ISSN: 2339-2541
JURNAL GAUSSIAN, Volume 5, Nomor 1, Tahun 2016, Halaman 91-97
Online di: http://ejournal-s1.undip.ac.id/index.php/gaussian

# PERAMALAN DINAMIS PRODUKSI PADI DI JAWA TENGAH MENGGUNAKAN METODE KOYCK DAN ALMON

**Firdha Rahmatika Pratami[1], Sudarno[2], Dwi Ispriyanti [3]**
[1]Mahasiswa Jurusan Statistika FSM Universitas Diponegoro
[2,3]Staff Pengajar Jurusan Statistika FSM Universitas Diponegoro
**Email:** rahmatikafirdha@gmail.com dsghani@gmail.com ispriyanti.dwi@gmail.com

## ABSTRACT

Paddy is one of the staple crops that have strategic value and has a great influence in economic, environmental, social and political. Almost of Indonesia's population consumes rice every day. Because of that, need models to determine or predict the amount of paddy production in Central Java for the future. Because the data used is the historical data, there will be a regression analysis that takes into account the time. If the regression model include not only the value of the independent variable *X* at this time, but also the value of the past (lagged), this model called a distributed-lag model. The methods used in determining the equation of distributed-lag are Koyck and Almon method. Koyck method used to determine the estimated dynamic model of distributed-lag time difference (lag) is unknown. Almon method used to determine the estimated dynamic model of distributed-lag time difference (lag) is known. Selection of the best model is using Mean Absolut Percentage Error criteria. According the result of the analysis, using Almon model has better result than Koyck Model.

**Keyword**: *Paddy, Distributed-lag model, Koyck, Almon*

## 1. PENDAHULUAN

### 1.1 Latar Belakang

Indonesia adalah negara agraris dengan sebagian besar penduduknya hidup dari hasil bercocok tanam atau bertani, sehingga pertanian merupakan sektor yang memegang peranan penting dalam kesejahteraan kehidupan penduduk Indonesia. Padi adalah salah satu tanaman pangan pokok yang memiliki nilai strategis dan mempunyai pengaruh yang besar dalam bidang ekonomi, lingkungan, dan sosial politik. Hampir semua penduduk Indonesia mengkonsumsi hasil padi atau beras setiap harinya.

Provinsi Jawa Tengah merupakan salah satu provinsi penyangga pangan nasional, oleh karena itu produktivitas padi lebih diutamakan untuk terus dipacu. Pada tahun 2013, produktivitas padi sekitar 56,06 kuintal per hektar, menurun 2,84 persen dibanding produktivitas tahun sebelumnya. Luas panen padi dan jumlah produksi padi mengalami peningkatan masing-masing sebesar 4,05 persen dan 1,09 persen. Sebagian produksi merupakan padi sawah, yaitu sekitar 96,74 persen (BPS, 2014).

Berdasarkan hal tersebut, perlu adanya model untuk mengetahui atau meramalkan jumlah produksi padi di Jawa Tengah untuk waktu mendatang. Obyek penelitian fokus pada Provinsi Jawa Tengah sebagai penyangga pangan nasional yang diharapkan mampu mencukupi ketersediaan di wilayahnya sendiri bahkan berkontribusi untuk wilayah lain.

Data yang digunakan merupakan data historis dari produksi padi dan luas lahan di Provinsi Jawa Tengah. Karena data yang digunakan adalah data historis, maka akan dilakukan analisis regresi linier yang memperhitungkan waktu. Jika model regresi memasukkan tidak hanya nilai variabel bebas *X* saat ini, tetapi juga nilai masa lalu (*lagged*), model ini disebut model terdistribusi-*lag*. Sedangkan jika model tersebut memasukkan satu atau lebih nilai masa lalu (*lagged*) dari variabel tak bebas di antara variabel bebasnya, model ini disebut model autoregresif (Gujarati, 1978). Waktu yang diperlukan bagi variabel *X* dalam mempengaruhi variabel tak bebas *Y* disebut beda kala atau *lag* (Supranto, 1995). Metode-metode yang digunakan dalam menentukan persamaan terdistribusi-*lag* antara lain metode Koyck dan Almon, sebab kedua metode ini lebih mudah diterapkan dalam membuat estimasi model dinamis distribusi *lag*.

Metode Koyck digunakan untuk menentukan estimasi model dinamis terdistribusi-*lag* yang panjang beda kala (*lag*) tidak diketahui. Pada persamaan Koyck diakhiri dengan model autoregresif karena muncul variabel bebas $Y_{t-1}$.

Metode Almon digunakan untuk menentukan estimasi model dinamis terdistribusi-*lag* yang panjang beda kala (*lag*) diketahui. Almon merupakan alternatif pada model regresi *lag* yang menghindari permasalahan estimasi berkaitan dengan model autoregresif (Gujarati, 2013).

Berdasarkan latar belakang di atas, penulis akan melakukan penelitian mengenai pemodelan dinamis produksi padi di Jawa Tengah menggunakan metode Koyck dan Almon serta menentukan model terbaiknya.

## 1.2 Tujuan

Tujuan dari penelitian ini adalah menentukan model dinamis produksi padi di jawa Tengah menggunakan metode Koyck dan Almon serta mendapatkan estimasi model terbaiknya.

# 2. TINJAUAN PUSTAKA

## 2.1 Analisis Regresi Berganda

Analisis regresi adalah suatu teknik statistik yang digunakan untuk menyelidiki dan membangun suatu model yang menghubungkan antara variabel-variabelnya. Model regresi dengan lebih dari satu variabel independen disebut dengan model regresi berganda. Model regresi linier berganda dengan variabel dependen *Y* dan *p* variabel independen dapat dinyatakan sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + ... + \beta_k X_{ip} + \varepsilon_i \quad (1)$$

Metode kuadrat terkecil digunakan untuk menduga koefisien regresi dari model regresi linier berganda. Metode kuadrat terkecil (*Ordinary Least Square/OLS*) adalah suatu metode yang digunakan untuk menentukan persamaan regresi dengan meminimumkan jumlah kuadrat jarak vertikal antara nilai aktual Y dan nilai dugaan atau ramalan $\hat{Y}$ (Suharyadi dan Purwanto, 2004). Estimasi $\beta_0, \beta_1, \ldots, \beta_k$ didapatkan dengan metode kuadrat terkecil sehingga diperoleh rumus sebagai berikut:

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y} \quad (2)$$

Terdapat dua uji hipotesis yang digunakan untuk mengetahui apakah variabel-variabel bebas dalam model regresi secara signifikan mempengaruhi variabel tak bebas, yaitu uji signifikansi regresi (uji F) dan uji koefisien regresi individual (uji t).

Ada beberapa asumsi yang digunakan dalam metode kuadrat terkecil, diantaranya yaitu residual berdistribusi normal, nonmultikolinearitas, nonautokorelasi, dan homoskedastisitas.

## 2.2 Model Dinamis Terdistribusi-*lag*

Dalam analisis regresi yang menggunakan data *time series,* model regresi melibatkan data pada waktu sekarang dan waktu lampau/selang waktu (*lagged/past*) dari variabel bebas *X*, maka dinamakan model terdistribusi-*lag*. Menurut Gujarati (2013), model terdistribusi-*lag* adalah sebagai berikut

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \varepsilon_t \quad (3)$$

Persamaan (3) menggambarkan bahwa nilai $Y_t$ tergantung atau dipengaruhi oleh nilai *X* pada saat *t* ($X_t$), nilai *X* pada satu unit ukuran sebelumnya ($X_{t-1}$), dan nilai *X* pada dua unit ukuran waktu sebelumnya ($X_{t-2}$). Model terdistribusi-*lag* telah menunjukkan kegunaan yang sangat besar dalam ilmu ekonomi empiris karena model ini telah membuat teori ekonomi yang bersifat statis menjadi yang bersifat dinamis dengan memperhitungkan peranan dari waktu.

Ada 2 jenis model terdistribusi-*lag*, yaitu:

a. Model *Lag Infinite*
   Model : $Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \ldots + \varepsilon_t$
   Model ini disebut model *lag infinite* sebab panjang *lag* tidak diketahui.
b. Model *Lag Finite*
   Model : $Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \ldots + \beta_k X_{t-k} + \varepsilon_t$
   Model ini disebut model *lag finite* sebab panjang *lag* diketahui sebesar *k*.

### 2.2.1 Pendekatan Koyck pada Model Terdistribusi-*lag*

Menurut Gujarati (2013), untuk menaksir parameter model terdistribusi-*lag* dilakukan dengan pendekatan Koyck, dimana persamaannya adalah

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \ldots + \varepsilon_t \quad (4)$$

Koyck telah mengusulkan metode cerdas dalam melakukan estimasi model terdistribusi-*lag*. Dengan mengasumsikan bahwa seluruh $\beta$ mempunyai tanda yang sama, Koyck mengasumsikan nilai koefisien tersebut akan berkurang secara geometris sebagai berikut

$$\beta_k = \beta_0 \lambda^k \text{ , k= 0,1,2,... dan } 0 < \lambda < 1 \quad (5)$$

dimana $\lambda$ adalah tingkat penurunan dari terdistribusi-*lag*.

Model Koyck dirumuskan dalam bentuk sebagai berikut

$$Y_t = (1-\lambda)\,\alpha + \beta_0 X_t + \lambda\, Y_{t-1} + v_t \quad (6)$$

dengan $v_t = \varepsilon_t - \lambda\varepsilon_{t-1}$

Menurut Nachrowi (2013) dalam Rangkuti (2007), adapun asumsi-asumsi dari aturan Koyck yaitu:

a. Nilai $\lambda$ positif sehingga $\beta_k$ selalu positif
b. Karena $0 < \lambda < 1$ maka semakin jauh periodenya semakin kecil bobot $\beta_k$
c. Aturan Koyck menjamin bahwa jumlah $\beta$ adalah penjumlahan jangka panjang, yaitu

$$\sum_{k=0}^{n} \beta_k = \beta_0 \left(1 + \lambda + \lambda^2 + \lambda^3 + \cdots\right) = \beta_0 \left(\frac{1}{1-\lambda}\right) \quad (8)$$

### 2.2.2 Pendekatan Almon pada Model terdistribusi-*lag*

Menurut Gujarati (2013), meskipun banyak digunakan, model terdistribusi-*lag* Koyck berdasarkan asumsi bahwa koefisien-koefisien $\beta$ menurun secara geometris ketika *lag* bertambah. Namun, apabila diagram pencar antara $\beta$ dengan *lag* itu naik kemudian turun maka metode Koyck tidak dapat digunakan.

Untuk menaksir parameter model terdistribusi-*lag* didekati dengan model *lag* finite, yakni

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \ldots + \beta_k X_{t-k} + \varepsilon_t \quad (9)$$

Yang dapat ditulis secara ringkas sebagai berikut

$$Y_t = \alpha + \sum_{i=0}^{k} \beta_i X_{t-i} + \varepsilon_t \quad (10)$$

Almon mengasumsikan bahwa $\beta_i$ bisa dikira-kira oleh polinomial derajat yang sesuai dari *i*, panjangnya *lag*. Misalnya, jika $\beta$ mengikuti polinomial derajat kedua model dapat dituliskan sebagai berikut

$$\beta_i = \alpha_0 + \alpha_1 i + \alpha_2 i^2 \quad (11)$$

Untuk menjelaskan bagaimana Almon bekerja, diasumsikan bahwa $\beta$ mengikuti polinomial derajat kedua. Dengan mendistribusi persamaan (11) ke persamaan (10), diperoleh

$$Y_t = \alpha + \sum_{i=0}^{k}(\alpha_0 + \alpha_1 i + \alpha_2 i^2) X_{t-i} + \varepsilon_t$$

$$= \alpha + \alpha_0 \sum_{i=0}^{k} X_{t-i} + \alpha_1 \sum_{i=0}^{k} i X_{t-i} + \alpha_2 i^2 \sum_{i=0}^{k} X_{t-i} + \varepsilon_t \quad (12)$$

didefinisikan bahwa

$$Z_{0t} = \sum_{i=0}^{k} X_{t-i}$$
$$Z_{1t} = \sum_{i=0}^{k} i X_{t-i}$$
$$Z_{2t} = \sum_{i=0}^{k} i^2 X_{t-i} \quad (13)$$

Maka (12) menjadi

$$Y_t = \alpha + \alpha_0 Z_{0t} + \alpha_1 Z_{1t} + \alpha_2 Z_{2t} + \varepsilon_t \quad (14)$$

Persamaan (14) dapat diperkirakan koefisiennya dengan metode kuadrat terkecil. Perkiraan $\hat{\alpha}$ dan $\alpha_i$ yang diperoleh akan mempunyai sifat-sifat yang diinginkan asalkan kesalahan pengganggu $\varepsilon_t$ memenuhi asumsi dari model linier yang klasik. Setelah semua $\alpha_i$ diperkirakan dari persamaan (14), koefisien $\hat{\beta}$ dapat dihitung berdasarkan persamaan (11) sebagai berikut

$$\hat{\beta}_0 = \hat{\alpha}_0$$
$$\hat{\beta}_1 = \hat{\alpha}_0 + \hat{\alpha}_1 + \hat{\alpha}_2$$
$$\hat{\beta}_2 = \hat{\alpha}_0 + 2\hat{\alpha}_1 + 4\hat{\alpha}_2$$
$$\hat{\beta}_3 = \hat{\alpha}_0 + 3\hat{\alpha}_1 + 9\hat{\alpha}_2$$
$$\vdots$$
$$\hat{\beta}_k = \hat{\alpha}_0 + k\hat{\alpha}_1 + k^2\hat{\alpha}_2 \quad (15)$$

Jadi, model estimasi terdistribusi-*lag* dari Almon yang sesuai persamaan (9) adalah

$$\hat{Y}_t = \hat{\alpha} + \hat{\beta}_0 X_t + \hat{\beta}_1 X_{t-1} + \hat{\beta}_2 X_{t-2} + \ldots + \hat{\beta}_k X_{t-k}$$

## 2.3 Pemilihan Model Terbaik

Pemilihan model terbaik dalam peramalan ditentukan oleh *error* yang dihasilkan. MAPE (*Mean Absolut Percentage Error*) merupakan salah satu alternatif untuk penentuan model terbaik dari beberapa model yang dipilih. Semakin kecil persentase *error* yang dihasilkan semakin bagus pula peramalan yang dilakukan. Menurut Makridakis (1999), persamaan MAPE dirumuskan sebagai berikut:

$$MAPE = \frac{\sum_{t=1}^{n}|PE_t|}{n}$$

$$PE_t = \left(\frac{Y_t - F_t}{Y_t}\right) x\ 100\%$$

dengan

$PE_t$ = Persentase Kesalahan periode ke-t
$Y_t$ = Data aktual periode ke-t
$F_t$ = Nilai peramalan periode ke-t
n = Jumlah data

## 3. METODOLOGI PENELITIAN

### 3.1 Data

Data yang digunakan dalam penelitian ini merupakan data sekunder, yaitu data tahunan produksi dan luas lahan padi di Provinsi Jawa Tengah dari tahun 1987 hingga 2013 yang diperoleh dari Dinas Pertanian Tanaman Pangan dan Holtikultura.

### 3.2 Metode Analisis

Adapun tahapan analisis yang dilakukan dalam penelitian ini adalah sebagai berikut:

1. *Input* data sekunder dari Dinas Pertanian Tanaman Pangan Provinsi Jawa Tengah.
2. Membuat persamaan regresi Koyck.
3. Melakukan uji statistik F dan uji t pada regresi Koyck.
4. Menentukan model akhir regresi Koyck.
5. Melakukan uji asumsi pada model akhir regresi Koyck.
6. Menentukan model dinamis terdistribusi-*lag* Koyck.
7. Membuat persamaan regresi Almon.
8. Melakukan uji statistik F dan uji t pada regresi Almon.
9. Menentukan panjang maksimum *lag k*.
10. Menentukan model akhir regresi Almon.
11. Melakukan uji asumsi pada model akhir regresi Almon.
12. Menentukan model dinamis terdistribusi-*lag* Almon.
13. Membandingkan nilai MAPE Koyck dan Almon.
14. Mendapatkan estimasi model terbaik.

## 4. HASIL DAN PEMBAHASAN

### 4.1 Pendekatan Koyck pada Model Terdistribusi-*Lag*

Didapatkan persamaan awal regresi Koyck adalah sebagai berikut

$$\hat{Y}_t = -3.143.373{,}595 + 5{,}04316803\ X_t + 0{,}411863629\ Y_{t-1} \quad (16)$$

Terdapat dua macam pengujian parameter yaitu uji signifikansi regresi (uji F) dan uji koefisien regresi Individual (uji t). Pada uji F menunjukkan hasil yang signifikan pada taraf signifikansi 5% dengan nilai F= 106,22 dan nilai *p-value* sebesar 0,000. Hal tersebut menunjukkan bahwa model regresi yang diperoleh dapat digunakan. Pada uji t menunjukkan hasil yang signifikan pada taraf signifikansi 5% dapat dilihat pada tabel 1.

**Tabel 1.** Nilai $t_{hitung}$ untuk Setiap Variabel pada Model Koyck

| Variabel Bebas | $t_{hitung}$ | *p-value* |
|---|---|---|
| $X_t$ | 5,221987138 | 0,000 |
| $Y_{t-1}$ | 3,568081956 | 0,002 |

Pengujian asumsi klasik (Galat menyebar normal, nonautokorelasi, nonmultikolinieritas, dan homoskedastisitas) pada model terpenuhi. Maka persamaan awal regresi Koyck sama dengan persamaan akhirnya.

Persamaan (16) dapat dituliskan dalam bentuk persamaan dinamis terdistribusi-*lag* sebagai berikut

$$\hat{Y}_t = -5.344.633{,}915 + 5{,}04316803\ X_t + 2{,}077097486\ X_{t-1} + 0{,}855480909\ X_{t-2} + 0{,}352341472\ X_{t-3} + 0{,}145116637\ X_{t-4} + \ldots \quad (17)$$

### 4.2 Pendekatan Almon pada Model Terdistribusi-*Lag*

Langkah pertama yang harus dilakukan adalah dengan menghitung nilai $Z_t$ sesuai dengan panjang $k$ untuk mencari model regresi yang mungkin terbentuk. Kemudian ditentukan derajat polinomial $m$ dalam hal ini Almon menyarankan menggunakan derajat kedua dan ketiga untuk menentukan persamaan Almon. Maka akan didapat 43 model dengan 22 model merupakan derajat polinomial kedua dan 21 model merupakan derajat polinomial ketiga.

Pada uji F dengan taraf signifikansi 5% pada $m$ = 2 dan $m$ = 3, *lag* 1 sampai *lag* 18 secara bersama-sama variabel bebas yang signifikan mempengaruhi variabel tak bebasnya atau model cocok untuk digunakan. Sedangkan pada $m$ = 2, *lag* 19 sampai *lag* 22 dan $m$ = 3, *lag* 19 sampai *lag* 21 variabel bebas tidak signifikan mempengaruhi variabel tak bebasnya atau model tidak cocok digunakan. Berdasarkan uji F, selanjutnya dilakukan uji t dengan hasil dari uji ini akan mempengaruhi pengambilan keputusan dalam menentukan panjang maksimum *lag k*.

Berdasarkan uji hipotesis yang telah dilakukan, pada $m$ = 2 panjang maksimum yang dapat diperoleh adalah pada *lag* ke-12, karena pada *lag* ke-13 variabel bebasnya mulai tidak signifikan. Sedangkan pada $m$ = 3, panjang maksimum yang dapat diperoleh adalah pada *lag* ke-2. Namun polinomial derajat ketiga tidak memberikan pendugaan yang cukup baik, karena untuk *lag* ke-2 terdapat variabel yang dihapus sehingga tidak dapat digunakan untuk membuat persamaan dinamis Almon. Jadi dapat disimpulkan panjang maksimum *lag* $k$ terdapat pada model Almon *lag* ke-12 dengan polinomial derajat kedua yang digunakan.

Setelah menentukan panjang maksimum $k$ yaitu pada *lag* ke-12, didapatkan persamaan akhir dari regresi Almon sebagai berikut

$$\hat{Y}_t = -18.351.652 + 2{,}9511\ Z_{ot} - 0{,}8163\ Z_{1t} + 0{,}06469\ Z_{2t} \qquad (18)$$

Pengujian asumsi klasik (Galat menyebar normal, nonautokorelasi, dan homoskedastisitas) pada model terpenuhi. Sedangkan Variabel Z sangat mungkin mengalami multikolinieritas karena ditentukan dari derajat polinomialnya.

Persamaan (18) dapat dituliskan dalam bentuk persamaan dinamis terdistribusi-*lag* sebagai berikut

$$\hat{Y}_t = -18351652 + 2{,}95\ X_t + 2{,}1987\ X_{t-1} + 1{,}5768\ X_{t-2} + 1{,}0843X_{t-3} + 0{,}7212X_{t-4} + \ldots + 2{,}4748\ X_{t-12} \qquad (19)$$

### 4.3 Pemilihan Model Terbaik

Setelah didapat model akhir dari Koyck dan Almon, kemudian dilakukan pemilihan model terbaik dengan menghitung *Mean Absolut Percentage Error* (MAPE) dan didapatkan persentase MAPE untuk Koyck adalah 1,919389104 % dan untuk Almon adalah 1,786257247 %. Berdasarkan nilai MAPE Koyck dan Almon, didapat persentase MAPE terkecil terdapat pada model Almon. Maka dapat disimpulkan bahwa model terbaik adalah model Almon.

## 5. KESIMPULAN

Pemilihan model terbaik dilakukan menggunakan kriteria *Mean Absolute Percentage Error*. Model dengan MAPE terkecil ditetapkan sebagai model terbaik. Hasil yang didapat adalah model Almon pada *lag* ke-12 dengan polinomial derajat kedua yang digunakan. Persamaan dinamis terdistribusi-*lag* Almon yang didapat adalah

$$\hat{Y}_t = -18351652 + 2{,}95\ X_t + 2{,}1987\ X_{t-1} + 1{,}5768\ X_{t-2} + 1{,}0843X_{t-3} + 0{,}7212X_{t-4} + \ldots + 2{,}4748\ X_{t-12}$$

## DAFTAR PUSTAKA

Badan Pusat Statistik. 2014. *Jawa Tengah dalam Angka 2014*. BPS Provinsi Jawa Tengah.

Dinas Pertanian Tanaman Pangan dan Holtikultura Provinsi Jawa Tengah. 2014.

Gujarati, D., 1978. *Ekonometrika Dasar*. Jakarta: Erlangga.

Gujarati, D., 2013. *Basic Econometrics*. New York: McGraw- Hill Education.

Makridakis, S. dan Wheelwright S.C., 1999. *Metode dan Aplikasi Peramalan*. Jakarta: Erlangga.

Rangkuti, A. 2007. *Kombinasi Penaksiran Model Lag Terdistribusi dengan Ekspektasi Adaptif dan Penyesuaian Parsial*. Jurnal Matematika. Makassar: Universitas Hassanudin.

Suharyadi dan Purwanto S.K., 2004. *Statistika: Untuk Ekonomi & Keuangan Modern*. Jakarta: Salemba Empat.

Supranto, J., 1995. *Ekonometrik*. Jakarta: Lembaga Penerbit Fakultas Ekonomi Universitas Indonesia.