

Kode>Nama Rumpun Ilmu : 123/Ilmu Komputer Bidang Fokus : TIK

**USULAN
PENELITIAN DISERTASI DOKTOR**



**MODEL PENDUGA CURAH HUJAN
MENGUNAKAN HYBRID MULTIVARIATE
SPATIO TEMPORALARIMA
DAN ARTIFICIAL NEURAL NETWORKS**

PENGUSUL
Ahmad Saikhu, S.Si, MT/0018077105

INSTITUT TEKNOLOGI SEPULUH NOPEMBER
Juli, 2017

HALAMAN PENGESAHAN
PENELITIAN DISERTASI DOKTOR

Judul Penelitian	: Pengembangan Model Penduga Curah Hujan menggunakan hybrid Spatio-Temporal Multivariat ARIMA dan Jaringan Syaraf Tiruan
Bidang Fokus	: Teknologi Informasi dan Komunikasi
Judul Disertasi	: Pengembangan Model Penduga Curah Hujan menggunakan hybrid Spatio-Temporal Multivariat dan Jaringan Syaraf Tiruan
Kode>Nama Rumpun Ilmu	: 459/Illmu Komputer
Peneliti	:
Nama Lengkap	: AHMAD SAIKHU S.Si, M.T
NIDN	: 0018077105
Jabatan Fungsional	: Lektor Kepala
Program Studi	: Teknik Informatika
Nomor HP/Surel	: 081357976106/saikhu@its-sby.edu
NIM	: 5115301007
Semester ke	: 4
PT Penyelenggara	: Institut Teknologi Sepuluh Nopember
Program Studi Doktor	: Ilmu Komputer
Nama Promotor	: Dr AGUS ZAINAL ARIFIN S.Kom, M.Kom
NIDN Promotor	: 0009087205
Biaya yang Diusulkan ke DRPM	: Rp 58,870,000.00
Biaya Luar Tambahan	: Rp 0.00

Kota Surabaya, 07-06-2017

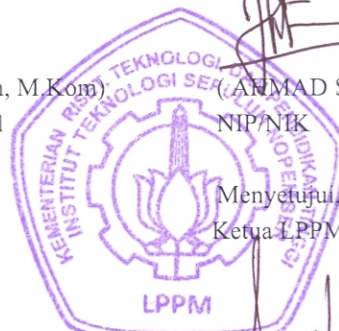


Mengetahui,
Dekan

(Dr. Agus Zainal Arifin, S.Kom, M.Kom)
NIP/NIK 197208091995121001

Ketua Peneliti

(AHMAD SAIKHU S.Si, M.T)
NIP/NIK



Menyetujui,
Ketua LPPM

(Prof. Dr. Ir. Adi Soeprijanto, MT.)
NIP/NIK 196404051990021001

IDENTITAS DAN URAIAN UMUM

1. Judul Penelitian : Pengembangan model penduga curah hujan menggunakan hybrid spatio-temporal multivariat ARIMA dan Jaringan Syaraf Tiruan

2. Tim Peneliti

No.	Nama	Jabatan	Bidang Keahlian	Instansi Asal	Alokasi Waktu (jam minggu)
1.	Ahmad Saikhu	Ketua	Data Mining	Institut Teknologi Sepuluh Nopember	40

3. Subyek Penelitian (jenis material yang akan diteliti dan segi penelitian)
Pemodelan dan analisis data observasi cuaca dan iklim (harian/bulanan) di wilayah Indonesia
4. Masa Pelaksanaan
Mulai : bulan:Januari 2018
Berakhir : bulan:Oktober 2018
5. Usulan Biaya DRPM Ditjen Penguatan Risbang
• Tahun ke-1 : Rp 58.870.000,00
6. Lokasi Penelitian (lab/studio/lapangan) : laboratorium dasar dan terapan komputasi (DTK)
7. Prediksi lulus S-3. Dilengkapi lampiran surat pernyataan pengusul diketahui promotor bahwa pengusul diprediksi belum lulus S3 ketika proposal mendapat persetujuan didanai.
Prediksi lulus S-3: September 2019. Surat keterangan sebagai mahasiswa aktif dan surat rekomendasi dari promotor, terlampir.
8. Temuan yang ditargetkan dan/atau implikasi teoretik (termasuk penjelasan gejala atau kaidah, metode, teori, pendekatan-perspektif, atau antisipasi yang dikontribusikan pada bidang ilmu)
Dari penelitian ini ditargetkan menemukan model penduga yang efektif dan efisien untuk berbagai data time series multivariat yang bersifat spatio-temporal (hidrologi, ekonomi, biomedis, transportasi, dll) dengan pendekatan *data mining* (dan *machine learning*). Selain berkontribusi utama pada bidang ilmu komputer, pada studi kasus yang dipilih, yaitu data sintetik dan data cuaca/iklim, akan berkontribusi kepada bidang *geoscience* dan lingkungan.
9. Kontribusi mendasar pada suatu bidang ilmu (uraikan tidak lebih dari 50 kata, tekankan pada gagasan fundamental dan orisinal yang akan mendukung pengembangan iptek sosbud)

Pengembangan teoritis untuk ekstraksi fitur dan seleksi fitur *time-lags* berdasarkan relationship linier dan non linier. Proses ini penting pada data mining untuk berbagai tujuan analisis. Kedua, model prediksi hybrid MSTARIMAX dan ANN dengan dekomposisi data input berdasar relationship dan optimalisasi ANN. Kontribusi aplikatifnya adalah pemodelan simultan untuk bidang meteorologi dengan fitur lokal/global.
10. Jurnal ilmiah yang menjadi sasaran (tuliskan nama terbitan berkala ilmiah internasional bereputasi, dan tahun rencana publikasi)

Ecological Informatics (ISSN: 15749541, Elsevier BV), Q2

SURAT REKOMENDASI PENELITIAN DISERTASI DOKTOR

Yang bertanda tangan di bawah ini :

Nama : Dr. Agus Zainal Arifin, S.Kom, M.Kom
NIP : 19720809 199512 1 001
Jabatan : Promotor

Memberikan rekomendasi untuk penelitian disertasi doktor atas mahasiswa

Nama : Ahmad Saikhu
NRP : 5115301007
Program Studi : S3 Ilmu Komputer
Judul Penelitian : Pengembangan Model Penduga Curah Hujan
menggunakan hybrid Spatio-Temporal Multivariat ARIMA
dan Jaringan Syaraf Tiruan
Instansi Asal : Institut Teknologi Sepuluh Nopember

Demikian surat rekomendasi ini dibuat untuk dapat digunakan sebagaimana mestinya.

Surabaya, 7 Juni 2017

Promotor,



Dr. Agus Zainal Arifin, S.Kom, M.Kom
NIP. 19720809 1995121 001

Mengetahui,
Koordinator Program Studi S3 Ilmu Komputer,



Waskitho Wibisono, S.Kom, M.Eng, Ph.D
19741022 200003 1 001

SURAT KETERANGAN

Yang bertanda tangan di bawah ini :

Nama Promotor : Dr. Agus Zainal Arifin, S.Kom, M.Kom
NIP : 19720809 199512 1 001
Jabatan : Dekan Fakultas Teknologi Informasi dan Komunikasi


Menerangkan bahwa mahasiswa Program Pascasarjana Institut Teknologi Sepuluh Nopember di bawah ini :

Nama : Ahmad Saikhu
NRP : 5115301007
Program Studi : S3 Ilmu Komputer
Judul Penelitian : Pengembangan Model Penduga Curah Hujan
menggunakan hybrid Spatio-Temporal Multivariat ARIMA
dan Jaringan Syaraf Tiruan
Instansi Asal : Institut Teknologi Sepuluh Nopember

Adalah mahasiswa Program Doktor di Institut Teknologi Sepuluh Nopember tahun masuk 2015. Yang bersangkutan hingga saat ini masih aktif studi/proses penelitian/penulisan disertasi dan sudah menempuh selama 4 semester terhitung sampai dengan bulan Juni 2017.

Demikian surat keterangan ini dibuat untuk dapat digunakan sebagaimana mestinya.

Surabaya, 7 Juni 2017

Mengetahui
Koordinator Program Studi S3 Ilmu Komputer,

Waskitho Wibisono, S.Kom, M.Eng, Ph.D
19741022 200003 1 001

Promotor,


Dr. Agus Zainal Arifin, S.Kom, M.Kom
NIP. 19720809 1995121 001

DAFTAR ISI

LEMBAR PENGESAHAN	iii
Identitas dan Uraian Umum	v
DAFTAR ISI.....	vii
RINGKASAN	ix
BAB 1 PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Perumusan Masalah.....	4
1.3. Batasan Masalah.....	4
1.4. Tujuan dan Manfaat Penelitian	5
1.5. Kontribusi dan Target Penelitian.....	5
BAB 2 KAJIAN PUSTAKA	6
2.1. Penelitian Terkait	6
2.2. Cuaca dan Iklim	7
2.3. Multivariate Time Series	8
2.4. Seleksi Fitur.....	9
2.5. Studi Pendahuluan dan Kebaharuan.....	10
BAB 3 METODE PENELITIAN	11
BAB 4 BIAYA DAN JADWAL PENELITIAN	14
4.1. Anggaran Biaya	14
4.2. Jadwal Penelitian.....	15
DAFTAR PUSTAKA	16
Lampiran 1. Dukungan Sarana dan Prasarana	
Lampiran 2. Biodata Peneliti	
Lampiran 3. Surat Pernyataan Peneliti	

RINGKASAN

Cuaca, iklim dan musim senantiasa akrab dalam kehidupan manusia sehari-hari. Salah satu komponen penting dalam cuaca dan iklim adalah curah hujan. Curah hujan di suatu wilayah diakibatkan oleh interaksi berbagai variabel cuaca dan iklim global. Curah hujan merupakan sistem yang kompleks di mana hubungan antar variabel lokal bersifat simultan dan dipengaruhi waktu, wilayah dan kondisi iklim global yang dinamis. Oleh karenanya, hubungan antar variabel tidak hanya bersifat sesaat tetapi terkait dengan kondisi wilayah sekitar dan observasi cuaca/iklim pada periode sebelumnya.

Pemodelan curah hujan di suatu wilayah cukup penting untuk perencanaan sumber daya air secara optimal, perkiraan cadangan air di musim kering/kemarau dan merupakan penelitian penting dalam bidang hidrologi. Berbagai penelitian telah dilakukan untuk pemodelan curah hujan, baik pendekatan model teoritis/konseptual maupun pendekatan data empiris/histori. Pemodelan konseptual bertumpu pada interaksi/proses fisis dinamika atmosfer, suhu permukaan laut, dan posisi matahari. Model empiris lebih disukai untuk menemukan hubungan antara data input dan output ketika pengguna tidak memahami secara lengkap hubungan faktor-faktor fisik yang terjadi

Salah satu pendekatan model empiris adalah metode ARIMA yang memiliki kelemahan kurang sesuai untuk data non-stasioner dan non linier sehingga akurasinya rendah. Untuk mengatasi hal ini, pendekatan lain adalah menggunakan berbagai metode *machine learning* yang juga memiliki kelemahan lambat mencapai konvergensi maupun terjadinya optimum lokal. Untuk mengatasi hal tersebut, diantaranya adalah mengidentifikasi sebanyak mungkin variabel prediktor dengan mempertimbangkan faktor waktu dan wilayah (spatio temporal). Pilihan lainnya adalah dengan dan mengembangkan model gabungan (hybrid). Namun hal ini menyebabkan komputasi tidak efisien.

Dalam penelitian ini, diajukan sejumlah pengembangan yaitu : Ekstraksi dan seleksi fitur input yang memperhitungkan pola hubungan antar variabel yang bersifat linier dan non linier serta memasukkan variabel waktu dan lokasi. Variabel input diekstraksi dengan metode *Detrend Partial Cross Correlation* dan *Partial Cross Mutual Information*. Sedangkan untuk seleksi fitur yang relevan dan tidak redundan, digunakan metode *Pearson Correlation and Symmetrical Uncertainty Filter* (PCSUF). Hasil atas pra-proses ini digunakan sebagai input dalam pengembangan model *hybrid Multivariate Spatio-Temporal Autoregressive Integrated Moving Average with Exogenous Variables* (MST-ARIMAX) dan *Artificial Neural Network*. Proses training ANN pada model hybrid tersebut dilakukan pengembangan dengan PSO ataupun GA.

Uji coba kinerja model penduga curah hujan dilakukan pada data cuaca dan indeks iklim meliputi 12 variabel pada 9 lokasi (stasiun) di Jawa Timur dengan berbagai scenario metode pembandingan, yaitu metode MST-ARIMAX saja, ANN saja dan metode hybrid. Evaluasi atas kinerja model meliputi RMSE, MAE, R^2 , dan korelasi.

Luaran atas hasil penelitian ini adalah publikasi pada konferensi internasional, jurnal internasional terindeks *scopus* dan draft laporan penelitian disertasi. Diharapkan dengan pendanaan penelitian disertasi doktor, peneliti dapat menyelesaikan studi S3 Ilmu Komputer pada Februari 2019.

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Cuaca, iklim dan musim senantiasa akrab dalam kehidupan manusia sehari-hari. Gejala alam tersebut berpengaruh kepada banyak sektor kehidupan, diantaranya pertanian, kehutanan, kemaritiman, perikanan, kesehatan, transportasi, pariwisata, energi dan infrastruktur. Dalam kondisi yang ekstrim, cuaca dan iklim bahkan dapat menyebabkan terjadinya bencana, diantaranya wabah penyakit, banjir, angin kencang, tanah longsor, kekeringan dan kebakaran hutan/lahan. Oleh karenanya, informasi, analisis dan pengetahuan terhadap kondisi cuaca dan iklim serta pola perubahan iklim sangat penting untuk dipelajari agar manusia dapat beradaptasi dengan perubahan yang dinamis, menghindarkan atau mengurangi terjadinya resiko akibat bencana dan melakukan tindakan mitigasi.

Dalam mengenali pola cuaca dan iklim dilakukan pengukuran terhadap suhu udara, kelembaban relatif, tekanan udara, penguapan air, kecepatan angin, suhu permukaan laut, lamanya penyinaran matahari dan curah hujan(Makmur, 2013). Nilai variabel tersebut berbeda bergantung pada sejumlah faktor, yaitu ketinggian tempat, garis lintang, tekanan, arus laut dan permukaan tanah. Pengaruh timbal balik antar faktor akan menentukan pola dari nilai variabel. Batas nilai variabel untuk keadaan yang stabil telah ditentukan, namun dapat terjadi penyimpangan yang mengakibatkan masalah tertentu, misalnya angin dengan kecepatan tinggi dan suhu ekstrim.

Salah satu komponen penting dalam cuaca dan iklim adalah curah hujan. Curah hujan di suatu wilayah diakibatkan oleh interaksi berbagai variabel cuaca dan iklim global. Curah hujan merupakan sistem yang kompleks di mana hubungan antar variabel lokal bersifat simultan dan dipengaruhi waktu, wilayah dan kondisi iklim global yang dinamis. Oleh karenanya, hubungan antar variabel tidak hanya bersifat sesaat tetapi terkait dengan kondisi wilayah sekitar dan observasi cuaca/iklim pada periode sebelumnya.

Pemodelan curah hujan di suatu wilayah cukup penting untuk perencanaan sumber daya air secara optimal, perkiraan cadangan air di musim kering/kemarau dan merupakan penelitian penting dalam bidang hidrologi. Berbagai penelitian telah dilakukan untuk pemodelan curah hujan, baik pendekatan model teoritis/konseptual maupun pendekatan data empiris/histori. Pemodelan konseptual bertumpu pada interaksi/proses fisis dinamika atmosfer, suhu permukaan laut, dan posisi matahari(Swarinoto, 2013).

Model empiris lebih disukai untuk menemukan hubungan antara data input dan output ketika pengguna tidak memahami secara lengkap hubungan faktor-faktor fisik yang terjadi (Asadi, 2013). Nilai akurasi prediksi curah hujan lebih utama pada riset hidrologi. Di samping itu, pendekatan konseptual memerlukan proses komputasi yang mahal, kompleksitas tinggi dan padat peralatan *scientific* yang dalam beberapa tahun memerlukan pembaharuan peralatan (Zou, 2014).

Saat ini pemodelan empiris yang banyak digunakan untuk prediksi curah hujan adalah pendekatan statistik dan *data mining/machine learning*. Prediksi curah hujan dengan metode statistik adalah menggunakan pendekatan *time-series*, antara lain dengan *Autoregressive Integrated Moving Average* (ARIMA) dan *Seasonal ARIMA with Exogenous Input* (SARIMAX). Pendekatan ini memiliki sejumlah keterbatasan diantaranya adalah hanya berlaku untuk pola data yang bersifat stasioner dan pola hubungan yang linier (Asadi dkk, 2013).

Untuk memperbaiki kelemahan dari model statistik, telah dikembangkan berbagai model prediksi curah hujan yang berbasis pada kecerdasan buatan, baik menggunakan metode *machine-learning* secara tunggal maupun gabungan (*hybrid*) antara metode statistik dan *machine-learning*. Prediksi curah hujan harian/bulanan menggunakan *Bayesian Network* untuk sejumlah stasiun cuaca di wilayah India dan Canada (Sharma dkk, 2014; Rasouli, 2012), Han dkk (2015) menggunakan *k-NN* dan *Mutual Information* (PMI) untuk prediksi curah hujan dari data meteorologi di wilayah Dalian (China).

Selanjutnya, Ramana dkk (2013) menggunakan metode *Wavelet Neural Network* untuk data curah hujan bulanan di wilayah India. *Support Vector Machine* juga digunakan untuk prediksi curah hujan harian(Monedero dkk, 2014; Garcia

dkk, 2014). Pendekatan *Neuro-Fuzzy* antara lain dilakukan oleh Hashim dkk (2016) untuk prediksi curah hujan bulanan di wilayah Patna, India. Kekurangan dari pendekatan *Machine Learning* antara lain adalah belum dapat mengakomodasi pola linier sebaik metode statistic.

Pendekatan model *hybrid* dan *ensemble* antara lain menggabungkan metode *Self Organized Map* (SOM) dan *Nonlinear Autoregressive with Exogenous Inputs* (NARX) *network* (Chang dkk, 2016), pendekatan *neurocomputing* dengan menggabungkan *Multiple Linear Regression* dan *Extreme Learning Machine* untuk memprediksi curah hujan bulanan dengan prediktor anomali suhu permukaan laut (Acharya dkk, 2014). Model hybrid DARIMA-NARX digunakan oleh Banihabib dkk (2017) untuk prediksi curah hujan harian dengan prediktor *North Atlantic Oscillation* (NAO). Penggabungan antara ANN, *Mutual Information* dan PSO digunakan untuk prediksi curah hujan bulanan di wilayah Australia (He dkk, 2015). Asadi dkk menggabungkan ANN dengan GA dan algoritma Levenverg-Marquardt untuk prediksi curah hujan harian di wilayah India.

Pemodelan curah hujan dengan studi kasus wilayah Indonesia telah dilakukan dengan pendekatan ARIMA, *machine-learning* dan berbagai variasinya. Sejumlah penelitian prediksi curah hujan di wilayah Semarang, Tenggarong, Jakarta, Bandung Malang, dll. (Setyaningrum, 2014; Nugroho, 2015; Lee, 2015; Hardiwinarto, 2015; Nhita, 2015; Utami, 2016; Munandar, 2015; Nisak, 2016). Metode yang digunakan a.l. ARIMA, ANN, ANFIS, sistem Fuzzy dan SVR.

Berdasarkan studi curah hujan di Indonesia (Toggweiler, 2001; Hamada, 2002; Aldrian, 2003; Hendon, 2003; Supari, 2012), dinyatakan bahwa curah hujan dipengaruhi oleh berbagai faktor lokal (cuaca) dan iklim regional (global) di sekitar wilayah tersebut. Oleh karenanya penentuan variabel prediktor yang relevan adalah penting dengan mempertimbangkan pola hubungan antar variabel.

Meskipun beberapa penelitian pemodelan iklim di suatu wilayah/negara memiliki akurasi yang tinggi, belum tentu model tersebut sesuai untuk wilayah lainnya dikarenakan perbedaan iklim, posisi geografis, topografi, dan proporsi daratan/maritim. Sebagai contoh, model iklim untuk wilayah/negara lintang tinggi, akan berbeda dengan negara pada lintang rendah (tropis). Hal ini menyebabkan

bahwa model untuk mengenal pola iklim dan prediksi hujan menjadi khas untuk suatu wilayah tertentu.

Oleh karenanya dalam penelitian ini diusulkan suatu pendekatan (skema) komprehensif untuk model penduga curah hujan di wilayah Indonesia. Skema dimaksud yaitu dengan melibatkan sejumlah fitur iklim dalam ruang lingkup lokal, regional/global dengan memperhitungkan faktor *time-lag* dan lokasi (*spatial*).

Periode data yang digunakan untuk pemodelan adalah pengamatan harian, dan per bulan. Kedua adalah dengan melakukan pemodelan secara bertahap, yaitu proses interpolasi data kosong, ekstraksi fitur *time-lags*, pemilihan input variabel (*input selection*). Ketiga adalah dengan mengembangkan model hybrid *Spatio-Temporal* Multivariat dan Jaringan Syaraf Tiruan untuk prediksi curah hujan.

1.2. Perumusan Masalah

Permasalahan pada penelitian ini adalah:

- a. Bagaimana melakukan pra-proses data untuk kondisi *missing-value*?
- b. Bagaimana melakukan ekstraksi variabel prediktor dari curah hujan dengan pola hubungan yang linier dan non linier?
- c. Bagaimana melakukan seleksi variabel prediktor sehingga terpilih yang relevan dan tidak redundan dengan pola hubungan yang linier dan non linier?
- d. Bagaimana mengembangkan model hybrid MST-ARIMA dan ANN untuk prediksi curah hujan dengan memperhitungkan waktu dan lokasi?

1.3. Batasan Masalah

Dalam penelitian ini, batasan masalah yang digunakan yaitu data observasi cuaca harian yang diperoleh dari BMKG (<http://dataonline.bmkg.go.id/>) untuk sejumlah wilayah di Jawa Timur. Data indeks iklim global periode per bulan diperoleh dari <http://climexp.knmi.nl>. Periode data observasi yang digunakan untuk mengembangkan model penduga curah hujan adalah periode Januari 1981 – Desember 2013, sedangkan untuk pengujian kesesuaian model, digunakan data periode Januari 2014-Desember 2016.

Variabel prediktor lokal yang digunakan dari data BMKG adalah minimum/mean/maksimum temperatur, kelembaban udara, lama penyinaran,

kecepatan angin dan maksimum kecepatan angin. Sedangkan variabel prediktor global yang digunakan dari data KNMI adalah data anomali suhu permukaan laut ENSO (Nino1.2, Nino3, Nino3.4, Nino4) dan DMI (Dipole Mode Index).

1.4. Tujuan dan Manfaat Penelitian

Tujuan dari penelitian ini adalah membangun model penduga curah hujan menggunakan *hybrid* Spatio-Temporal Multivariat ARIMA dan Jaringan Syaraf Tiruan.

Manfaat dari penelitian ini, secara khusus adalah tersedianya model penduga curah hujan dengan akurasi tinggi untuk periode harian/bulanan yang melibatkan sejumlah variabel input secara simultan, baik variabel cuaca/iklim, waktu maupun wilayah. Model ini diharapkan sesuai digunakan untuk wilayah Indonesia yang berkarakteristik ZOM. Sedangkan manfaat umum dari penelitian ini adalah dapat digeneralisasi bahwa untuk data time-series yang bersifat *spatio-temporal* dapat dimodelkan dengan pendekatan *hybrid* Spatio-Temporal Multivariat ARIMA dan Jaringan Syaraf Tiruan.

1.5. Kontribusi dan Target Penelitian

Kontribusi penelitian ini dapat ditinjau dari 2 hal, yaitu kontribusi teoritis dan kontribusi aplikatif.

- a. Ekstraksi Fitur Time-Lags berdasarkan relationship linier dan non linier.
- b. Seleksi Fitur yang terpisah antara pola relationship linier dan non linier untuk memperoleh fitur prediktor yang relevan dan tidak redundan.
- c. Pemodelan hybrid MSTARIMAX dan ANN dengan dekomposisi data input berdasar relationship dan optimalisasi proses training pada ANN.

Sedangkan kontribusi aplikatif adalah:

- a. Prediktor tidak hanya 7 variabel lokal dari BMKG, namun juga ditambahkan fitur regional/global dari sumber lain.
- b. Pendekatan Spatio-Temporal dengan memperhitungkan dependensi Curah Hujan tidak hanya terhadap waktu (time-lags), tetapi juga lokasi (Space) di antara pengamatan cuaca (dan iklim) pada wilayah yang saling berdekatan.

BAB 2 KAJIAN PUSTAKA

2.1 Penelitian terkait

Ada banyak studi metode *machine learning* untuk prediksi curah hujan. Selain yang telah diuraikan pada sub bab 1.1, berikut disajikan ringkasan tentang studi penggunaan model empiris yang telah dilakukan sejumlah peneliti. Chauhan (2014) telah melakukan studi perbandingan berbagai teknik data mining untuk prediksi cuaca, khususnya dari 13 penelitian iklim/cuaca. Dari penelitian tersebut, teknik data mining yang digunakan didominasi oleh *Decision Tree* (6), *Clustering* (4) dan selebihnya menggunakan Regresi. Sedangkan Algoritma yang digunakan cukup beragam, yaitu C4.5/C5.0, *k-means*, *kNN*, ANN dan MLR.

Variabel yang digunakan sebagai prediktor pada umumnya adalah variabel lokal dengan jumlah minimum 3 variabel yaitu tekanan udara, curah hujan dan kelembaban. Data histori yang digunakan dari pengamatan 1 s/d 50 tahun dengan ukuran 48 s/d 400 ribu *records*. Akurasi hasil prediksi yang dihasilkan sangat beragam diantara 53% - 93%. Keberagaman nilai akurasi ini antara lain disebabkan ukuran data, tidak adanya pra-proses dan beberapa metode gagal pola non linier.

Sedangkan Kaur (2012) telah melakukan studi perbandingan berbagai metode *machine learning* untuk pemodelan data meteorologi/klimatologi, a.l. ANN, GA, *Fuzzy Logic*, dan CART. Dari kajian ini direkomendasikan bahwa algoritma berbasis ANN menghasilkan kinerja yang lebih baik. Kinerja akan meningkat apabila dilakukan seleksi fitur terlebih dahulu dengan menggunakan teknik statistik. Kohail (2011) dalam penelitiannya tentang penggunaan teknik data mining untuk data meteorologi menerapkan teknik *pre-processing*, analisis *outliers*, *clustering*, prediksi, klasifikasi dan *association rules mining*. Algoritma dengan hasil akurasi lebih baik adalah menggunakan ANN (85%), dibandingkan dengan Naive-Bayes, *kNN* dan *Decision Tree*.

Abbot (2013) melakukan seleksi input atas variabel lokal dan global serta mengoptimasi prediksi hujan bulanan menggunakan ANN untuk studi kasus di *Queensland*. Sedangkan Tran (2015) menggunakan *Partial Linear Correlation* dan *Partial Mutual Information* untuk memilih variabel input yang signifikan pada

model penduga curah hujan. Han (2015) menggunakan *Joint Mutual Information* untuk seleksi variabel input. Seleksi variabel input menggunakan *binary-coded PSO* dan *ELM* oleh May (2008) dan Tourmina (2015) dan menggunakan pendekatan non-linear untuk seleksi variabel input.

Pendekatan terbaru untuk model penduga curah hujan adalah menggunakan metode hybrid berbasis *machine learning* untuk mengatasi keterbatasan metode statistik dan *machine learning* sekaligus meningkatkan akurasi hasil prediksi. Asadi (2013) menggabungkan metode Neural Network dengan GA serta Algoritma Levenverg-Marquardt untuk prediksi curah hujan harian dan berhasil mengurangi RMSE dibanding model tunggal. Chang dkk (2016) membuat model hybrid PSO dan *Autoregressive Non Linear with exogenous input* (NARX) dengan hasil $R^2 > 0,9$. Gabungan antara ARIMA dan Autoregressive Non Linear (NARX) berhasil memperbaiki akurasi ARIMA maupun Neural Network (Banihabib, 2017). Dari pendekatan model hybrid, diusulkan pendekatan spatio-temporal dengan variabel input yang simultan untuk memperoleh akurasi lebih baik (efektif) dan jumlah prediktor yang efisien.

2.2 Cuaca dan Iklim

Cuaca adalah keadaan udara (atmosfir bumi) pada saat tertentu di wilayah tertentu yang relatif sempit dan pada jangka waktu yang singkat. Cuaca terjadi karena suhu dan kelembaban yang berbeda antara satu tempat dengan tempat lainnya. Perbedaan ini dapat terjadi karena sudut pemanasan matahari yang berbeda dari satu tempat ke tempat lainnya akibat perbedaan lintang bumi. Pemantauan gejala perilaku cuaca yang terjadi di atas permukaan bumi dilakukan antara lain oleh satelit, termometer, barometer, anemometer dan instrumentasi lainnya (Swarinoto, 2012; Betung, 2013).

Iklim adalah keadaan cuaca rata-rata dalam waktu yang relatif lama dan meliputi wilayah yang lebih luas. Iklim di suatu tempat di bumi dipengaruhi oleh letak geografis dan topografi tempat tersebut. Pengaruh posisi relatif matahari terhadap suatu tempat di bumi menimbulkan musim, suatu penciri yang membedakan iklim satu dari yang lain. Ratag (2008) menjelaskan bahwa sistem iklim terdiri atas atmosfer daratan, salju dan es, lautan dan badan-badan air lainnya,

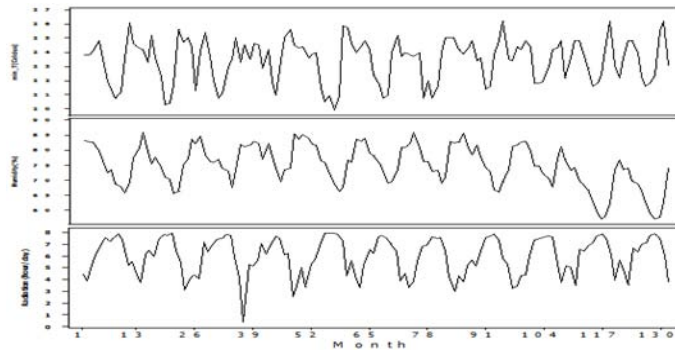
serta makhluk hidup. Iklim merupakan suatu sistem yang kompleks dan interaktif. Sistem berevolusi terhadap waktu oleh adanya pengaruh-pengaruh yang berasal dari dinamika internalnya dan juga sebagai akibat terjadinya perubahan pada faktor-faktor luar yang mempengaruhi iklim.

2.3 Multivariate Time Series

Time Series adalah representasi data yang terurut berdasarkan waktu kejadian dengan interval waktu yang sama di mana antar nilai saling berhubungan (Vasimalla, 2014). Data time series dapat dinotasikan sebagai

$$X_i(t); [I = 1, 2, 3, \dots, n, t = 1, 2, 3, \dots, m]$$

Jika nilai $n > 2$, disebut sebagai MTS, sementara apabila $n = 1$ dinamakan *univariate time series*. Gambar 2.1 adalah merupakan contoh data time series multivariat.



Gambar 2.1 Contoh data time series meteorologi

Untuk mengidentifikasi hubungan antar variabel berdasarkan *time-lags* tertentu, alat analisis yang umum digunakan adalah Cross Correlation Function (CCF). Untuk menghitung CCF digunakan persamaan (1).

$$r_{xy}(T) = \frac{\sigma_{xy}(T)}{\sqrt{\sigma_{xx}(0)}\sqrt{\sigma_{yy}(0)}} \quad (1)$$

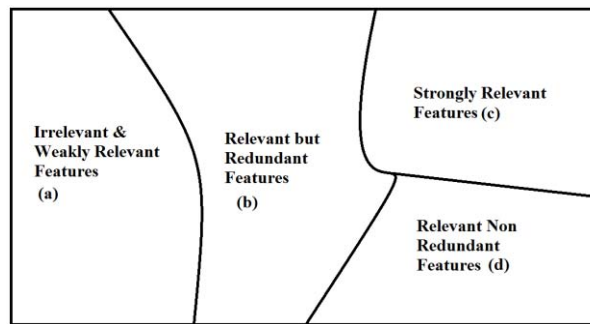
di mana $r_{xy}(T)$ adalah korelasi variabel x dan variabel y pada periode ke T dan σ_{xy} adalah kovariansi(x,y), σ_{xx} adalah variansi(x) dan σ_{yy} adalah variansi(y).

Dalam hubungan antara dua time series (y_t dan x_t), series y_t mungkin berhubungan dengan lags yang lalu dari series-x. Sampel CCF membantu mengidentifikasi lags dari variabel x yang mungkin merupakan kandidat prediktor dari y_t . Sampel CCF didefinisikan sebagai himpunan korelasi sampel antara x_{t+h} dan y_t untuk $h = 0, \pm 1, \pm 2, \pm 3$, dan seterusnya. Nilai negative dari h menunjukkan bahwa

korelasi antara variabel x pada waktu sebelum t dan variabel y pada saat t . Misalnya, $h = -2$ berarti bahwa nilai korelasi yang dihitung adalah x_{t-2} and y_t . Hasil dari nilai korelasi dapat diuji dengan membandingkan nilai korelasi silang dengan *standard error* $1/\sqrt{n-k}$.

2.4 Seleksi fitur

Dari studi (Yu dkk, 2004), fitur input untuk pemodelan MTSC dibagi menjadi 3 subset yang saling lepas, yaitu relevan yang sangat kuat, relevan yang lemah dan fitur yang tidak relevan. Subset pertama mutlak/harus ada sebagai prediktor dari label klas dan tidak dapat dihapus karena akan berpengaruh kepada hasil prediksi. Subset kedua kurang diperlukan sebagai prediktor, kecuali pada kondisi tertentu. Sedangkan subset ketiga tidak diperlukan sebagai prediktor pada kondisi apapun. Pembagian fitur input atas 3 subset dijelaskan pada gambar 2.2.



Gambar 2.2 fitur input yang relevan dan tidak relevan

Untuk melakukan seleksi fitur input yang relevan, pada umumnya asosiasi antar fitur diasumsikan linier. Oleh karenanya, suatu fitur disebut relevan terhadap label klas jika korelasinya signifikan. Formula untuk menentukan fitur relevan adalah dengan menggunakan korelasi Pearson, yang ditulis pada persamaan 2.

$$\rho_{xy} = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}} \quad (2)$$

ρ_{xy} adalah korelasi antara variabel x dan variabel y . ρ_{xy} disebut signifikan jika nilainya lebih dari nilai kritis tabel korelasi dengan derajat bebas dan level signifikansi tertentu.

Sedangkan untuk mengevaluasi asosiasi (hubungan) antar fitur yang bersifat non linier, digunakan *symmetrical uncertainty* (SU).

2.5 Studi Pendahuluan dan Kebaharuan

Peneliti telah melakukan studi ekstraksi dan seleksi fitur dengan menggunakan metode CCF dan PCSUF pada data meteorologi stasiun BMKG Juanda untuk curah hujan bulanan berdasarkan 9 variabel prediktor lokal dan indeks iklim. Data yang digunakan adalah periode Januari 2001 – Desember 2015. Hasil penelitian dalam proses *review* untuk dipublikasikan pada Jurnal Nasional JIKI Universitas Indonesia.

Hasil penelitian tersebut menunjukkan bahwa dengan metode PCSUF dalam seleksi variabel input, berhasil mengurangi jumlah variabel prediktor sebesar 77,9% (68 menjadi 15 fitur prediktor) sekaligus meningkatkan nilai akurasi menjadi 2,25 kali lipat.

Peneliti juga telah melakukan studi perbandingan kinerja 3 model penduga curah hujan, yaitu metode ARIMA, *Single-Input* dan *Multi-Input Transfer Function*. Studi kasus yang digunakan adalah data curah hujan dari periode Januari 1981 – Desember 2016 dan 7 fitur prediktor lokal dan 5 prediktor global. Hasil dari studi ini menunjukkan bahwa model *Multi-Input Transfer Function* merupakan model terbaik dengan 4 variabel prediktor, yaitu temperature rata-rata, temperature maksimum, kelembaban udara dan lama penyinaran. Hasil studi ini dalam proses *review* untuk seminar internasional ISICO, Oktober 2017.

Kontribusi dan kebaruan dari penelitian ini adalah diperolehnya sebuah pengembangan metode ekstraksi dan seleksi fitur secara simultan untuk pola hubungan yang bersifat linier dan non linier dari data time-series multivariat. Hasil atas pra-proses ini digunakan sebagai input dalam pengembangan model *hybrid* MST-ARIMAX dan ANN. Proses training ANN pada model hybrid tersebut dilakukan pengembangan dengan PSO ataupun GA.

BAB 3

METODE PENELITIAN

Sejumlah langkah diperlukan untuk pemodelan iklim dan prediksi curah hujan di wilayah Indonesia dengan studi kasus 9 stasiun cuaca di Jawa Timur. Pada gambar 3.1, dijelaskan diagram alir untuk pemodelan sebagai berikut:

a. Pengumpulan data.

Untuk data curah hujan, suhu min/mean/maks, kelembaban udara, kecepatan angin, dan kecepatan angin maksimum diperoleh dari BMKG (<http://dataonline.bmkg.go.id/>) untuk 9 stasiun cuaca di Jawa Timur. Sedangkan untuk data indeks iklim bulanan, yaitu anomali suhu permukaan laut yang meliputi Nino1.2, Nino2, Nino3, Nino3.4, Nino4 dan DMI diperoleh dari The Royal Netherlands Meteorological Institute (KNMI) secara online pada <http://climexp.knmi.nl/>. Periode data yang digunakan pada penelitian ini adalah Januari 1981 – Desember 2016, baik harian maupun bulanan.

b. Tahap pra-proses

Dilakukan interpolasi dari data yang tidak tersedia (*missing value*) maupun pencilan (*outliers*), khususnya data observasi cuaca, di mana *missing value* dan *outliers* antara 4,7% - 16,3%. Selanjutnya, untuk pemodelan data deret waktu bulanan, pengamatan cuaca harian dilakukan agregasi menjadi data bulanan.

c. Tahap ekstraksi dan seleksi fitur

Pemodelan dan prediksi curah hujan merupakan sistem yang kompleks dimana melibatkan interaksi berbagai faktor/kondisi di darat, atmosfer dan laut dengan banyak fitur yang diukur dari faktor tersebut. *Time-lag* dan pengaruh iklim global juga diperhitungkan dalam pemodelan. Pada tahap ini, dilakukan ekstraksi fitur pengaruh *time-lags* terhadap curah hujan dengan pola relasi yang linier/non linier. Membangun model *hybrid* MST-ARIMA dan ANN

Model ini merupakan pengembangan dari prediksi data deret waktu multivariate yang memiliki pola hubungan linier dan non linier antara variabel prediktor dengan variabel respons. Pada *Multivariate Spatio-Temporal Autoregressive Integrated Multivariate* (MST-ARIMA), hubungan antar lokasi dan waktu

diperhitungkan dalam model dan mengasumsikan bahwa hubungan antar variabel bersifat linier. Sedangkan untuk pola hubungan yang bersifat non-linier, digunakan metode *Artificial Neural Networks* (ANN) dengan berbagai pengembangan arsitektur dan pembelajaran. Hasil keduanya digabungkan untuk memperoleh hasil prediksi terbaik.

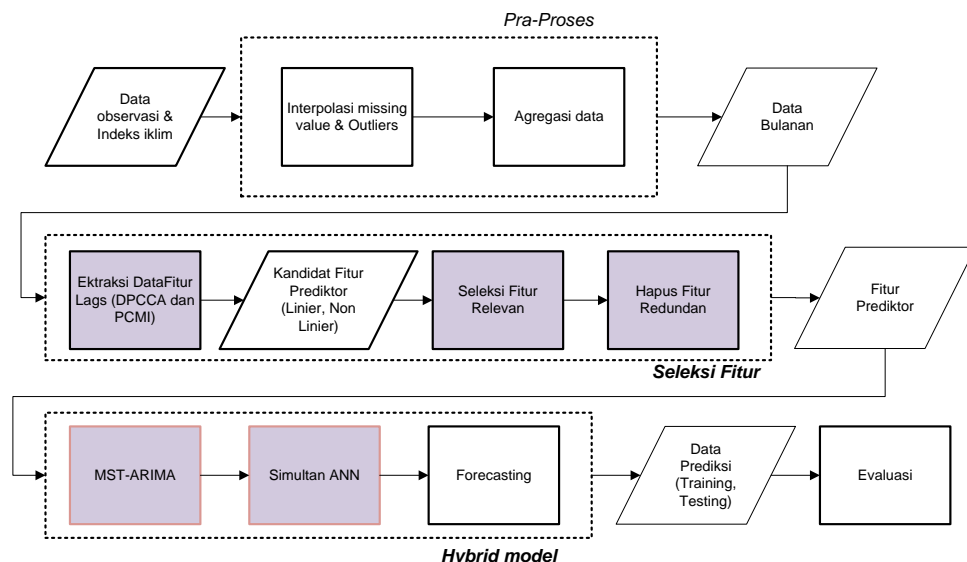
d. Prediksi data training dan data testing

Model *hybrid* diterapkan pada data training dan data testing untuk mengetahui kinerja dari model dalam melakukan peramalan curah hujan.

e. Tahap evaluasi

Evaluasi kinerja model diukur melalui nilai RMSE, MAE, R-square dan Korelasi (r) antara nilai actual dan nilai prediksi. Model memiliki kinerja tinggi apabila memiliki nilai RMSE dan MAE yang rendah serta R-square dan r yang tinggi. Uji coba model dilakukan melalui beberapa scenario, yaitu mengevaluasi kinerja model untuk MST-ARIMA, model ANN dan model *hybrid*. Skenario lainnya adalah uji coba tanpa menggunakan ekstraksi fitur maupun seleksi fitur untuk mengevaluasi efektifitas dari tahapan ini. Proporsi antara data training dan data testing juga diuji coba pada berbagai nilai rasio.

Tahap a-f dijelaskan pada gambar 3.1.

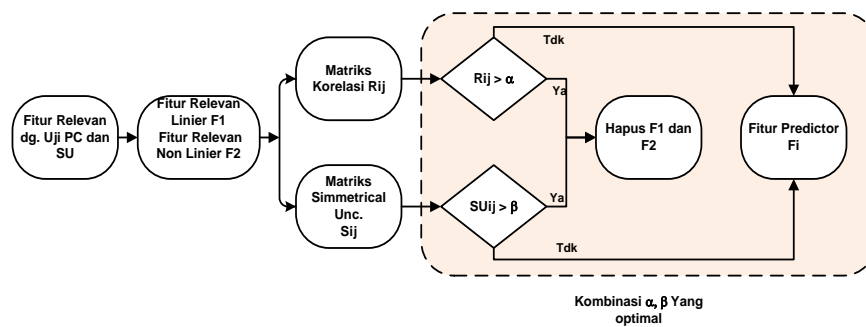


Gambar 3.1 Diagram alir model penduga curah hujan

Tahap penting dari pengembangan model ini adalah ekstraksi fitur, seleksi fitur relevan/tidak redundant, dan pengembangan model *hybrid*. Tahap ekstraksi fitur

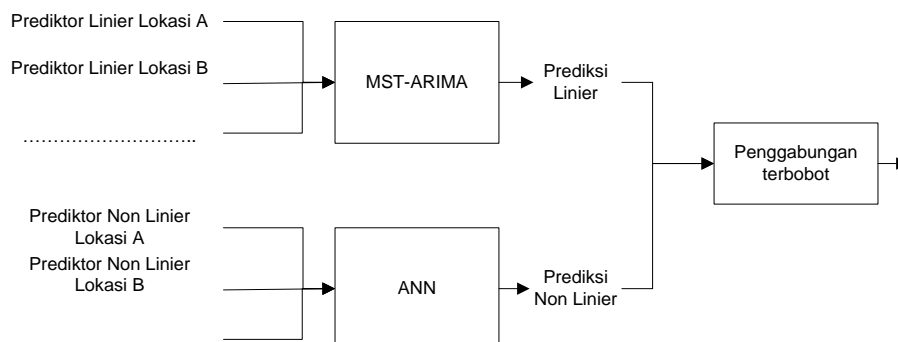
time-lags dari prediktor, terdiri atas ekstraksi fitur linier menggunakan DPCCA dan ekstraksi non linier menggunakan PCMI.

Pada tahap seleksi fitur relevan, kandidat fitur prediktor dipilih menggunakan korelasi pearson (PC) untuk relevan secara linier. Sedangkan untuk relevan secara non linier, digunakan *Symmetrical Uncertainty* (SU). Dari hasil keduanya, diseleksi kembali dengan menggunakan perbandingan terhadap nilai *therhold* α dan β untuk menghapus fitur yang tidak relevan sehingga diperoleh prediktor optimal. Gambar 3.2 menjelaskan proses seleksi fitur prediktor.



Gambar 3.2 Seleksi fitur relevan dan hapus fitur redundan

Sedangkan untuk pengembangan model penduga curah hujan dengan menggunakan metode MST-ARIMA dan ANN dijelaskan pada gambar 3.3. Input untuk metode MST-ARIMA adalah sejumlah variabel prediktor linier yang telah diidentifikasi pada tahap seleksi fitur dengan variabel respon adalah curah hujan pada bulan ke-t. Sedangkan input dari ANN adalah sejumlah variabel prediktor non linier yang telah diidentifikasi dengan menggunakan PCMI dan PCSUF. Hasil keduanya digabungkan untuk memperoleh prediksi akhir.



Gambar 3.3 hybrid MST-ARIMA dan ANN

BAB 4

BIAYA DAN JADWAL PENELITIAN

4.1 Biaya Penelitian

Tabel 4.1 Rekapitulasi Anggaran Penelitian

1. Honor Tim Peneliti					
No.	Honor	Honor / Jam (Rp.)	Waktu(j/mg)	Minggu	Biaya
1	Ketua Peneliti (10 Bulan)	0	40	40	-
Sub Total (Rp.)					-
2. Peralatan Penunjang					
No.	Material	Justifikasi Pemakaian	Kuantitas	Harga Sat	Biaya
1	Buku Referensi MTS Forecasting	Studi literatur	3	750000	2,250,000.00
2	Buku Referensi Matlab	Studi literatur	2	800000	1,600,000.00
3	Apple MacBook Pro MF839ID/A	Olah dan Simpan Data	1	15700000	15,700,000.00
Sub Total (Rp.)					19,550,000.00
3. Bahan Habis Pakai					
No.	Material	Justifikasi Pemakaian	Kuantitas	Harga Sat	Biaya
1	Kertas F4 70 gr	Adm & Dok Penelitian	5	40000	200,000.00
2	Kertas A4 70 gr	Adm & Dok Penelitian	5	40000	200,000.00
3	Tinta Printer Black (L220 EPSON)	Adm & Dok Penelitian	2	92000	184,000.00
4	Tinta Printer Color(L220 EPSON)	Adm & Dok Penelitian	3	96000	288,000.00
5	Sambungan Internet (1 Mbps)	Adm & Dok Penelitian	10	105000	1,050,000.00
6	Rem Hardisk 4 TB WD MyPasport	Penyimpan Data	1	1898000	1,898,000.00
Sub Total (Rp.)					3,820,000.00
4. Perjalanan					
No.	Kegiatan	Justifikasi Pemakaian	Kuantitas	Harga Sat	Biaya
1	Diseminasi : Conference	Pub. di Seminar Inter	1	5000000	5,000,000.00
2	Transportasi Conference	Transportasi Sby - Malaysia (pp)	2	3000000	6,000,000.00
3	Akomadasi untukConference	Penginapan (2 hari)	2	1500000	3,000,000.00
4	Survey dan diskusi NarSum BMKG	Transportasi (Sby-Jkt, 2x, pp)	4	1000000	4,000,000.00
5	Akomadasi selama survey	Inap (2 hari), transport dlm kota	2	850000	1,700,000.00
Sub Total (Rp.)					19,700,000.00
5. Lain-lain: Dok, Publikasi dan Laporan					
No.	Kegiatan	Justifikasi	Kuantitas	Harga Sat	Biaya
1	Penggandaan & Jilid Lap Interim	Akuntabilitas penelitian	8	250000	2,000,000.00
2	Penggandaan & Jilid Lap Interim	Akuntabilitas penelitian	8	250000	2,000,000.00
3	Cetak Foto Dokumentasi	Akuntabilitas penelitian	100	3000	300,000.00
4	Publikasi jurnal internasional	Akuntabilitas penelitian	1	9000000	9,000,000.00
5	Tenaga Administrasi (Laboran)	Kegiatan administrasi	10	250000	2,500,000.00
Sub Total (Rp.)					15,800,000.00
Total Anggaran yang Diperlukan (Rp.)					58,870,000.00

4.2 Jadwal Penelitian

No	Jenis Kegiatan	Bulan ke-									
		1	2	3	4	5	6	7	8	9	10
1	Studi literature										
2	Pengembangan <i>framework</i> untuk memangkitkan variabel <i>time-lags</i> Multivariate Spatio-Temporal (MST) dengan DPCCA dan PCMI										
3	Menyiapkan <i>framework</i> untuk seleksi fitur data MST dengan PCSUF										
4	Mengembangkan algoritma hybrid MST-ARIMAX dan ANN										
5	Mengumpulkan data MTS sintetik dan data meteorologi/klimatologi										
6	Mengimplementasikan algoritma <i>time-lags</i> , seleksi fitur dan hybrid										
7	Uji coba dan evaluasi hybrid										
8	Penulisan makalah untuk conference dan jurnal internasional										
9	Penulisan Laporan Penelitian (untuk laporan Interim dan Final)										

DAFTAR PUSTAKA

- Abbot, J. & Marohasy, J., 2014. Input selection and optimisation for monthly rainfall forecasting in queensland, australia, using artificial neural networks. *Atmospheric Research*, 138, pp.166–178. Available at: <http://dx.doi.org/10.1016/j.atmosres.2013.11.002>.
- Acharya, N. et al., 2014. Development of an artificial neural network based multi-model ensemble to estimate the northeast monsoon rainfall over south peninsular India: An application of extreme learning machine. *Climate Dynamics*, 43(5–6), pp.1303–1310.
- Asadi, S. et al., 2013. Neurocomputing A new hybrid arti fi cial neural networks for rainfall – runoff process modeling. *Neurocomputing*, 121, pp.470–480. Available at: <http://dx.doi.org/10.1016/j.neucom.2013.05.023>.
- Banihabib, M.E., Ahmadian, A. & Jamali, F.S., 2017. Hybrid DARIMA-NARX model for forecasting long-term daily inflow to Dez reservoir using the North Atlantic Oscillation (NAO) and rainfall data. *GeoResJ*, 13, pp.9–16. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S2214242816300584>.
- Chang, F.J. et al., 2016. Prediction of monthly regional groundwater levels through hybrid soft-computing techniques. *Journal of Hydrology*, 541, pp.965–976. Available at: <http://dx.doi.org/10.1016/j.jhydrol.2016.08.006>.
- Hamada, J.-I. et al., 2002. Spatial and Temporal Variations of the Rainy Season over Indonesia and their Link to ENSO. *Journal of the Meteorological Society of Japan*, 80(2), pp.285–310.
- Han, M., Ren, W. & Liu, X., 2015. Joint mutual information-based input variable selection for multivariate time series modeling. *Engineering Applications of Artificial Intelligence*, 37, pp.250–257. Available at: <http://dx.doi.org/10.1016/j.engappai.2014.08.011>.
- Hashim, R. et al., 2016. Selection of meteorological parameters affecting rainfall estimation using neuro-fuzzy computing methodology. *Atmospheric Research*, 171, pp.21–30.
- Hendon, H.H., 2003. Indonesian rainfall variability: Impacts of ENSO and local air-sea interaction. *Journal of Climate*, 16(11), pp.1775–1790.
- Kaur, G., 2012. Meteorological Data Mining Techniques: A Survey. *International Journal of Emerging Technology and Advanced Engineering*, 2(8), pp.325–327. Available at: http://www.ijetae.com/files/Volume2Issue8/IJETAE_0812_56.pdf.
- Kohail, S. & El-Halees, A., 2011. Implementation of Data Mining Techniques for Meteorological Data Analysis (A case study for Gaza Strip). *International Journal of Information*, 1(3), pp.96–100. Available at: [http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Implementation+of+Data+Mining+Techniques+for+Meteorological+Data+Analysis+\(+A+case+study+for+Gaza+Strip+\)+#0](http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Implementation+of+Data+Mining+Techniques+for+Meteorological+Data+Analysis+(+A+case+study+for+Gaza+Strip+)+#0).
- Lee, H., 2015. General Rainfall Patterns in Indonesia and the Potential Impacts of Local Seas on Rainfall Intensity. *Water*, 7(4), pp.1751–1769. Available at: <http://www.mdpi.com/2073->

4441/7/4/1751/.

- Munandar, D., 2015. Optimation Weather Parameters Influencing Rainfall Prediction Using Adaptive Network-Based Fuzzy Inference System (Anfis) and Linear Regression. , pp.154–159.
- Nhita, F., Annisa, S. & Kinasih, S., 2015. Comparative Study of Grammatical Evolution and Adaptive Neuro-Fuzzy Inference System on Rainfall Forecasting in Bandung. , pp.6–10.
- Nisak, S.C., 2016. Seemingly Unrelated Regression Approach for GSTARIMA Model to Forecast Rain Fall Data in Malang Southern Region Districts. *Cauchy*, 4(2), p.57. Available at: <http://ejournal.uin-malang.ac.id/index.php/Math/article/view/3488>.
- Nugroho, A. & Simanjuntak, B.H., 2014. ARMA (Autoregressive Moving Average) Model for Prediction of Rainfall in Regency of Semarang - Central Java - Republic of Indonesia. *International Journal of Computer Science Issues (IJCSI)*, 11(3), pp.27–32. Available at: ijcsi.org/papers/IJCSI-11-3-1-27-32.pdf.
- Nur'utami, M.N. & Hidayat, R., 2016. Influences of IOD and ENSO to Indonesian Rainfall Variability: Role of Atmosphere-ocean Interaction in the Indo-pacific Sector. *Procedia Environmental Sciences*, 33, pp.196–203.
- Ortiz-garcía, E.G., Salcedo-sanz, S. & Casanova-mateo, C., 2014. Accurate precipitation prediction with support vector classifiers : A study including novel predictive variables and observational data. *Atmospheric Research*, 139, pp.128–136. Available at: <http://dx.doi.org/10.1016/j.atmosres.2014.01.012>.
- Ramana, R.V., Krishna, B. & Kumar, S.R., 2013. Monthly Rainfall Prediction Using Wavelet Neural Network Analysis. , pp.3697–3711.
- Sánchez-Monedero, J. et al., 2014. Simultaneous modelling of rainfall occurrence and amount using a hierarchical nominal-ordinal support vector classifier. *Engineering Applications of Artificial Intelligence*, 34, pp.199–207. Available at: <http://dx.doi.org/10.1016/j.engappai.2014.05.016>.
- Setyaningrum, A.H. & Swarinata, P.M., 2014. Weather prediction application based on ANFIS (Adaptive neural fuzzy inference system) method in West Jakarta region. *2014 International Conference on Cyber and IT Service Management, CITSM 2014*, pp.113–118.
- Sharma, A. & Goyal, M.K., 2015. Bayesian Network Model for Monthly Rainfall Forecast. *2015 IEEE International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), Kolkata, India*, pp.241–246. Available at: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7434243&isnumber=7434193>.
- Supari, 2012. SPATIOTEMPORAL CHARACTERISTICS OF EXTREME RAINFALL EVENTS OVER JAVA ISLAND , Case: East Java Province.
- Swarinoto, Y.S. et al., 2012. Model Sistem Prediksi Ensemble Total Hujan Bulanan Dengan Nilai Pembobot (Kasus Wilayah Kabupaten Indramayu). *Jurnal Mkg*, 13(3), pp.189–200.
- Toggweiler, J. & Key, R., 2001. Ocean circulation: Thermohaline circulation. *Encyclopedia of Atmospheric Sciences*, 4(December 2007), p.1549–1555. Available at:

http://cdiac.esd.ornl.gov/oceans/GLODAP/glodap_pdfs/Thermohaline.web.pdf.

Vasimalla, K., 2014. A Survey on Time Series Data Mining. *International Journal of Innovative Research in Computer and Communication Engineering*, 2(5), pp.170–179.

Yu, L. & Liu, H., 2003. Feature Selection for High-Dimensional Data: A Fast Correlation-Based Filter Solution. *International Conference on Machine Learning (ICML)*, pp.1–8. Available at: <http://www.aaai.org/Papers/ICML/2003/ICML03-111.pdf>.

Lampiran 1. Dukungan Sarana dan Prasarana Penelitian

Penelitian ini direncanakan akan dilakukan di Laboratorium Dasar dan Terapan Komputasi(DTK) dan Laboratorium Residensi Program Studi S3 Ilmu Komputer, Jurusan Teknik Informatika, Fakultas Teknologi Informasi. Laboratorium DTK memiliki sarana hardware dan software sebagai berikut:

Nama Perangkat	Jumlah	Keterangan/Spesifikasi
Komputer client	16	2 server dan 20 PC workstation
Komputer server	2	Windows server dan Windows XP
Scanner	1	Simul8, ARENA, Visual Studio.Net, Java Netbean, Matlab, Xpert Rule, My SQL, SQL Server 2005 dan Oracle 10g DBMS
Printer	1	-
Keyboard	16	-
Mouse	16	-
Speaker	2	-
Switch HUB	3	-
Router	1	-
Sistem Operasi	3	Windows 7, Solaris, Linux
Tools/Software		Matlab R2015a, R, SPSS 17, Visual Studio.Net, SQL Server



Lampiran 2. Biodata Ketua

Ketua Peneliti

A. Identitas Diri

1	Nama Lengkap	Ahmad Saikhu, S.Si, MT (L/P)
2	Jabatan Fungsional	Lektor Kepala
3	Jabatan Struktural	Kepala Laboratorium Dasar dan Terapan Komputasi
4	NIP	197107182006041001
5	NIDN	0018077105
6	Tempat dan Tanggal Lahir	Gresik, 18 Juli 1971
7	Alamat Rumah	Griya Kartika A-27 Sedati Sidoarjo
8	Nomor Telp/Fax/HP	+62081357976106
9	Alamat Kantor	Jurusan Teknik Informatika, ITS Kampus ITS Sukolilo Surabaya
10	Nomor Telp/Fax	+62 31 5939214
11	Alamat Email	saikhu@its-sby.edu
12	Lulusan yang Telah Dihasilkan	S1 = 72 orang
13	Mata Kuliah yang Diampu	1. Simulasi Sistem Diskrit
		2. Riset Operasional
		3. Kecerdasan Buatan
		4. Analisa Data Multivariat

B. Riwayat Pendidikan

	S-1	S-2	S-3
Nama Perguruan Tinggi	ITS	ITS	ITS

Bidang Ilmu	Statistika	Teknik Informatika	Ilmu Komputer
Tahun Masuk-Lulus	1989-1994	1997-2000	2015-
Judul Skripsi/Tesis/Disertasi	Analisis Log Linier untuk evaluasi pola pilihan layanan kesehatan di Kota Surabaya	Studi perbandingan kinerja model syaraf tiruan dan ARIMA untuk peramalan	
Nama Pembimbing/Promotor	Prof. Dr. Susanti Linuwih, M.Stat	Prof. Dr. Drs. Ir. Rianarto Sarno, M.Sc.	Dr. Agus Zainal Arifin, S.Kom, M.Kom

C. Pengalaman Penelitian dalam 5 Tahun Terakhir

No.	Tahun	Judul Penelitian	Pendanaan	
			Sumber	Jml (Juta Rp)
1	2008	Development of Content-Based Image Retrieval System using Object Recognition	JICA-PREDICT	200
2	2010	Pengembangan Sistem penelusuran Produk pada PT. Aneka Tuna Indonesia	Kemristek	265
3	2011	Pemodelan Event Log menggunakan Regresi Logistik	DIPA-ITS	35
4	2011	Analisis hasil pembangkitan bilangan acak menggunakan Java, VB, dan C++	DIPA-ITS	35
5	2013	Pemodelan Aplikasi Monitoring Arus Lalu Lintas dengan SIG-T Berbasis Web di Kota Kupang	EPI-UNET BOPTN	60

6	2013	Aplikasi Manajemen Penyakit Diabetes Menggunakan Konsep Software as a Service	IPTEKS-DIKTI	100
7	2014	Sistem Manajemen Arus Lalu Lintas Cerdas, Studi Kasus: di Kota Kupang, NTT	EPI-UNET BOPTN	100
8	2014-2015	Pengembangan Model Optimasi Untuk Penjadualan Dokter Pada Keadaan Darurat Berbasis Aplikasi Mobile	BOPTN	60
9	2015	Traffic Management System by Computer Vision, Computer Simulation and Statistical Data Analysis	KLN	100

D. Pengalaman Pengabdian kepada Masyarakat dalam 5 Tahun Terakhir

No.	Tahun	Judul Pengabdian kepada Masyarakat	Pendanaan	
			Sumber	Jml (Juta Rp)
1	2011	Technical Assistant pada Audit Rintisan Sekolah Bertaraf Internasional tingkat SMA di Surabaya. Didanai Dinas Pendidikan, tahun 2012 Sebagai Anggota (Narasumber).	Dinas Pendidikan	10
2	2012	Pelatihan Pemasaran Online bagi Anggota Koperasi Wanita “Setia Bhakti Surabaya”. Didanai Jurusan Teknik Informatika ITS,	Teknik Informatika	2

		Juli 2012. Sebagai Anggota (Narasumber).		
3	2012	Counselor ICTPura untuk Kabupaten Sidoarjo dan Kota Mojokerto	Depkominfo	6
4	2013	Pemanfaatan IT untuk Optimasi Produksi dan Pemasaran pada UKM Kerajinan Sulaman di Surabaya dan sekitarnya	BOPTN	20
5	2013	Implementasi Program Laporan Akuntansi Dasar dan Pelatihan Excel, Powerpoint serta Internet untuk Unit Pengelola Keuangan Badan Keswadayaan Masyarakat (BKM) Ngagel Rejo Mulyo	BOPTN	20
6	2015	IbM guru-guru dan pengelola PAUD	BOPTN	19

E. Pengalaman Penulisan Artikel Ilmiah dalam Jurnal dalam 5 Tahun Terakhir

No	Judul Artikel Ilmiah	Volume/ Nomor/Tahun	Nama Jurnal
1	Rancang Bangun Optimasi Kebutuhan Bahan Baku menggunakan Algoritma Wagner-Within	Volume 7, Nomor 4, ISSN 1412-6389	JUTI, Jurnal Ilmiah Teknologi Informasi
2	Penyelesaian Masalah Penjadualan Flowshop Fleksibel 3 Tahap untuk N-Job dengan Algoritma Outer dan Inner Game	Volume 4 No. 2, ISSN 0216-0544	Jurnal Kursor

3	Kombinasi Metode Sintesis Tekstur dan Metode Image Painting untuk Restorasi Citra	Volume V, Nomor 1, Februari 2010	Jurnal SCAN
4	Implementasi Pengembangan Metode Differential Evolution untuk Clustering Pixel	Volume 9, Nomor 2, Juli 2011	JUTI, Jurnal Ilmiah Teknologi Informasi
5	Pendekatan Algoritma Heuristik dan Neural Network untuk Screening Test pada Urinalysis	Vol. 1 No. 2 Des 2013	Jurnal Cybernatika ITB

F. Pengalaman Penyampaian Makalah Secara Oral pada Pertemuan/Seminar Ilmiah dalam 5 Tahun Terakhir

No	Nama Pertemuan Ilmiah/Seminar	Judul Artikel Ilmiah	Waktu dan Tempat
1	SNATI 2009, Yogyakarta	Simulasi Evakuasi Keadaan Darurat: Studi Kasus Apartemen XYZ, Surabaya. Seminar Nasional Aplikasi Teknologi Informasi	20 Juni 2009, UII
2	SNATIKA 2011, Malang	Uji Bilangan Acak dari Fungsi Pembangkit Bilangan Acak pada Bahasa Pemrograman JAVA. Seminar Nasional Teknologi, Informasi, Komunikasi dan Aplikasinya	11 Nopember 2011, STIKI

3	KNS&I	Klasifikasi <i>Event</i> Pada <i>Process Logs</i> Menggunakan Model Regresi Logistik	Pebruari 2012, STIKOM Bali
4	SENTIKA, Yogyakarta	Klasterisasi Data Diskrit Menggunakan Algoritma Multinomial Generalized Dirichlet Mixture Pada Citra	Juni 2012, Atmajaya
5	SEMANTIK	Penentuan Jenis Fumigasi Dengan Menggunakan Metode Decision Tree	23 Juni 2012, Semarang
6	SNATIA	Pendekatan Algoritma Heuristik dan Neural Network untuk Screening Test pada Urinalysis	21 September 2013, Universitas Surabaya
7	Seminar Nasional Sains & Teknologi V, Lampung	Perangkat Lunak untuk deteksi Jumlah Kendaraan di Jalan dengan Transceiver SRF02	19-20 November 2013, Universitas Lampung
8	DISC	Optimasi Permasalahan Penugasan Dokter Menggunakan Representasi Graf Bipartit Berbobot	3-4 Oktober 2014, Universitas Maranatha
9	ITECHS 2014	Implementasi metode integer programming untuk penjadualan tenaga medis pada situasi darurat berbasis aplikasi mobile	Desember 2014, STIKI Malang
10	ICTS 2015		September 2015, ITS Surabaya

G. Pengalaman Penulisan Buku dalam 5 Tahun Terakhir

No	Judul Buku	Tahun	Jumlah halaman	Penerbit
-	-	-	-	-

H. Pengalaman Perolehan HKI dalam 5 - 10 Tahun Terakhir

No	Judul/Tema HKI	Tahun	Jenis	Nomor P/ID
1	Metode Penemuan Kembali Citra Berbasis Pengenalan Objek Dan Estimasi <i>Kansei</i>	2008	Paten	P00200800224

I. Pengalaman Merumuskan Kebijakan Publik/Rekayasa Sosial Lainnya dalam 5 Tahun Terakhir

No	Judul/Tema/Jenis Rekayasa Sosial Lainnya yang Telah Diterapkan	Tahun	Tempat Penerapan	Respons Masyarakat
1	-	-	-	-

J. Penghargaan yang Pernah Diraih dalam 10 Tahun Terakhir (dari pemerintah, asosiasi atau institusi lainnya)

No	Jenis Penghargaan	Institusi Pemberi Penghargaan	Tahun
1	Dosen berprestasi III	Fakultas	2008
2	Penyaji Makalah Terbaik pada Seminar Hasil Penelitian Dosen Muda, Regional VII Jawa Timur	LPPM ITS	2008

Semua data yang saya isikan dan tercantum dalam biodata ini adalah benar dan dapat dipertanggungjawabkan secara hukum. Apabila di kemudian hari ternyata dijumpai ketidaksesuaian dengan kenyataan, saya sanggup menerima risikonya.

Demikian biodata ini saya buat dengan sebenarnya untuk memenuhi salah satu persyaratan dalam pengajuan Penelitian Program Pasca Sarjana Dana Lokal ITS.

Surabaya, 6 Maret 2017

Pengusul,

A handwritten signature in black ink, appearing to be 'Ahmad Saikhu', enclosed within a large, loopy oval shape.

Ahmad Saikhu, S.Si, MT.
NIP:197107182006041001

Lampiran 5. Surat Pernyataan Ketua Pengusul

SURAT PERNYATAAN KETUA PENGUSUL

Yang bertanda tangan di bawah ini:

Nama : Ahmad Saikhu, S.Si, MT.

NIDN : 0018077105

Pangkat / Golongan : Pembina / IVA

Jabatan Fungsional : Lektor Kepala

Dengan ini menyatakan bahwa proposal saya dengan judul:

Pengembangan Model Penduga Curah Hujan menggunakan hybrid Spatio-Temporal Multivariat ARIMA dan Jaringan Syaraf Tiruan

yang diusulkan dalam skema Penelitian Disertasi Doktor (PDD) untuk tahun anggaran 2018 bersifat original dan belum pernah dibiayai oleh lembaga/sumber dana lain.

Bilamana di kemudian hari ditemukan ketidak sesuaian dengan pernyataan ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku dan mengembalikan seluruh biaya penelitian yang sudah diterima ke kas negara.

Demikian pernyataan ini dibuat dengan sesungguhnya dan dengan sebenar-benarnya.



Mengetahui
Ketua LPPM

Prof. Dr. Ir. Adi Soeprijanto, MT
NIP. 196404051990021001

Surabaya, 7 Juni 2017
Yang menyatakan,



Ahmad Saikhu, S.Si, MT.
NIP. 197107182006041001