

**PENDUGAAN DATA TIDAK LENGKAP CURAH HUJAN**  
**DI KABUPATEN INDRAMAYU**  
**DENGAN KRIGING & RATA-RATA BERGERAK (*MOVING AVERAGE*)**  
**(BERDASARKAN DATA TAHUN 1980 – 2000)**

**Dewi Retno Sari Saputro<sup>1</sup>, Ahmad Ansori Mattjik<sup>2</sup>, Rizaldi Boer<sup>3</sup>**

**Aji Hamim Wigena<sup>4</sup>, Anik Djuraidah<sup>5</sup>**

<sup>1)</sup>Mahasiswa S3 Statistika Program Pascasarjana IPB, Jurusan Matematika FMIPA UNS

dewi\_rss@yahoo.com

<sup>2,4,5)</sup>Departemen Statistika FMIPA IPB

<sup>3)</sup>Departemen Geofisika dan Metereologi FMIPA IPB

### **Abstrak**

Berdasarkan amatan data curah hujan tahun 1980-2000 yang tersebar di 27 stasiun penakar hujan di Kabupaten Indramayu, terdapat ketidaklengkapan data. Ketidaklengkapan data curah hujan pada Kabupaten tersebut mencapai rata-rata 3.72% dengan persentase data tidak lengkap terbesar terjadi pada bulan Januari dan Pebruari sebesar 5.46% dan 5.29%. Meskipun persentasenya data tidak lengkapnya relatif kecil, namun menjadi penting dalam berbagai bidang penelitian karena dapat menyebabkan bias dan inefisiensi dalam memprediksi respon dari amatan. Oleh karena itu, diperlukan suatu pendugaan terhadap data tidak lengkap tersebut. Terdapat berbagai teknik dan cara untuk menduga data tidak lengkap. Di antaranya yaitu dengan Kriging dan metode rata-rata bergerak (*moving average/MA*). Hasil validasi model semivariogram dengan metode Kriging menunjukkan bahwa dugaan data tidak lengkap yang dihasilkan relatif menyimpang dari nilai aktualnya, sehingga model semivariogram tidak dapat digeneralisasi untuk pendugaan pada tahun lainnya. Hasil dengan MA menunjukkan bahwa rata-rata galat pendugaan (*Mean Absolute Deviation/MAD*) yang persentase galatnya lebih dari 40% sebanyak 29.81% artinya 70,19% sisanya memadai sebagai data pelengkap curah hujan. Selanjutnya dengan data yang telah lengkap tersebut dapat diperoleh pola curah hujan monsoon

dan ini sesuai hasil penelitian tentang pola curah hujan di Indonesia bahwa untuk daerah di pulau Jawa curah hujannya bertipe monsoon yakni tipe curah hujan yang bersifat unimodal.

**Kata Kunci** : Data tidak lengkap, Kriging, Deret waktu, Rata-rata bergerak, *Mean absolute deviation (MAD)*, Curah hujan bertipe monsoon

## 1. Pendahuluan

Data yang tidak lengkap atau data hilang mengindikasikan bahwa tidak ada data apapun yang tersimpan pada peubah amatan. Hal tersebut dapat disebabkan berbagai hal di antaranya: alat ukur yang kurang akurat, tidak tercatat dan masalah-masalah teknis lainnya. Data hilang merupakan masalah yang penting dalam berbagai bidang penelitian karena dapat menyebabkan bias dan inefisiensi dalam memprediksi respon dari amatan. Menurut Little & Rubin (1990), Scheffer (2002), Tsiatis (2006), serta Daniels & Hogan (2008), terdapat beberapa jenis data tidak lengkap berdasarkan mekanismenya, yakni: MCAR (*Missing Completely at Random*), MAR (*Missing at Random*), *Nonignorable*.

Sampai dengan tahun 1970, data yang tidak lengkap diselesaikan dengan pengeditan. Rubin (1976) mengembangkan kerangka inferensi dari data tidak lengkap yang dipergunakan selama ini (Schafer & Graham 2002). Dalam beberapa kasus, cara yang paling sederhana untuk mengatasi data tidak lengkap dengan menghilangkan data tersebut dan membatasi perhatian pada data yang mengandung amatan lengkap saja. Namun, jika data yang hilang dalam jumlah yang cukup besar menyebabkan terjadinya peningkatan kesalahan secara keseluruhan dan dapat menurunkan ketepatan pendugaan.

Salah satu kasus data tidak lengkap yaitu data curah hujan di Kabupaten Indramayu berdasarkan amatan tahun 1980-2000 yang tersebar di 27 stasiun penakar hujan. Ketidaklengkapan data curah hujan pada Kabupaten tersebut mencapai rata-rata 3.72% dengan persentase data tidak lengkap terbesar terjadi pada bulan Januari dan Pebruari sebesar 5.46% dan 5.29%. Meskipun persentasenya relatif kecil, namun menjadi penting dalam berbagai bidang penelitian karena dapat menyebabkan bias dan

inefisiensi dalam memprediksi respon dari amatan. Oleh karena itu, diperlukan suatu pendugaan terhadap data tidak lengkap tersebut.

Terdapat berbagai teknik dan cara untuk menduga data tidak lengkap, seperti yang dinyatakan oleh Little & Rubin (1987), penanganan data hilang dapat dilakukan berdasarkan prosedur : amatan lengkap, imputasi, pembobotan dan model. Di antara teknik yang berbasiskan prosedur tersebut yaitu Kriging dan metode rata-rata bergerak (*moving average/MA*). Kedua teknik, Kriging dan MA dipergunakan untuk menduga data tidak lengkap curah hujan mengingat data curah hujan dianggap sebagai data yang berbasis *spatio-temporal* yaitu proses stokastik yang terjadi berkorelasi secara serentak dalam lokasi (*spatial*) dan waktu (*temporal*). Hal ini berarti bahwa model statistik yang dipergunakan untuk menduga data hilang dapat menggunakan Kriging dengan model variogramnya atau dengan metode rata-rata bergerak (Khoerudin 2010).

Dalam penelitian ini dikaji pendugaan data tidak lengkap curah hujan dengan rata-rata bergerak, ditunjukkan pula hasil kajian dari Khoeruddin (2010) tentang metode pendugaan dengan ordinary Kriging pada data yang tidak lengkap tersebut serta didiskusikan hasil dua kajian antara Kriging dan MA.

## 2. Metodologi Penelitian

Data yang dipergunakan merupakan data curah hujan di Kabupaten Indramayu tahun 1980-2000 yang tersebar dalam 27 stasiun penakar hujan (lampiran), dan posisi stasiun curah hujan (*latitude dan longitude*). Data ini diperoleh dari Lab Geofisika dan Metereologi, Departemen Metereologi dan Geofisika, FMIPA IPB dan juga dari Badan Metereologi, Klimatologi dan Geofisika (BMKG).

Metode yang dipergunakan yakni metode rata-rata bergerak (*moving average*), metode ini memadai untuk kasus dengan pola datanya tidak mengikuti pola sebaran tertentu dan digunakan pada data yang stasioner atau data yang konstan terhadap ragam, merupakan salah satu cara untuk mengubah pengaruh data masa lalu terhadap nilai tengah sebagai pendugaan. Pendugaannya berbasiskan pemulusan (*smoothing*) yakni dengan melakukan rata-rata untuk menghilangkan pengaruh data irreguler yang bersifat

acak. Tekniknya dengan menentukan sejak awal berapa jumlah nilai observasi masa lalu yang akan dimasukkan untuk menghitung nilai tengah, selanjutnya setiap muncul nilai observasi baru, nilai rata-rata baru dihitung dengan tidak mempergunakan nilai observasi yang paling tua dan memasukkan nilai observasi yang terbaru. Nilai rata-rata bergerak ini kemudian akan menjadi dugaan untuk periode mendatang.

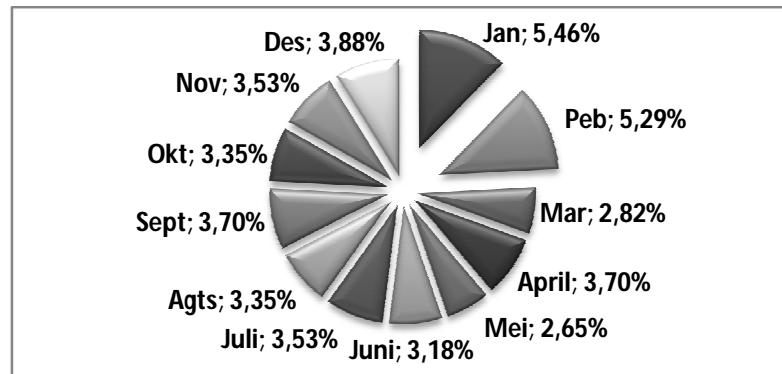
Secara garis besar langkah-langkah dalam penelitian ini dinyatakan sebagai berikut.

- a. Melakukan review hasil kajian Kriging.
- b. Menyusun data curah hujan berdasarkan rata-rata bulanan dari 27 stasiun penakar hujan per bulan.
- c. Menentukan length (panjang waktu) MA dan menghitung pendugaan data tidak lengkapnya, nilai MA ditentukan dengan  $\hat{Y}_{t+1} = \frac{1}{n} \sum_{i=t+1-k}^t Y_i$ ,  $\hat{Y}_{t+1}$  yakni nilai pendugaan periode mendatang,  $Y_t$  nilai pada periode  $t$ ,  $n$  merupakan length.
- d. Menentukan galat dengan *mean absolute deviation (MAD)* terkecil berdasarkan lengthnya. Nilai MAD ditentukan dengan  $MAD = \frac{1}{n} \sum_{t=1}^n |Y_t - \hat{Y}_t|$
- e. Menyusun dan melengkapi data pendugaannya, menentukan rata-ratanya selanjutnya melakukan *plotting* untuk dua dimensi dan dalam bentuk spasial.
- f. Menganalisis hasil plot.

### 3. Hasil Penelitian Dan Pembahasan

#### 3.1. Karakteristik Data

Stasiun-stasiun penakar curah hujan telah dibangun untuk memantau dan mengetahui pola curah hujan dari waktu ke waktu. Pada beberapa waktu tertentu terdapat beberapa data yang tidak lengkap. Seperti telah dinyatakan di pendahuluan, data tidak lengkap dapat terjadi karena faktor-faktor teknis. Dengan demikian, penelaahan pola perkembangan curah hujan dari waktu ke waktu menjadi kurang akurat. Berdasarkan data curah hujan tahun 1980-2000, diperoleh karakteristik data yang ditunjukkan pada Gambar 1 berikut.



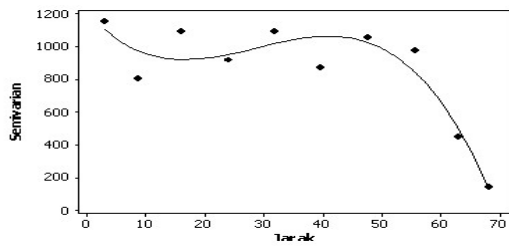
**Gambar 1.** Persentase Ketaklengkapan Data Curah Hujan Tahun 1980-2000

Gambar 1 menunjukkan bahwa rata-rata data tidak lengkap tersebar di semua bulan, bahkan rata-rata data tidak lengkap mencapai 3.7% dengan presentase tertinggi terjadi pada bulan Januari sebesar 5.46%.

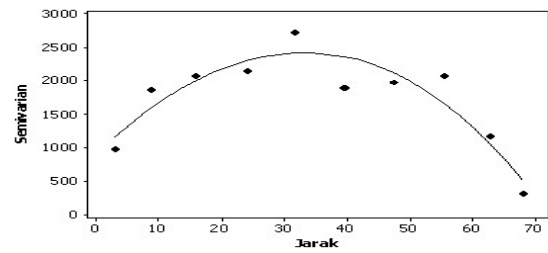
### 3.2. Kriging

Seperti telah dinyatakan dalam pendahuluan bahwa data curah hujan dianggap sebagai data yang berbasis *spatio-temporal* yakni proses stokastik yang terjadi berkorelasi secara serentak dalam lokasi (*spatial*) dan waktu (*temporal*). Hal ini berarti bahwa model statistik yang dipergunakan untuk menduga data hilang dapat menggunakan Kriging.

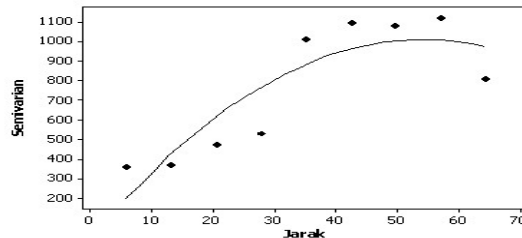
Oleh Khoeruddin (2010) dilakukan pendugaan data tidak lengkap pada kasus ini dengan Kriging, namun hasilnya tidak menunjukkan adanya pengaruh spasial. Hal ini dibuktikan dengan gambar semivariogram yang ditunjukkan pada Gambar 2-4 berikut. Berdasarkan gambar tersebut, dapat ditunjukkan bahwa semakin jauh jarak antar stasiun yang dipergunakan, semakin kecil ragamnya.



**Gambar 2.** Variogram Bulan Januari  
 (Bulan Basah)



**Gambar 3.** Variogram Bulan Mei



**Gambar 4** Variogram Bulan Agustus (Bulan Kering)

Selanjutnya, hasil dari penelitian Khoerudin tersebut juga menyatakan bahwa galat relatif pendugaan akan meningkat seiring dengan meningkatnya persen kehilangan data. Hasil validasi model untuk kehilangan data 5% menunjukkan bahwa rata-rata galat relatif terbesar terdapat pada bulan Agustus, yaitu sebesar 76,85%. Nilai ini berarti bahwa nilai dugaan yang dihasilkan menyimpang sebesar 76,85% dari nilai aktualnya. Secara keseluruhan rata-rata galat relatif untuk pendugaan data hilang 5% sebesar 51,82%. Rata-rata galat relatif terbesar pada kehilangan data 10% terdapat pada Bulan Agustus, yaitu sebesar 80,17%. Nilai ini berarti bahwa nilai dugaan yang dihasilkan menyimpang sebesar 80,17% dari nilai aktualnya. Secara keseluruhan rata-rata galat relatif untuk pendugaan data hilang 10% sebesar 49,48%. Rata-rata galat relatif terbesar pada kehilangan data 15% terdapat pada bulan Agustus yaitu sebesar 88,54%. Nilai ini berarti bahwa nilai dugaan yang dihasilkan menyimpang sebesar 108,08% dari nilai aktualnya. Secara keseluruhan rata-rata galat relatif untuk pendugaan data hilang 15% sebesar 58,96%.

Secara keseluruhan dugaan yang dihasilkan relatif menyimpang dari nilai aktualnya. Hal ini disebabkan oleh model semivariogram yang tidak menunjukkan peningkatan keragaman curah hujan dengan semakin meningkatnya jarak antar stasiun. Kesimpulan dari penelitian tersebut, dikutip seperti berikut.

*Hasil validasi model menunjukkan bahwa dugaan yang dihasilkan relatif menyimpang dari nilai aktualnya. Hal ini disebabkan oleh model semivariogram yang tidak menunjukkan peningkatan keragaman curah hujan dengan semakin meningkatnya jarak antar stasiun. Galat relatif pendugaan akan meningkat seiring dengan meningkatnya persen kehilangan data. Model Semivariogram tidak dapat digeneralisasi untuk pendugaan pada tahun lainnya.*

### **3.3. Rata-rata Bergerak (*Moving Average*)**

Berdasarkan hasil dengan metode Kriging, model semivariogram tidak dapat dipergunakan untuk menduga data tahun berikutnya dikarenakan akurasi yang relatif kecil. Oleh karena itu dipergunakan metode lain, yang dapat meningkatkan akurasi pendugaannya. Metode tersebut berbasiskan waktu (*time series*), dalam penelitian ini dipergunakan metode MA yang merupakan salah satu teknik pemulusan (*smoothing*) dalam runtun waktu. Pemulusan pada dasarnya merupakan suatu proses yang secara sistematis dapat menghilangkan pola data yang kasar (berfluktuasi) dan selanjutnya dapat mengambil pola data yang dijelaskan secara umum.

Metode rata-rata bergerak, sesuai dengan yang namanya bergerak dilakukan dengan pengelompokan periode waktu dihitung rata-ratanya menurut pengelompokan periode waktu dihitung. Jika menggunakan rentang waktu yang lebih pendek maka hasil rata-rata bergerak yang akan diperoleh akan lebih mendekati kondisi sifat data yang sebenarnya dan rata-rata yang ditemukan terdistribusi atau tersebar pada kelompok data faktual. Sedangkan jika satuan waktu yang lebih panjang, rata-rata yang kita peroleh akan lebih mewakili sejumlah data yang lebih banyak dan beraneka macam

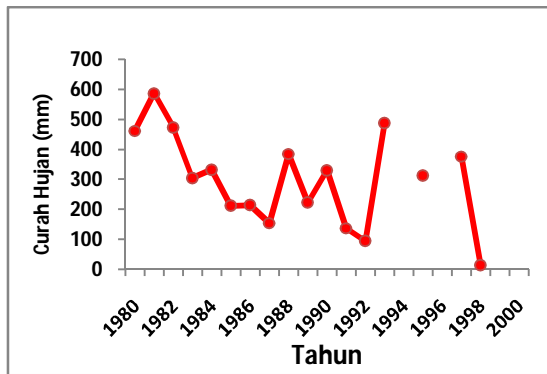
fluktuasinya, sehingga rata-rata Bergerak lebih tersebar dan kurang mewakili fakta sifat data yang tersebar tersebut. Terkecuali sifat data lebih homogen dan tidak terlalu fluktuatif. Proses ini merupakan konsep dasar untuk pemulusan yang lebih umum untuk peubah bebas nonkategori, di mana rata-rata kategori dikembangkan menjadi perhitungan rata-rata lokal, yakni menentukan nilai rata-rata respon pengamatan pada nilai-nilai peubah bebas yang dekat dengan suatu titik target tertentu. Jadi rata-rata dihitung pada suatu lingkungan (*neighborhood*) nilai tertentu yang dijadikan target.

Hal yang menjadi pusat perhatian dalam suatu proses pemulusan yakni (a) bagaimana merata-ratakan nilai respon pada suatu lingkungan tertentu dan (b) seberapa besar lingkungan yang harus diambil. Hal pertama terkait dengan metode pemulusan yang akan dipergunakan. Secara umum perbedaan pemulus disebabkan oleh perbedaan pada metode rata-ratanya, sedangkan hal kedua dikarenakan adanya kecenderungan bahwa selang yang besar akan menghasilkan ragam yang kecil, tetapi sangat potensial meningkatkan bias dan sebaliknya.

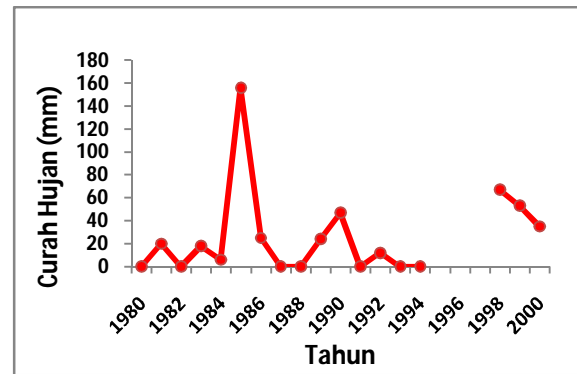
Metode pemulusan rata-rata Bergerak merupakan metode peramalan dengan bobot masing-masing nilai pengamatan yang sama. Pada pemulusan rata-rata Bergerak, pengaruh data masa lalu terhadap nilai tengah sebagai ramalan dapat diubah dengan menentukan sejak awal berapa jumlah nilai pengamatan masa lalu yang akan dimasukkan untuk menghitung nilai tengah. Prosedur ini dinamakan rata-rata Bergerak karena setiap muncul pengamatan baru, nilai rata-rata baru dapat dihitung dengan membuang nilai observasi yang paling tua dan memasukkan nilai pengamatan yang terbaru (Wei 1990).

Sebelum melakukan proses perhitungan dengan MA, ditunjukkan terlebih dahulu pola curah hujan dengan data tidak lengkapnya seperti pada Gambar 5 dan Gambar 6 berikut.



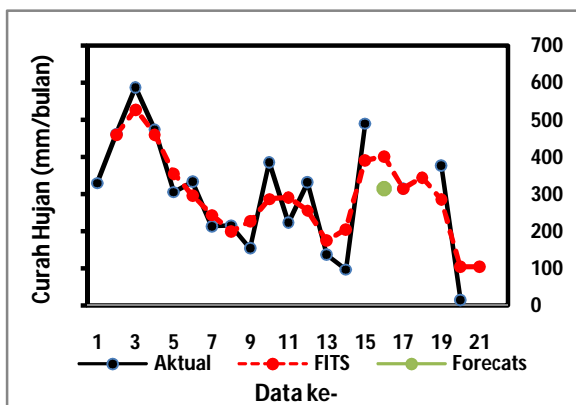


**Gambar 5.** Pola Data Hilang pada Bulan Januari di Stasiun Tamiyang

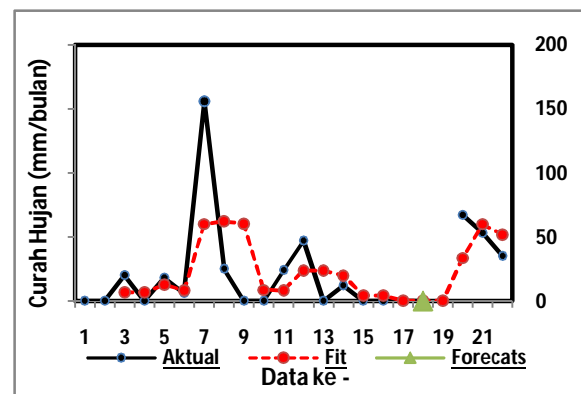


**Gambar 6.** Pola Data Hilang pada Bulan Januari di Stasiun Tamiyang

Pada pola yang terputus tersebut dilakukan pendugaannya dengan MA. Hasil perhitungan dengan MA ditunjukkan pada Gambar 7 dan 8 berikut merupakan pemulusan dengan metode MA di stasiun Anjatan di bulan Juli dengan MAD 18.426 dengan *length* 3 dan stasiun Anjatan di bulan Juli dengan MAD 71.05 dengan *length* 2.

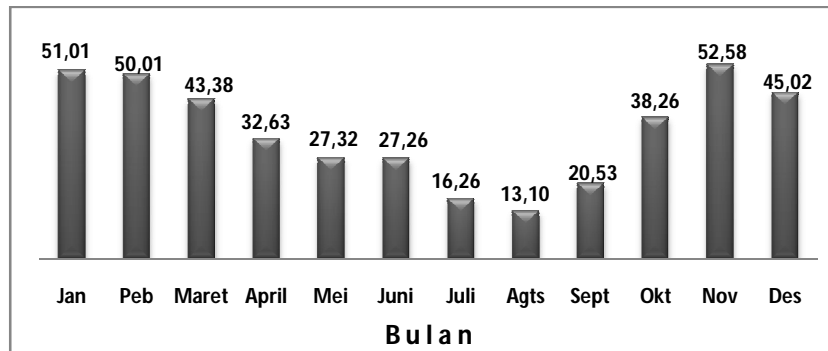


**Gambar 7.** MA untuk Stasiun Tamiyang di Bulan Januari



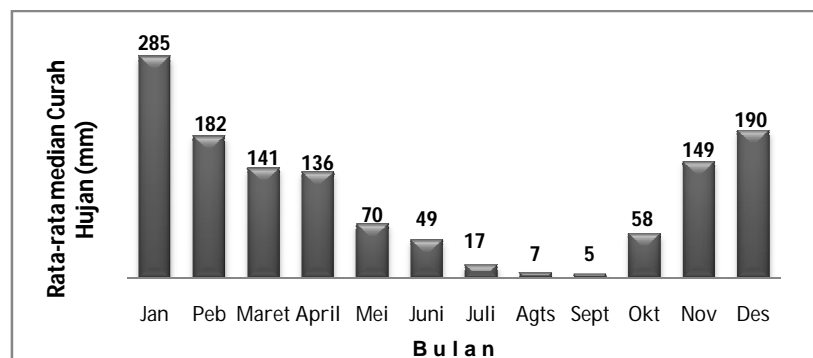
**Gambar 8.** MA untuk Stasiun Anjatan di Bulan Juli

Proses perhitungan pendugaan data tidak lengkap dengan MA dilakukan terhadap semua stasiun-stasiun penakar hujan yang memiliki data tidak lengkap. Berdasarkan proses tersebut, diperoleh rata-rata galat seperti ditunjukkan pada Gambar 9.



**Gambar 9.** Mean Absolute Deviation (MAD) dengan Rata-Rata Bergerak

Selanjutnya, hasil pendugaan data tidak lengkapnya dapat dilengkapi dan diperoleh pola data seperti pada Gambar 10 berikut.



**Gambar 10.** Rata-Rata Curah Hujan Setelah Data Dilengkapi

Gambar 10 menunjukkan pola curah hujan monsoon dan sesuai dengan hasil penelitian Aldrian dan Susanto (2003) bahwa untuk daerah di pulau Jawa tipe curah hujannya monsoon yakni tipe curah hujan yang bersifat unimodal. Secara umum pola curah hujan wilayah Kabupaten Indramayu mempunyai puncak musim hujan pada bulan Januari dan puncak musim kering pada bulan Agustus-September (Kadarsah 2007).

Selanjutnya berdasarkan Gambar 10 pula dapat disusun kategori bulan basah, lembab dan kering seperti ditunjukkan pada Tabel 1. Adapun kriteria yang digunakan untuk menentukan bulan kering, bulan lembab, dan bulan basah berdasarkan Schmidt-Fergusson dengan katagori sebagai berikut bulan kering (BK) : bulan dengan curah hujan  $< 60$  mm, bulan lembab (BL) : bulan dengan curah hujan antara 60 sampai dengan 100 mm, bulan basah (BB) : bulan dengan curah hujan  $> 100$  mm.

**Tabel 1.** Pembagian Wilayah dan Kategori Bulan Basah, Lembab dan Kering

Bulan Basah	Bulan Lembab	Bulan Kering
Jan-April, Nov-Des	Mei	Juni-Okt

#### 4. Kesimpulan

Berdasarkan data amatan curah hujan 1980-2000, pola curah hujan yang ditunjukkan tidak mengikuti pola tertentu bahkan terdapat beberapa lonjakan pola. Oleh karena itu tidak mudah menemukan teknik pendugaan yang tepat yang dapat meminimalkan galat pendugaan. Pendugaan yang dilakukan dengan Kriging, kurang memadai karena kurangnya pengaruh spasial pada data tersebut meskipun data curah hujan dianggap sebagai data yang berbasis *spatio-temporal*. Kehilangan data sebesar 5% dengan Kriging memiliki galat pendugaan rata-rata sebesar 48.39%, dengan metode rata-rata bergerak, kehilangan rata-rata data sebesar 3.7% memiliki rata-rata galat pendugaan sebesar 34.78%. Dengan demikian, metode rata-rata bergerak memadai untuk melakukan pendugaan terhadap data tidak lengkap pada kasus dalam penelitian ini. Hasil dengan metode ini pula dapat ditunjukkan bahwa rata-rata galat pendugaan (*Mean Absolute Deviation/MAD*) yang persentase galatnya lebih dari 40% sebesar

29.81% artinya 70,19% sisanya memadai sebagai data pelengkap untuk curah hujan. Demikian juga pola data curah hujan yang terbentuk dalam kurun waktu 21 tahun menunjukkan pola curah hujan monsoon.

### Daftar Pustaka

- Aldrian E, Susanto RD . 2003. Identification Of Three Dominant Rainfall Regions Within Indonesia and Their Relationship To Sea Surface Temperature. *International Journal Of Climatology* **23**: 1435–1452 (2003)
- Daniels MJ, Hogan JW. 2008. *Missing Data in Longitudinal Studies*. London: Chapman & Hall/CRC.
- Kadarsah. 2007. Tiga Pola Curah Hujan Indonesia <http://www.kadarsah.wordpress.com/2007/06/29/tiga-daerah-iklim-indonesia/>. [20 Juli 2009]
- Khoerudin M. 2010. Pendugaan Data Hilang dengan Menggunakan Metode Ordinary Kriging [Skripsi]. Bogor: Departemen Statistika FMIPA IPB
- Little RL, Rubin, DB. 1990. *Statistical analysis with missing data*. New York: Wiley.
- Schafer, JL and John W. Graham. 2002. Missing Data: Our View of the State of the Art. *Psychological Methods*. Vol. 7, No. 2, 147-177.
- Scheffer J. 2002. Dealing with Missing Data. *Res. Lett. Inf. Math. Sci.* Vol 3, 153-160
- Tsiatis A .2006. *Semiparametric Theory and Missing Data*. New York: Springer.
- Wei WWS. 1990. *Time Series Analysis: Univariate and Multivariate Methods*. USA: Addison-Wesley Publishing Co.

Lampiran

