{x}_i \hat{\beta} + \bar{z}_i \hat{v} + d_i \hat{u}$.
 - j. Menghitung MSE hasil estimasi produktivitas padi untuk setiap kecamatan menggunakan metode *parametric bootstrap* yang tertera pada formula (16).
 - k. Menghitung RMSE dan RSE hasil estimasi produktivitas padi untuk setiap kecamatan sesuai formula (17) dan (18).
3. Tahap analisis lanjutan hasil estimasi tidak langsung

- Menghitung produksi padi di setiap kecamatan di kabupaten Tulungagung.
- Mengonversi padi dalam bentuk gabah kering gilig (GKG) ke beras untuk mendapatkan angka produksi beras.
- Mengidentifikasi surplus dan defisit beras di setiap kecamatan di Kabupaten Tulungagung.

HASIL DAN PEMBAHASAN

Secara geografis, Kabupaten Tulungagung terletak di posisi $7^{\circ} 51'$ hingga $8^{\circ} 18'$ Lintang Selatan dan $111^{\circ} 43'$ hingga $112^{\circ} 07'$ Bujur Timur. Luas wilayah Kabupaten Tulungagung mencapai $1.055,65 \text{ km}^2$ terdiri atas 19 kecamatan, 257 desa, dan 14 kelurahan (BPS Tulungagung, 2021). Luas panen padi di Kabupaten Tulungagung tahun 2019 didominasi oleh padi yang ditanam di lahan sawah. Berdasarkan data Dinas Pertanian Kabupaten Tulungagung, luas panen padi sawah tahun 2019 mencapai 41.343 hektare atau sebesar 93,51 persen dari total luas panen padi.

1. Pendugaan Produktivitas Padi Level Kecamatan Menggunakan Metode *Geoadditive* SAE



Gambar 1. Peta tutupan lahan menggunakan GEE platform dari Landsat 8 di Kota Jakarta Utara

Tabel 2. Jumlah Knot dan Nilai GCV

Jumlah Knot (K)	GCV untuk <i>Latitude</i>	GCV untuk <i>Longitude</i>
K = 10	1,9708	2,6802
K = 20	1,8385	2,6384
K = 29	1,7518	2,6059
K = 30	1,7497	2,6034

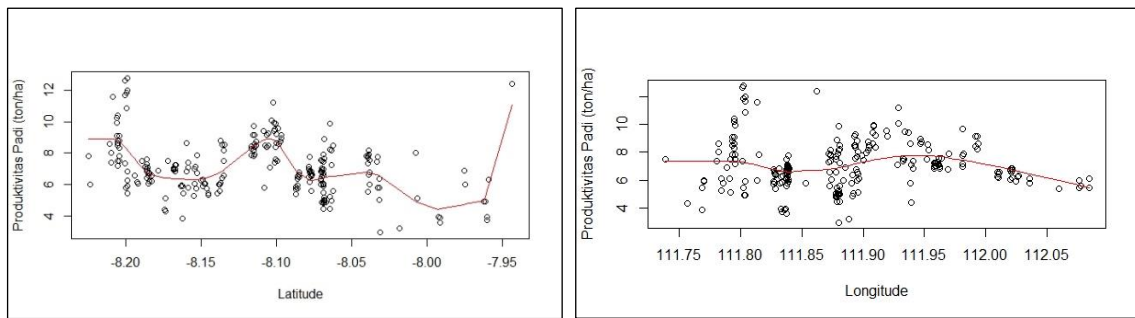
Eksplorasi Hubungan Antar Variabel

Estimasi tidak langsung dimulai dengan melakukan eksplorasi pola hubungan antara variabel respon dan variabel penyerta. Gambar 2 menunjukkan bahwa pola hubungan yang terbentuk antara *latitude* dan *longitude* terhadap produktivitas padi tidak mengikuti pola hubungan linier, tetapi memiliki perilaku pola data yang berubah-ubah. Dengan demikian, variabel *latitude* dan *longitude* merupakan variabel nonparametrik.

Kemudian, dilakukan pengecekan asumsi normalitas pada variabel respon. Berdasarkan uji *shapiro-wilk*, *pvalue* yang dihasilkan yaitu $5,781 \times 10^{-5}$. Pada taraf signifikansi 5 persen, sampel sudah cukup membuktikan bahwa data produktivitas padi tidak berdistribusi normal. Dengan demikian, metode yang digunakan pada penelitian ini yaitu *geoadditive* SAE yang mampu menangani pelanggaran asumsi normalitas dan linieritas.

Penentuan Titik Knot Optimum

Pemodelan pendugaan area kecil menggunakan *geoadditive* SAE dimulai dengan menentukan jumlah knot dan lokasi knot. Berdasarkan *fixed selection method*, jumlah knot maksimum penelitian ini yaitu 35 knot. Setiap knot yang telah ditentukan



Gambar 2. Peta tutupan lahan menggunakan GEE platform dari Landsat 8 di Kota Jakarta Utara

Tabel 3. Hasil Estimasi Parameter Model Geoadditive SAE

Kecamatan	Fixed Effects	Random Effects	
		$\bar{z}_i\hat{v}$	$d_i\hat{u}$
Besuki		0,3786	$2,9943 \times 10^{-9}$
Bandung		0,4121	$-5,0094 \times 10^{-10}$
Campurdarat		0,3209	$-2,3050 \times 10^{-9}$
Kalidawir		-0,3744	$-9,2398 \times 10^{-9}$
Pucanglaban		-1,0345	$-4,2597 \times 10^{-9}$
Rejotangan		-1,5501	$-6,6024 \times 10^{-9}$
Ngunut	$\hat{\beta}_0 = -1238,8551^*$	-0,8882	$2,6260 \times 10^{-9}$
Sumbergempol	$\hat{\beta}_1 = 1,6175^*$	-0,3438	$4,9167 \times 10^{-9}$
Boyolangu	$\hat{\beta}_2 = -11,8135^*$	0,0305	$5,4822 \times 10^{-9}$
Tulungagung	$\hat{\beta}_3 = 10,2668^*$	-0,0042	$-1,4855 \times 10^{-9}$
Kedungwaru		-0,4320	$1,9692 \times 10^{-9}$
Karangrejo		-0,6550	$7,4577 \times 10^{-10}$
Kauman		-0,2590	$2,1631 \times 10^{-9}$
Gondang		0,1904	$-1,0469 \times 10^{-9}$
Pagerwojo		-0,8223	$2,3682 \times 10^{-9}$
Sendang		-0,7983	$4,5086 \times 10^{-9}$

Keterangan : * signifikan pada taraf signifikansi 5 persen

dari hasil *fixed selection method* dihitung nilai GCVnya. Nilai GCV setiap knot di penelitian ini disajikan pada tabel 2.

Berdasarkan GCV minimum, jumlah knot optimum yang digunakan yaitu 30 knot. Nilai GCV minimum yang diperoleh adalah 1,7497 untuk *latitude* dan 2,6034 untuk *longitude*. Pola hubungan variabel *latitude* dan *longitude* terhadap produktivitas padi ketika didekati *spline-2* menggunakan 30 knot disajikan pada gambar 2.

Pendugaan Parameter Model

Model *geoadditive* SAE terdiri atas pengaruh tetap dan dua pengaruh acak. Hasil estimasi parameter pada model *geoadditive* SAE disajikan pada tabel 3.

Tabel 3 menunjukkan bahwa semua variabel penyerta yang digunakan memiliki

pvalue kurang dari taraf signifikansi 5 persen. Hal ini berarti ketiga variabel penyerta yang digunakan berpengaruh signifikan terhadap produktivitas padi. Variabel penyerta ini menunjukkan pengaruh tetap pada model. Model *geoadditive* SAE memiliki dua pengaruh acak yaitu pengaruh acak spasial dari pemulus *spline-2* dan pengaruh acak area kecil. $\bar{z}_i\hat{v}$ menunjukkan pengaruh acak spasial dari pemulus *spline-2* sedangkan $d_i\hat{u}$ menunjukkan pengaruh acak area kecil. Setiap kecamatan memiliki nilai estimasi pengaruh acak berbeda. Ada dua varians pengaruh acak yang diestimasi yaitu varians pengaruh acak spasial dari pemulus *spline-2* (σ_v^2) dan varians pengaruh acak area kecil (σ_u^2). Estimasi terhadap σ_v^2 dan σ_u^2 dilakukan menggunakan metode *restricted maximum likelihood* (REML) sehingga

didapatkan $\hat{\sigma}_v^2 = 34,4070$ dan $\hat{\sigma}_u^2 = 5,5941 \times 10^{-10}$.

Hasil Estimasi Produktivitas Padi Level Kecamatan Menggunakan Geoadditive SAE

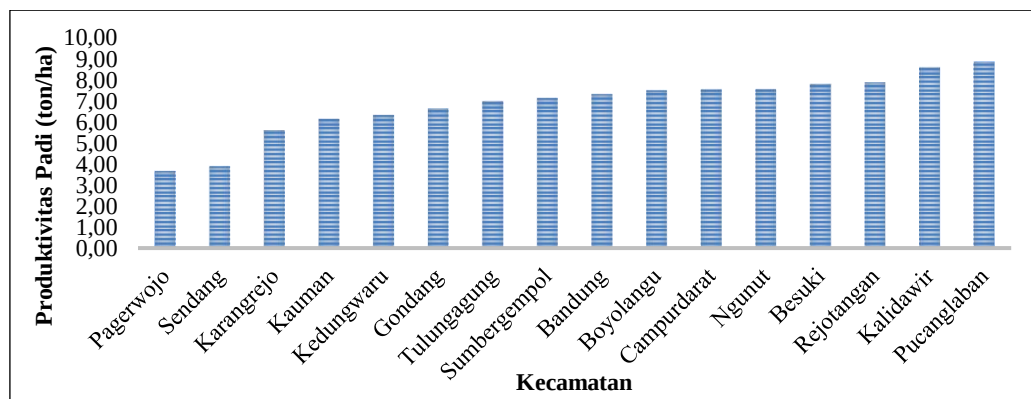
Gambar 3 menunjukkan bahwa Kecamatan Pucanglaban memiliki angka produktivitas padi terbesar yaitu 8,8648 ton/ha sedangkan Kecamatan Pagerwojo memiliki angka produktivitas padi terkecil yaitu 3,6576 ton/ha. Jika dikaitkan dengan luas panen, Kecamatan Pucanglaban merupakan kecamatan yang memiliki luas panen kategori rendah tetapi Kecamatan Pucanglaban justru menghasilkan produktivitas padi paling tinggi. Hal ini mengindikasikan bahwa Kecamatan Pucanglaban memiliki manajemen usaha pertanian lebih baik daripada kecamatan lain. Rata-rata produktivitas padi level kecamatan di Kabupaten Tulungagung tahun 2019 menggunakan metode *geoadditive* SAE sebesar 6,8364 ton/ha.

Kabupaten Tulungagung memiliki 19 kecamatan, tetapi terdapat 3 kecamatan yang tidak memiliki sampel pada Survei Ubinan 2019. Tiga kecamatan nirsampel meliputi Pakel, Tanggunggunung, dan

Ngantru. Ketiga kecamatan tersebut diestimasi menggunakan *geoadditive* SAE dengan cara meminjam pengaruh acak v_i dan u_i pada area- i yang memiliki kemiripan terdekat dengan area lain berdasarkan variabel penyerta. Ukuran kemiripan terdekat ini dihitung menggunakan *minimum euclidean distance* berdasarkan variabel penyerta yang digunakan (Ardiansyah, 2018). Berdasarkan ukuran kemiripan *euclidean distance*, Kecamatan Pakel dan Tanggunggunung memiliki kemiripan terdekat dengan Kecamatan Campurdarat, sedangkan Kecamatan Ngantru memiliki kemiripan terdekat dengan Kecamatan Karangrejo.

2. Evaluasi Pendugaan Produktivitas Padi Level Kecamatan Hasil Estimasi Langsung dan Geo-SAE

Evaluasi terhadap kebaikan pendugaan produktivitas padi level kecamatan dilakukan dengan membandingkan nilai RMSE dan RSE antara model *geoadditive* SAE dan metode estimasi langsung. Gambar 4 dan 5 menunjukkan bahwa secara keseluruhan RMSE dan RSE di setiap kecamatan mengalami reduksi ketika menggunakan

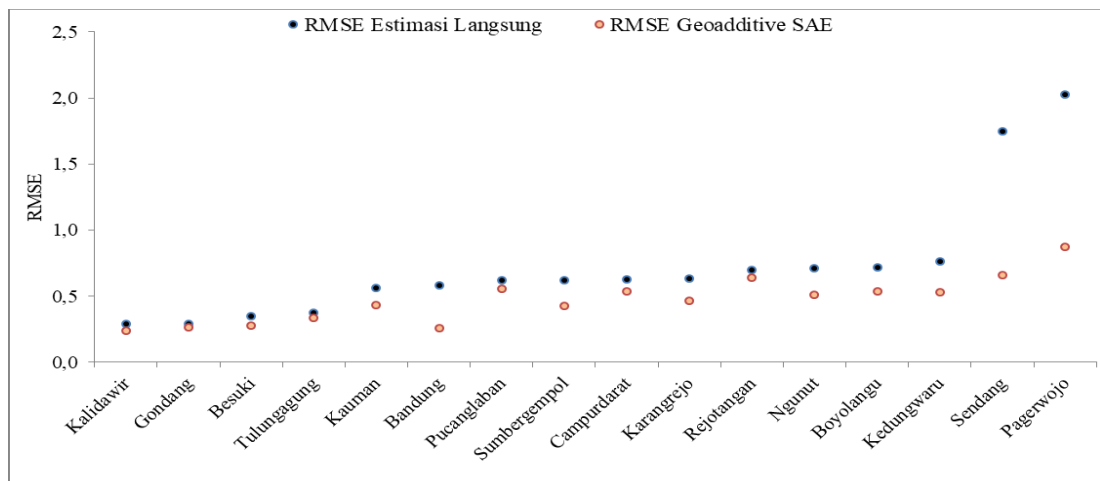


Sumber: Survei Ubinan 2019 dan Dinas Pertanian Kabupaten Tulungagung (diolah)

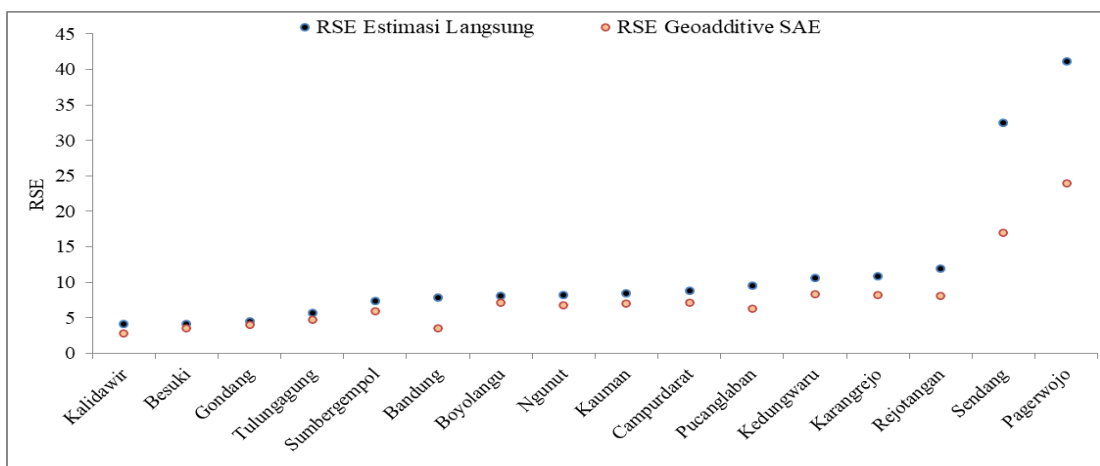
Gambar 3. Hasil Estimasi Produktivitas Padi Level Kecamatan di Kabupaten Tulungagung Tahun 2019 Menggunakan Model Geoadditive SAE.

Tabel 4. Hasil Estimasi Pengaruh Acak dan Produktivitas Padi di Kecamatan Nirsampel

Kecamatan Nirsampel	Random Effects		Produktivitas Padi (ton/ha)
	$\bar{z}_i \hat{u}$	$d_i \hat{v}$	
Pakel	0,3209	$-2,31 \times 10^{-9}$	6,9860
Tanggunggunung	0,3209	$-2,31 \times 10^{-9}$	6,8451
Ngantru	-0,6550	$7,46 \times 10^{-10}$	6,5578



Gambar 4. Perbandingan RMSE Estimasi Langsung dan Geoadditive SAE Menurut Kecamatan di Kabupaten Tulungagung 2019



Gambar 5. Perbandingan RMSE Estimasi Langsung dan Geoadditive SAE Menurut Kecamatan di Kabupaten Tulungagung 2019

metode *geoadditive* SAE. Rata-rata RMSE dari estimasi langsung yaitu 0,7259, sedangkan rata-rata RMSE dari *geoadditive* SAE yaitu 0,4705. Kemudian, rata-rata RSE dari estimasi langsung yaitu 11,4678 persen sedangkan rata-rata RSE dari *geoadditive* SAE yaitu 7,7799 persen. Rata-rata RMSE dan RSE dari metode *geoadditive* SAE lebih kecil daripada estimasi langsung. Hal ini menunjukkan bahwa hasil estimasi produktivitas padi level kecamatan di Kabupaten Tulungagung tahun 2019 menggunakan metode *geoadditive* SAE memiliki presisi lebih baik daripada estimasi langsung.

3. Kondisi Surplus-Defisit Beras Menurut Kecamatan di Kabupaten Tulungagung 2019

Produktivitas padi level kecamatan hasil estimasi *geoadditive* SAE digunakan untuk menghitung produksi padi di setiap kecamatan. Produksi padi merupakan perkalian antara produktivitas padi dan luas panen padi. Hasil penghitungan produksi padi ini memiliki satuan ton dalam bentuk Gabah Kering Panen (GKP), sedangkan data produksi padi yang digunakan oleh BPS dalam bentuk Gabah Kering Gilig (GKG). Oleh karena itu, nilai produksi padi harus dikonversi menggunakan angka konversi GKP ke GKG dari Survei Konversi Gabah ke Beras 2018 yaitu 83,17 persen untuk Jawa Timur (BPS RI, 2018). Alur penghitungan dan konversi dari produksi padi bentuk GKG menjadi beras untuk pangan penduduk disajikan pada gambar 1.

Tabel 5. Hasil Produksi Padi, Produksi Beras, Konsumsi Beras, dan Kondisi Surplus-Defisit Beras Menurut Kecamatan di Kabupaten Tulungagung Tahun 2019

Kecamatan	Produksi Padi (ton GKG)	Produksi Beras (ton)	Produksi Beras untuk Pangan Penduduk (ton)	Penduduk (jiwa)	Konsumsi Beras (ton)	Keterangan
Besuki	12.874,27	7.649,98	7.395,24	39.229	4.272,04	Surplus
Bandung	15.795,94	9.386,06	9.073,51	49.147	5.352,11	Surplus
Pakel	18.168,67	10.795,95	10.436,45	55.774	6.073,79	Surplus
Campurdarat	7.381,34	4.386,05	4.239,99	59.437	6.472,69	Defisit
Tanggunggunung	85,40	50,74	49,05	26.360	2.870,60	Defisit
Kalidawir	23.406,42	13.908,26	13.445,11	76.665	8.348,82	Surplus
Pucanglaban	2.359,32	1.401,93	1.355,24	25.919	2.822,58	Defisit
Rejotangan	28.049,43	16.667,17	16.112,15	81.979	8.927,51	Surplus
Ngunut	18.899,09	11.229,97	10.856,01	84.438	9.195,30	Surplus
Sumbergempol	16.702,46	9.924,72	9.594,22	74.189	8.079,18	Surplus
Boyolangu	16.970,15	10.083,78	9.747,99	85.500	9.310,95	Surplus
Tulungagung	4.775,76	2.837,79	2.743,29	67.513	7.352,16	Defisit
Kedungwaru	1.266,66	752,66	727,59	93.984	10.234,86	Defisit
Ngantru	6.086,80	3.616,82	3.496,38	58.384	6.358,02	Defisit
Karangrejo	12.063,95	7.168,49	6.929,77	44.981	4.898,43	Surplus
Kauman	13.635,90	8.102,55	7.832,73	53.940	5.874,07	surplus
Gondang	14.352,62	8.528,43	8.244,43	61.443	6.691,14	Surplus
Pagerwojo	10.948,16	6.505,48	6.288,84	32.173	3.503,64	Surplus
Sendang	16.951,21	10.072,53	9.737,11	47.759	5.200,96	Surplus
Total	240.773,56	143.069,33	138.305,12	1.118.814	121.838,84	Surplus

Identifikasi surplus-defisit beras di setiap kecamatan membutuhkan nilai konsumsi beras penduduk. Kecamatan dikatakan surplus beras apabila produksi beras di kecamatan tersebut lebih banyak dibanding konsumsi berasnya, sedangkan kecamatan dikatakan defisit beras apabila produksi beras di kecamatan tersebut lebih sedikit dibanding konsumsi berasnya (Ardiansyah, 2018). Konsumsi beras merupakan perkalian antara jumlah penduduk dan konsumsi beras per kapita. Penelitian ini menggunakan angka konsumsi beras per kapita dalam satu tahun berdasarkan Kajian Konsumsi Bahan Pokok yaitu 108,90 kg untuk Jawa Timur (BPS RI, 2018). Konsumsi beras penduduk dibandingkan dengan produksi beras untuk pangan penduduk untuk mengetahui kondisi surplus-defisit beras di suatu wilayah.

Produksi padi tertinggi berada di Kecamatan Rejotangan sebesar 28.049,43 ton GKG, sedangkan produksi padi terendah berada di Tanggunggunung

sebesar 85,40 ton GKG. Kecamatan Rejotangan menghasilkan produksi padi tertinggi karena memiliki luas panen padi terluas kedua dan angka estimasi produktivitas padi tertinggi ketiga. Sementara itu, Kecamatan Tanggunggunung menghasilkan produksi padi terendah karena memiliki luas panen padi terendah meskipun angka produktivitas padi di kecamatan ini cukup tinggi. Total produksi beras di Kabupaten Tulungagung 2019 adalah 143.069,33 ton sedangkan total produksi beras untuk pangan penduduk setelah mengalami susut/tercecer adalah 138.305,12 ton.

Jumlah penduduk Kabupaten Tulungagung tahun 2019 yaitu 1.118.814 jiwa. Dengan demikian, konsumsi beras penduduk dalam satu tahun di Kabupaten Tulungagung tahun 2019 diperkirakan sekitar 121.838,84 ton. Apabila konsumsi beras dibandingkan dengan produksi beras untuk pangan penduduk, Kabupaten Tulungagung mengalami surplus beras sebesar 16.466,28 ton pada tahun 2019.

Adanya surplus beras di Kabupaten Tulungagung ini dapat di alokasikan ke kabupaten/kota lain yang mengalami defisit beras. Hasil pendugaan ini menunjukkan bahwa ada 6 kecamatan yang mengalami defisit beras. Kecamatan yang mengalami defisit beras tahun 2019 meliputi Campurdarat, Tanggunggunung, Pucanglaban, Tulungagung, Kedungwaru, dan Ngantru. Sementara itu, 13 kecamatan lainnya mengalami surplus beras pada tahun 2019.

KESIMPULAN

Berdasarkan metode *geoadditive* SAE, rata-rata produktivitas padi level kecamatan di Kabupaten Tulungagung tahun 2019 yaitu 6,8364 ton/ha. Hasil pendugaan menggunakan *geoadditive* SAE menunjukkan bahwa produktivitas padi tertinggi berada di Kecamatan Pucanglaban sebesar 8,8648 ton/ha, sedangkan produktivitas padi terendah di Kecamatan Pagerwojo sebesar 3,6576 ton/ha. Penggunaan metode *geoadditive* SAE untuk mengestimasi produktivitas padi level kecamatan memberikan tingkat presisi lebih baik karena menghasilkan RMSE dan RSE lebih kecil daripada estimasi langsung. Hasil pendugaan ini juga menunjukkan bahwa terdapat 6 kecamatan yang masih mengalami defisit beras pada tahun 2019. Penelitian ini masih memiliki keterbatasan dalam mengestimasi produktivitas padi di kecamatan nirsampel. Oleh karena itu, penelitian selanjutnya disarankan dapat mengembangkan SAE *Clustering* yang melibatkan dua pengaruh acak seperti yang terdapat pada *geoadditive* SAE untuk mengestimasi produktivitas padi di kecamatan nirsampel.

DAFTAR PUSTAKA

- Apriani, F. (2017). Pemodelan Pengeluaran per Kapita Menggunakan Small Area Estimation dengan Pendekatan Semiparametrik Penalized Spline [Tesis]. Surabaya: Institut Teknologi Sepuluh Nopember.
- Ardiansyah, M. (2018). Pendugaan Produktivitas Padi di Tingkat Kecamatan Menggunakan Geoadditive Small Area Model [Tesis]. Bogor: Institut Pertanian Bogor.
- Biemer, P. P., & Lyberg, L. E. (2003). *Introduction to Survey Quality*. Hoboken: John Wiley & Sons, Inc.
- Bocci, C. (2009). *Geoadditive Models for Data with Spatial Information*. Florence: University of Florence.
- BPS Jawa Timur. (2020). *Provinsi Jawa Timur Dalam Angka 2020*. Surabaya: Badan Pusat Statistik Provinsi Jawa Timur.
- BPS RI. (2018). *Kajian Konsumsi Bahan Pokok 2017*. Jakarta: Badan Pusat Statistik.
- BPS RI. (2018). *Konversi Gabah ke Beras (SKGB) Tahun 2018*. Jakarta: Badan Pusat Statistik.
- BPS Tulungagung. (2021). *Kabupaten Tulungagung Dalam Angka 2021*. Tulungagung: Badan Pusat Statistik Kabupaten Tulungagung.
- Chand, N., & Alexander, C. H. (1995). Using Administrative Records for Small Area Estimation in the American Community Survey. 1999 *FCSM Research Conference* (hal. 1-9). Maryland, United States: U.S. Bureau of the Census.
- Ghosh, M., & Rao, J. (1994). Small Area Estimation: An Appraisal. *Statistical Science*, 9(1), 55-76.
- Kamman, E. E., & Wand, M. P. (2003). Geoadditive Models. *Journal of the Royal Statistical Society Applied Statistics*, 52(1), 1-18. doi:10.1111/1467-9876.00385
- Kurnia, A. (2009). Prediksi Terbaik Empirik untuk Model Transformasi Logaritma di dalam Pendugaan Area Kecil dengan Penerapan pada Data Susenas [Tesis]. Bogor: Institut Pertanian Bogor.
- Petrucci, A., & Pratesi, M. (2014). Spatial Models in Small Area Estimation in the Context of Official Statistics. *Statistica Applicata - Italian Journal of Applied Statistics*, 24(1), 9-27.
- Pusponegoro, N. H., Djuraidah, A., Fitrianto, A., & Sumertajaya, I. M.

- (2019). Geo-additive Models in Small Area Estimation of Poverty. *Journal of Data Science and Its Applications*, 2(1), 11-18. doi:10.21108/JDSA.2019.2.15
- Rao, J., & Molina, I. (2015). *Small Area Estimation Second Edition*. New Jersey: John Wiley & Sons, Inc.
- Ruppert, D., Wand, M. P., & Carroll, R. J. (2003). *Semiparametric Regression*. New York: Cambridge University Press.

