

PROPOSAL SEMINAR USULAN PENELITIAN

**PREDIKSI POLA PENYEBARAN PENYAKIT MELALUI PEMODELAN
KASUS SECARA SPATIAL TEMPORAL**

***“PREDICTION OF DISEASE SPREAD PATTERN THROUGH
TEMPORAL SPATIAL MODELING”***

**MUHAMMAD RESHA
P1600316007**



PROGRAM PASCASARJANA
PROGRAM STUDI DOKTOR
DEPARTEMEN TEKNIK ELEKTRO
UNIVERSITAS HASANUDDIN
MAKASSAR
2017

Pengesahan Proposal Seminar Usulan Penelitian

**PREDIKSI POLA PENYEBARAN PENYAKIT MELALUI PEMODELAN
KASUS SECARA SPATIAL TEMPORAL**

Diajukan Oleh :

**Muhammad Resha
NIM. P1600316007**

*Telah diperiksa dan dinyatakan memenuhi syarat untuk melaksanakan
seminar usul penelitian*

Menyetujui

Tim Promotor

**Prof. Dr. Ir. Ansar Suyuti, MT., IPU.
Promotor
Tanggal :.....**

**Dr. Ir. H. Rhiza S. Sadjad, MSEE
Kopromotor**

Tanggal :.....

**Dr. Eng. Muhammad Niswar ST.,MIT
Kopromotor**

Tanggal :.....

**Ketua Program Studi
S3 Teknik Elektro**

Prof. Dr. Ir. Ansar Suyuti, MT.

ABSTRAK

MUHAMMAD RESHA, Prediksi Pola Penyebaran Penyakit Melalui Pemodelan Kasus Secara Spatial Temporal (dibimbing oleh Ansar Suyuti, Rhiza S Sadjad, Muhammad Niswar)

Penyakit menular merupakan permasalahan kesehatan yang penting di hampir semua negara, termasuk Indonesia. Beberapa penyakit menular dapat menyebar di dalam populasi hingga menyebabkan epidemi. Karakteristik geografi, keadaan lingkungan beberapa kondisi yang menjadi tantangan untuk mendapatkan model yang valid dan reliabel dan digunakan untuk memetakan prediksi penyakit menular di setiap wilayah. Perbedaan yang terjadi harus dapat diakomodasi dalam pemodelan angka resiko penyakit menular. Salah satu upaya untuk mendapatkan model yang valid dan reliabel adalah membangun model pada kelompok data yang memiliki tingkat homogenitas yang tinggi.

Hal ini dapat diatasi dengan menggabungkan beberapa metode clustering dengan pemodelan pemetaan penyakit. Model spatial-temporal adalah teknik analisis pengelompokan yang memperhatikan distribusi probabilitas dari data yang memungkinkan melakukan pengelompokan data dengan distribusi arah dan bentuk yang berbeda.. Melalui cara ini dapat diprediksi angka resiko penyakit untuk setiap lokasi.

Keyword : Population-Based Incremental Learning, Spatial Temporal

DAFTAR ISI

HALAMAN JUDUL
HALAMAN JUDUL
HALAMAN PENGESAHAN
DAFTAR ISI
DAFTAR TABEL
DAFTAR GAMBAR

BAB 1 PENDAHULUAN

- 1.1. Latar Belakang 1
- 1.2. Rumusan masalah 7
- 1.3. Tujuan Penelitian 8
- 1.4. Manfaat Penelitian 8
- 1.5. Penelitian Terkait dan Kontribusi Hasil Penelitian 9

BAB 2 TINJAUAN PUSTAKA

- 2.1. Pengertian Penyakit Tuberculosis 18
- 2.2. Pola Penyebaran Penyakit Tuberculosis 19
- 2.3. K-Means 22
- 2.4. Analisa Data Spatio-Temporal 26

BAB 3 METODE PENELITIAN

- 3.1. Diagram Alir Penelitian 40
- 3.2. Kerangka Berpikir Penelitian 41
- 3.3. Teknik Pengumpulan Data 41
- 3.4. Jenis dan Sumber Data 44
- 3.5. Waktu dan Lokasi Penelitian 44
- 3.6. Bahan dan Alat yang digunakan 45

DAFTAR PUSTAKA

DAFTAR GAMBAR

Gambar 1. Perubahan Beban Penyakit di Indonesia

Gambar 2. Status Beberapa Penyakit Menular

Gambar 3. Data Pasien Tuberculosis Makassar 01

Gambar 4. Langkah 1: Kumpulkan dan Siapkan Data

Gambar 5. Peta A dan Peta B

Gambar 6. Diagram Alir Penelitian

Gambar 7. Peta Makassar

Daftar Tabel

Tabel 1. Indikator Pengendalian Penyakit Menular

Tabel 2. Penelitian terkait , state of the art

Tabel 3. Kerangka berpikir Penelitian

Tabel 4. Daftar Puskesmas di Makassar

BAB 1

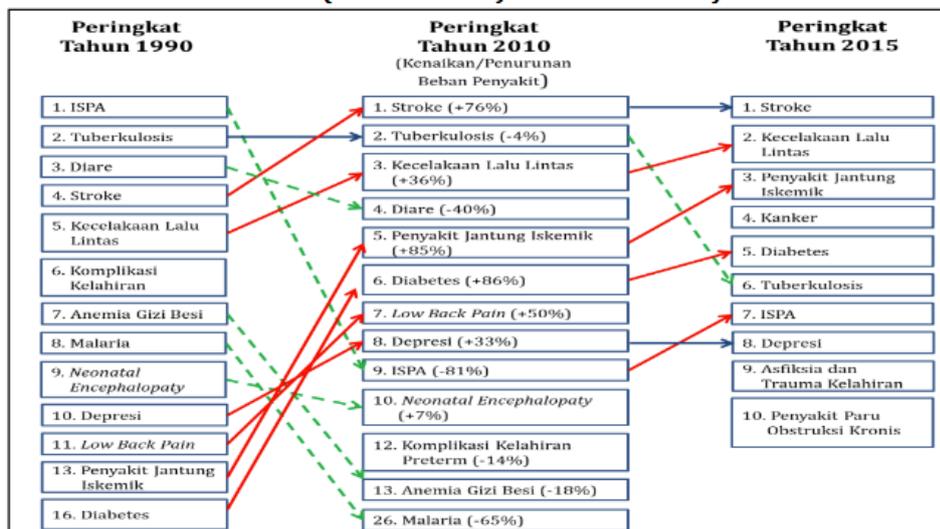
PENDAHULUAN

1.1. Latar Belakang

Kualitas Kesehatan di Indonesia masih menghadapi masalah yang serius, terutama penyakit menular masih menjadi masalah utama kesehatan di Indonesia. Kementerian kesehatan pada Rencana Program Jangka Menengah (RPJM) 2015-2019 telah menjelaskan penyakit menular yang menjadi prioritas adalah Demam Berdarah Dengue (DBD), Diare, Malaria, Tuberculosis, HIV dan AIDS menjadi prioritas penanggulangan utama sampai tahun 2019.

Pada gambar di bawah terlihat perubahan beban penyakit antara tahun 1990, 2010 dan 2015 dimana pada tahun 2015 terdapat penyakit menular yang menjadi masalah utama.

GAMBAR 2.5
PERUBAHAN BEBAN PENYAKIT ANTARA TAHUN 1990, 2010 DAN 2015
DI INDONESIA. BEBAN DIHITUNG SEBAGAI DALYS
(DISABILITY-ADJUSTED LIFE YEARS).



Sumber: *Global Burden of Disease, 2010* dan *Health Sector Review (2014)*

Gambar 1. Perubahan Beban Penyakit di Indonesia

Indonesia merupakan negara yang masuk dalam High Burden Countries, dimana sekarang ini berada pada rangking kelima negara dengan beban TB tertinggi di dunia. Lima negara dengan jumlah terbesar kasus insiden tersebut adalah India, Cina, Afrika Selatan, Nigeria dan Indonesia. Estimasi prevalensi TB semua kasus adalah sebesar 660.000 dan estimasi insiden berjumlah 430.000 kasus baru per tahun. Jumlah kematian akibat TB diperkirakan 61.000 kematian per tahunnya. (Kemenkes RI, 2011). Jika di lihat dari statusnya ada 5 Penyakit menular yang menjadi sasaran utama untuk dapat diturunkan kasusnya.

TABEL 2.2
STATUS BEBERAPA PENYAKIT MENULAR DAN TIDAK MENULAR TAHUN 2013

Jenis Penyakit	Indikator	Status
Penyakit Menular		
Demam Berdarah Dengue (DBD)	<i>Case Fatality Rate (CFR) DBD</i>	0,77 % ¹⁾
Diare	<i>Period Prevalence diare semua kelompok umur</i>	7 % ²⁾
Malaria	<i>Annual Parasite Index (API)</i>	1,38 per 1.000 penduduk ¹⁾
Tuberculosis	Prevalensi TB paru yang didiagnosis oleh nakes	0,4 % ²⁾
HIV dan AIDS	Jumlah kasus baru HIV positif	29.037 ¹⁾

Sumber: 1) *Profil Kesehatan 2013*; 2) *Riskesdas 2013*

Gambar 2. Status Beberapa Penyakit Menular

Kualitas Kesehatan harus mendapat perhatian khusus terutama mencegah penyakit menular lebih dini. Pada Rencana Strategis Kementerian Kesehatan 2015-2019 memiliki rencana Pengendalian Penyakit Menular Langsung. Sasaran kegiatan ini adalah menurunnya angka kesakitan dan kematian akibat penyakit menular langsung.

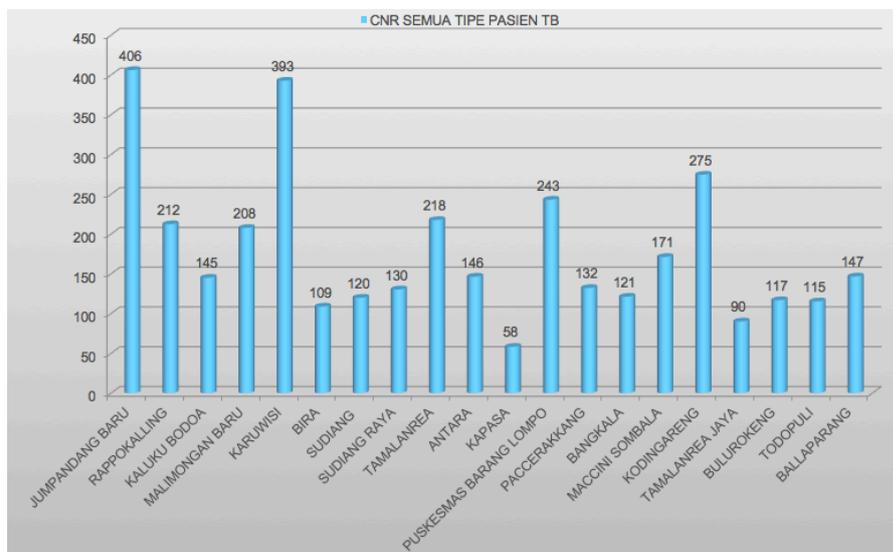
Sasaran yang ingin dicapai adalah meningkatkan derajat kesehatan dan status gizi masyarakat melalui upaya kesehatan dan pemberdayaan masyarakat yang didukung dengan pemeratan

pelayanan kesehatan. Sasaran pembangunan kesehatan pada RPJMN 2015-2019 sebagai berikut:

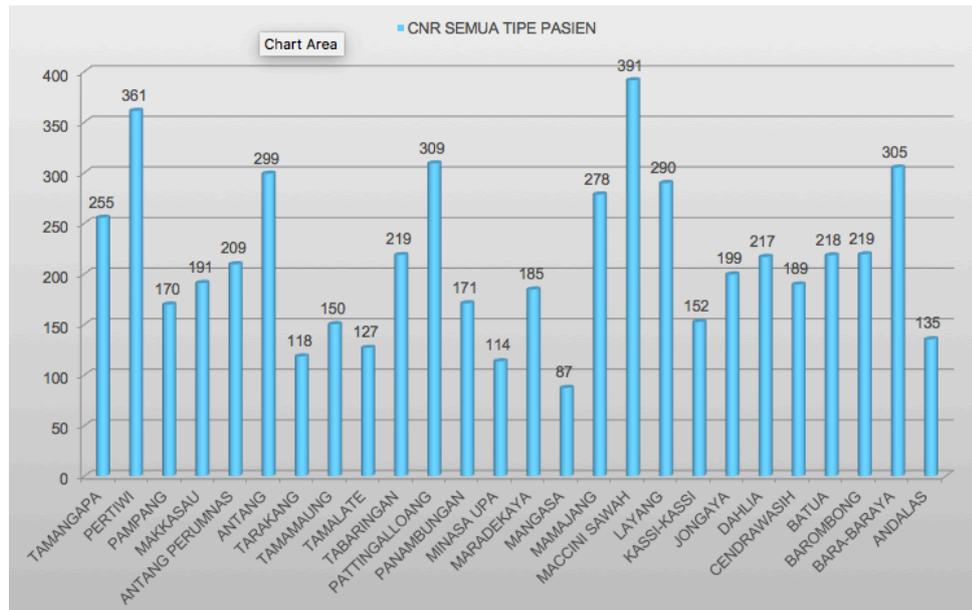
No	Indikator	Status Awal	Target 2019
2	Meningkatnya Pengendalian Penyakit Menular dan Tidak Menular		
a.	Prevalensi Tuberkulosis (TB) per 100.000 penduduk	297 (2013)	245
b.	Prevalensi HIV (persen)	0,46 (2014)	<0,50
c.	Jumlah kabupaten/kota mencapai eliminasi malaria	212 (2013)	300
d.	Prevalensi tekanan darah tinggi (persen)	25,8 (2013)	23,4
e.	Prevalensi obesitas pada penduduk usia 18+ tahun (persen)	15,4 (2013)	15,4
f.	Prevalensi merokok penduduk usia ≤ 18 tahun	7,2 (2013)	5,4

Tabel 1. Indikator Pengendalian Penyakit Menular

Data Pasien Tuberculosis Makassar dari Ekspose bidang P2PL Dinas Kesehatan Kota Makassar tahun 2016 di setiap Puskesmas.



Gambar 3. Data Pasien Tuberculosis Makassar 01



Gambar 4. Data Pasien Tuberculosis Makassar 02

Penelitian terkait Spatio Temporal Tuberculosis telah dilakukan (Daiane Leite da Roza, et al 2012). Mengevaluasi sebaran penyakit Tuberculosis di Sao Paulo, Brazil menggunakan pemodelan *Bayesian Regression* dengan mengasumsikan suatu *Poisson Distribution* untuk kasus baru TB.

Pada penelitian (Li Huang et al 2017) di Timur Laut Yunnan - China, pemetaan sebaran penyakit Tuberculosis menggunakan Time Series dan Analisa Spatial Cluster untuk mendeteksi pola spasial dan trend temporalnya. Dideteksi penderita tertinggi berada pada musim dingin dan musim semi pada satu cluster yang paling memungkinkan, 6 cluster diidentifikasi untuk registrasi pasien TB.

dari 6 cluster, 5 diantaranya registrasi baru pasien dan yang lain registrasi pasien berulang.

Defenisi Spatio-Temporal merupakan data dikumpulkan pada suatu waktu dan juga daerah tertentu dimana data dikumpulkan secara berkala pada setiap hari atau setiap minggu, sehingga analisis data harus memperhitungkan ketergantungan spasial di antara para pemantau, namun juga pengamatan pada masing-masing monitor biasanya tidak independen tapi format timeseries.

Dalam penelitian Spatio-Temporal Penyakit Tuberkulosis, (Kai Cao, et al 2016) menambahkan analisa ekologi yang mempengaruhi faktor penyebaran Tuberculosis disamping menggunakan Bayesian Framework. Pada Bayesian Framework digunakan 6 model statistik dengan parameter waktu, ruang, interaksi spatio-temporal dan kombinasinya.

Hasil analisa dengan menggunakan software SatScan dilakukan (Ulziijargal Gurjav et al 2015) Dengan parameter waktu retrospektif parameter analisis distribusi Poisson. Pemetaannya menggunakan Quantum GIS pada area hotspot. Area ini diidentifikasi melalui membandingkan kasus yang diamati di lokasi spatio-temporal dengan distribusi kasus Poisson yang diharapkan.

Penelitian Spatio-Temporal di Fukuoka - Jepang, (Daisuke Onozuka et. Al. 2007) identifikasi cluster menggunakan space-time scan statistik dimana data di analisa dengan software SatScan yang digunakan untuk menguji kelompok lokasi geografis dan untuk mengidentifikasi perkiraan lokasinya. Jumlah kejadian kasus, dapat diasumsikan memiliki distribusi Poisson atau Bernoulli. Data statistik digabungkan, yaitu area sensus dengan koordinat geografis yang tepat, di mana setiap area sensus hanya mengandung satu orang yang berisiko.

Penelitian Spatio-Temporal pada sebaran penyakit tuberculosis menggunakan Bayesian Trans-Gaussian and Log Empirical Bayesian Kriging di lakukan (Kasturi Basu, et al . 2016) di India dengan memprediksi risiko penyebaran penyakit secara terpisah pada setiap kuartal dalam 4 tahun. Kelompok pasien Positif diidentifikasi pada setiap kuartal dan perubahan pola temporal spatio di seluruh area diamati selama periode 4 tahun.

Dari berbagai analisa Spatio-Temporal di atas didapatkan beberapa teknik algoritma yang di gunakan yaitu :

- Bayesian Regression
- Distribution Poisson
- Bayesian Trans-Gaussian and Log Empirical Bayesian

Berdasarkan analisa di atas, maka penulis akan meneliti algoritma yang paling efisien untuk digunakan dalam menganalisa Spatio-Temporal penyebaran penyakit menular dengan cara memodelkan algoritma yang ada, sehingga dapat diketahui pola penyebaran penyakit Tuberkulosis di Makassar.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan di atas maka rumusan masalah dalam penelitian terkait Spatio-Temporal Penyakit Tuberculosis adalah sebagai berikut :

- Bagaimana menyeleksi data dasar pasien penyakit Tuberculosis dengan menggunakan Bayesian Regression dan Distributiron Poisson sehingga ditemukan posisi yang lebih akurat.
- Bagaimana memetakan data yang sudah terseleksi pada sebaran penyakit Tuberculosis melalui pemodelan Unsupervised Classification.
- Bagaimana merumuskan prediksi sebaran penyakit Tuberculosis terkait Spatio-Temporal dan menemukan polanya untuk mendapatkan pemetaan yang paling efisien.
- Bagaimana memodelkan Pola Sebaran Penyakit Tuberculosis pada peta yang mewakili Waktu dan Lokasi dalam kurun waktu yang ditentukan.

1.3. Tujuan Penelitian

Sesuai dengan latar belakang permasalahan dan rumusan masalah di atas, maka tujuan yang akan di lakukan dalam penelitian ini adalah sebagai berikut :

- Untuk menemukan posisi yang lebih akurat dengan menyeleksi data dasar pasien penyakit Tuberculosis dengan menggunakan bayesian regression dan Distribution Poisson.
- Untuk memetakan data sebaran penyakit Tuberculosis melalui pemodelan Unsupervised Classification.
- Untuk merumuskan prediksi sebaran penyakit Tuberculosis terkait Spatio-Temporal dan menemukan polanya.
- Untuk memodelkan pola sebaran penyakit Tuberculosis pada peta yang mewakili waktu dan lokasi dan kurun waktu yang ditentukan.

1.4. Manfaat Penelitian

Bedasarkan latar belakang permasalahan, rumusan masalah dan tujuan penelitian tersebut diatas maka diharapkan penelitian ini bermanfaat yang sebesar-besarnya pada dunia kesehatan masyarakat dan khususnya kepada para peneliti informatika kesehatan sebagai berikut :

- Untuk membantu mencegah penyebaran penyakit Tuberculosis lebih dini dengan memprediksi pola dari data dasar pasien.

- Diharapkan dapat meningkatkan keakuratan prediksi sebaran penyakit Tuberculosis.
- Berusaha menemukan pemodelan Spatio-Temporal yang lebih efisien.

1.5. Penelitian Terkait dan Kontribusi Hasil Penelitian

Pada penelitian ini diharapkan mampu memberikan solusi yang lebih efisien dan memberikan kontribusi dalam dunia kesehatan masyarakat, mengingat pentingnya menurunkan angka penderita penyakit Tuberculosis yang selalu berada pada 5 besar angka penderita terbanyak penyakit menular.

Kontribusi yang sama telah dilakukan oleh peneliti lainnya adalah sebagai berikut :

Peneliti (Daiane Leite da Roza et. al. 2012) mengevaluasi data dari TBWeb yaitu sebuah sistem notifikasi elektronik untuk kasus TB. Ukuran kerentanan sosial diperoleh dari SEADE Foundation, dan informasi tentang jumlah penduduk, pendidikan dan pendapatan rumah tangga diperoleh dari Institut Geografi dan Statistik Brasil. Analisis statistik dilakukan dengan model regresi Bayesian dengan asumsi distribusi Poisson untuk kasus TB baru yang diamati di setiap area. Struktur autoregresif bersyarat digunakan untuk struktur kovariansi spasial.

Peneliti (Li Huang et. al. 2017) mengambil data tentang kasus PTB dari sistem pendaftaran online tuberkulosis nasional. Menerapkan Analisa *Time Series* dan analisa *Spatial Cluster* untuk mendeteksi tren temporal dan pola spasial Penyakit Tuberculosis di area antara tahun 2011 dan 2015 di Zhaotong. Tiga indikator riwayat pendaftaran pengobatan PTB digunakan: tingkat pendaftaran pengobatan awal, tingkat registrasi ulang, dan Jumlah pasien pendaftar penyakit Tuberculosis.

Peneliti (Kai Cao et. al. 2016), mengeksplorasi efek interaksi spasial-temporal dalam *Bayesian Framework* dan untuk menyelidiki faktor-faktor yang mempengaruhi ekologis untuk tuberkulosis. Menggunakan Enam model statistik yang berbeda yang berisi parameter waktu, ruang, interaksi spasial-temporal dan gabungannya dibangun berdasarkan *Bayesian Framework*.

Peneliti (Ulziijargal Gurjav et. el. 2015) menganalisa data geografi, epidemiologi dan genotipe dari semua kasus tuberkulosis yang dikonfirmasi untuk mengidentifikasi faktor penentu bakteri dan demografis daerah hotspot tuberkulosis. Risiko Relatif (RR) untuk masing-masing area hotspot dihitung dengan Binomial Logistic Regression dengan menggunakan area non-hotspot sebagai referensi.

Peneliti (Daisuke Onozuka et. el. 2007) data dikoordinatkan dengan posisi *longitude* dan *lattice*. Menggunakan 109 kumpulan

data TB geografis yang berbeda untuk statistik pemindaian spasial. Statistik pemindaian spasial dikembangkan untuk menguji kelompok geografis dan untuk mengidentifikasi perkiraan lokasi mereka. Kami menggunakan statistik *space-time scan* untuk memeriksa kelompok TB.

Peneliti (Kasturi Basu et. al. 2016) menggunakan Metode Empirical Bayesian geostatistical pada Software ArcGIS 10.2 yang digunakan untuk mengurangi variasi data dan untuk menghindari interpolasi negatif, setelah transformasi logaritmik dari data yang diamati dan diterapkan untuk menganalisis dan memprediksi prevalensi pasien TB Bermasalah Positif berisiko baru.

Peneliti (Shiyi Cao et. al. 2013) menggunakan data kasus kejadian TB bulanan mulai Januari 2005 sampai Desember 2011 yang diperoleh dari Kementerian Kesehatan, China. Menggunakan pemodelan *Seasonal Autoregressive Integrated Moving Average* (SARIMA) dan model hybrid yang menggabungkan model SARIMA dan *Neural Network Regression* yang digunakan untuk menyesuaikan data dari tahun 2005 sampai 2010. Parameter kinerja simulasi *Mean Square Error* (MSE), *Mean Absolute Error* (MAE) dan *Mean Absolute Percentage Error* (MAPE) digunakan untuk membandingkan antara kedua model ini. Sumber data digunakan untuk memvalidasi model yang dipilih.

Peneliti (Rindra V Rendremanana et. al. 2013) menggunakan metode pendekatan *Bayesian* dan *Generalized Linear Mixed Model* (GLMM) yang dikembangkan untuk menghasilkan peta Tuberculosis yang beresiko dan untuk memodelkan hubungan antara kasus Tuberculosis baru dan indikator program pengendalian Tuberculosis nasional. Kasus baru Tuberculosis dikumpulkan dari catatan dari 16 Pusat Diagnostik dan Pengobatan Tuberkulosis (DTC) pada kurun 2004 - 2006. Dan lima indikator Tuberculosis dipertimbangkan dalam analisis: jumlah kasus yang menjalani peninjauan ulang, jumlah pasien dengan kegagalan pengobatan dan yang menderita kambuh setelah selesainya perawatan, jumlah rumah tangga dengan lebih dari satu kasus, jumlah pasien yang mangkir, dan dekat dengan DTC.

Peneliti (Guy Harling et. el. 2017) menganalisa data sosiodemografi dan klinis untuk semua kasus tuberkulosis baru pada *Notifiable Diseases Information System* (SINAN) dari Fortaleza antara tahun 2007 dan 2014. Penghitungan tingkat insiden untuk 117 lingkungan di Fortaleza, kemudian menilai pengelompokan spasialnya, dan menggunakan model *spatial regression* untuk mengukur hubungan antara kovariat tingkat lingkungan dan tingkat insiden. Menggunakan algoritma *Hierarchical Logistic Regression* untuk mengevaluasi kovariat tingkat individu dan lingkungan memperkirakan pengunduran pengobatan TB.

Peneliti (Dyah WSR Wardani et. al. 2014) menggunakan sumber data dari Catatan medis pasien TB BTA-positif yang menerima perawatan di fasilitas DOTS, berada di 27 pusat kesehatan primer dan satu rumah sakit, selama satu periode. Data alamat rumah dari semua kasus dikonversi menjadi koordinat *longitude* dan *latitude*, Menggunakan alat global positioning system (GPS). Data koordinat kemudian dianalisis dengan menggunakan SaTScan.

Peneliti (A. R. Abdul Rasam et. al. 2016) menggunakan Model indeks berbasis *geographical information system* (GIS) sebagai metode untuk menentukan area berisiko tinggi dari kasus Tuberculosis. Selain itu menggunakan metode pengambilan keputusan *multi-criteria decision making* (MCDM) untuk menentukan faktor risiko lingkungan terhadap penyakit ini dalam skala lima. Skala 1 dan 5 menggambarkan risiko TB terendah dan tertinggi, masing-masing skala 3 sampai 5 dimasukkan sebagai tingkat risiko potensial.

Peneliti (K. Touray et. al. 2010) Data pasien yang sudah diambil lokasinya menggunakan koordinat GPS untuk tempat tinggal dari setiap kasus Tuberculosis. Menggunakan *spatial scan statistic* untuk mengidentifikasi kelompok tuberkulosis.

Lokasi geografis menggunakan penerima GPS Positioning System - penerima GPS Garmin 12 Channel. Semua data ditransfer secara elektronik dan dimasukkan ke Aplikasi Microsoft Access dan diperiksa kesalahannya. Analisis geografi dilakukan dengan menggunakan

teknik Sistem Informasi Geografis (SIG). Versi ArcGIS 9.2 (ESRI, Redlands, CA) digunakan untuk pemetaan semua kasus dan fasilitas kesehatan di wilayahnya.

Peneliti (Adrian Zaragoza Bastida et. al. 2012) Menggunakan *geographic information system* (GIS) dan SCAN. Penentuan dan identifikasi kelompok Tuberculosis dilakukan dengan data statistik program SCAN spasial dan dihitung dengan paket komputer SaTScan (V8.0) yang memperoleh signifikansi statistik dan perkiraan lokalisasi cluster. Program SaTScan memerlukan tiga file untuk inisiasi: berkas lokalisasi kasus, populasi, dan geografi. Model probabilitas Poisson digunakan dengan jendela pencarian 25% dan 50% untuk tingkat tinggi: umur, jenis kelamin, dan indeks marginasi adalah kovariat. Untuk representasi cluster grafis, peta diambil berdasarkan informasi yang diperoleh dengan SCAN dan perangkat lunak IDRISI Taiga.

Peneliti (Saroochi Agarwal et. al. 2017) Analisa data menggunakan Sistem Pengendalian Informasi Genotip (TB GIMS) dari Centers for Disease Control and Prevention (CDC). Analisis *Chi-Square* digunakan untuk menentukan hubungan antara kelompok yang dipilih dan karakteristik minat tertentu. Menggunakan *Geographical Information System* (GIS) ArcGIS 10.4 dan peta hot spot.

Peneliti (Ángela María Segura Cardona et. al. 2013) menggunakan analisa *Time Series* dan distribusi spasial mortalitas

Tuberculosis; Data yang diperoleh dari sertifikat kematian di mana tuberkulosis adalah penyebab kematian di Medellin antara 2000-2007. Analisa *Time Series* bersifat global dan strati oleh komunitas tempat tinggal. Distribusi spatio-temporal angka kematian TB dan HIV untuk periode yang sama, juga dihitung untuk menjelaskan risiko dan distribusi kedua penyakit tersebut, yang menguraikan peta untuk masing-masing penyakit.

Tabel 2. Penelitian terkait , state of the art

No	Peneliti dan Tahun	Judul	Pokok Permasalahan	Metode	Korelasi dengan usulan riset	Usulan
1	(Daiane Leite da Roza et. al. 2012)	Spatio-temporal patterns of tuberculosis incidence in Ribeirão Preto, State of São Paulo, southeast Brazil, and their relationship with social vulnerability: a Bayesian analysis	Analisa Spatio-Temporal dengan Bayesian Regression dalam mengevaluasi sebaran penyakit tuberculosis	Analisis statistik Regresi Bayesian, distribusi Poisson untuk kasus TB baru	Menganalisis sebaran penyakit tuberculosis secara spatio-Temporal	
2	(Li Huang et. al. 2017)	Spatial-temporal analysis of pulmonary tuberculosis in the northeast of the Yunnan province, People's Republic of China	Menganalisa Time Series kasus Tuberculosis untuk memberikan informasi yang berguna bagi para pembuat kebijakan.	Time series dan spatial cluster analyses	Menganalisis sebaran penyakit tuberculosis secara spatio-Temporal	
3	(Kai Cao et. Al. 2016)	Spatial-Temporal Epidemiology of Tuberculosis in Mainland China: An Analysis Based on Bayesian Theory	Mengeksplorasi efek interaksi spasio-temporal dalam Bayesian Framework, menyelidiki faktor-faktor yang mempengaruhi ekologis tuberculosis	Bayesian Methods dan Negative Binomial Distribution	Menganalisis sebaran penyakit tuberculosis secara spatio-Temporal	
4	(Ulziijargal Gurjav et. al. 2015)	Genotype heterogeneity of Mycobacterium tuberculosis within geospatial hotspots suggests foci of imported infection in Sydney, Australia	Menggabungkan hotspot geospasial rinci dan analisis cluster genotipe MIRU-24 dari tuberculosis kultur	Poisson distribution, SatScan : Spatial scan statistics	Menganalisis sebaran penyakit tuberculosis secara spatio-Temporal	
5	(Shiyi Cao et. al. 2013)	A hybrid seasonal prediction model for tuberculosis incidence in China	Mengembangkan model yang sesuai untuk memprediksi epidemi Tuberculosis dan menganalisis musimannya	Seasonal Autoregressive integrated moving average (SARIMA) model, regression neural network	Menganalisis sebaran penyakit tuberculosis secara spatio-Temporal	
6	(Muh Resha et. Al. 2017)	Prediction of Disease Spread Pattern Through Spatio-Temporal Modelling	Mengembangkan model prediksi pola penyakit Tuberculosis dengan Pemodelan Spatio-Temporal	Population Based Incremental Learning	Menganalisis sebaran penyakit tuberculosis secara spatio-Temporal	Pemodelan prediksi sebaran penyakit Tuberculosis secara Spatio-Temporal

Penelitian-penelitian di atas telah memberi gambaran kepada penulis untuk mengkaji pola sebaran penyakit Tuberculosis secara Spatio-Temporal. Penulis rencana membuat pemodelan kasus menggunakan Algoritma Population Based Incremental Learning.

Sehubungan dengan itu penulis rencana akan melakukan penelitian dan mempublikasikannya pada konferensi Nasional dan Internasional berkaitan dengan bidang penelitian.

BAB II

Tinjauan Pustaka

2.1. Pengertian Penyakit Tuberculosis

Penyakit TBC (Tuberkulosis) adalah sebuah penyakit infeksi yang terjadi pada saluran pernafasan manusia yang disebabkan oleh bakteri. Bakteri penyebab penyakit TBC ini merupakan jenis bakteri basil yang sangat kuat sehingga memerlukan waktu yang cukup lama untuk mengobati penyakit TBC ini. Secara umum, bakteri ini lebih sering menginfeksi organ pernapasan paru-paru (90%) dibandingkan dengan bagian lain pada tubuh manusia.

Penyakit TBC (Tuberkulosis) merupakan jenis penyakit menular yang sampai saat ini masih menjadi perhatian seluruh dunia. Sampai saat ini, bahkan belum ada satu negara pun yang bebas dari penyakit TBC. Angka kematian dan kesakitan akibat bakteri mycobacterium tuberculosis ini pun sangat tinggi. Kita harus mengetahui bahwa penyakit TBC dapat menyerang siapa saja (tua, muda, laki-laki, perempuan, miskin, atau kaya) dan juga dapat menyerang dimana saja. Berdasarkan data yang ada, setiap tahunnya di Indonesia bertambah dengan seperempat juta kasus baru TBC dan sekitar 140.000 kematian terjadi setiap tahunnya disebabkan oleh penyakit

TBC. Saat ini Indonesia berada pada urutan ketiga terbesar sebagai negara dengan masalah penyakit TBC di dunia.

2.2. Pola Penyebaran Penyakit Tuberculosis

Bakteri tuberculosis ditemukan oleh Robert Koch pada tahun 1882, tumbuh lambat dengan membelah diri setiap 18-24 jam pada suhu yang optimal. Bakteri ini hidup sebagai parasit intrasel sehingga pertahanan tubuh yang terpenting terhadap bakteri tersebut dilakukan oleh sistem imunitas seluler. Masa inkubasi penyakit TB paru, sejak terinfeksi lesi primer atau reaksi uji tuberkulin yang bermakna adalah 4-12 minggu. Risiko progresif menjadi tuberculosis paru atau tuberculosis di luar paru adalah 1-2 tahun setelah terinfeksi dan mungkin menetap sebagai infeksi laten (Amu, 2008).

Menurut Rieder et al. (2009), bakteri tuberculosis terdapat dalam butir-butir percikan dahak yang disebut droplet nuclei dan melayang di udara untuk waktu yang lama sampai terhisap oleh orang atau mati dengan sendirinya kena sinar matahari langsung. Pada percobaan yang dilakukan pada binatang menunjukkan bahwa droplet nuclei dapat melalui bronkiolus yang paling halus berukuran 2-3 mikron, sehingga diperkirakan jumlah bakteri yang dapat masuk ke alveolus dan menyebabkan penyakit tidak lebih dari satu kuman saja.

Menurut Depkes RI (2006), sumber penularan penyakit tuberkulosis paru dengan BTA positif, yang dapat menularkan kepada orang yang berada disekelilingnya, terutama kontak erat pada waktu batuk/bersin. Penderita menyebarkan bakteri ke udara dalam bentuk droplet (dalam bentuk percikan dahak). Droplet yang mengandung bakteri dapat bertahan di udara pada suhu kamar. Percikan dahak yang mengandung bakteri tuberkulosis yang dibatukan keluar, dihirup oleh orang sehat melalui jalan nafas dan selanjutnya berkembang biak di paru-paru.

Cara Penularan Penyakit Tuberkulosis Paru

Pada perjalanannya bakteri ini banyak mengalami hambatan antara lain di hidung (terhambat oleh bulu hidung) dan lapisan lendir yang melapisi seluruh saluran pernafasan dari atas sampai ke kantong alveoli. Bila penderita baru pertama kali ketularan bakteri tuberkulosis ini, terjadilah suatu proses dalam tubuhnya (paru) yang disebut Primary Complex of Tuberculosis (PCT) yang terdiri dari focus di paru dimana terjadi eksudasi dari sel karena proses dimakannya bakteri tuberkulosis oleh sel macrophag (Rieder et al., 2009).

Menurut Misnadiarly (2006), di negara dimana prevalensi tuberkulosis paru tinggi kebanyakan anak-anak sudah terinfeksi oleh tuberkulosis paru pada tahun-tahun pertama dari kehidupannya, namun yang kemudian menjadi penyakit tuberkulosis paru hanya sedikit saja. Terdapat dua kemungkinan yang terjadi menyusul pembentukan Primary Complex of Tuberculosis (PCT) ini, yaitu:

Dapat sembuh dengan sendirinya karena proses penutupan fokus primer oleh kapsul membran yang akhirnya akan terjadi perkapuran.

Beberapa bakteri akan ikut terlepas ke dalam pembuluh darah dan dapat menginfeksi organ-organ yang terkena. Infeksi yang demikian ini di sebut Post Primary Tuberculosis (PPT) berupa infeksi pada paru, laring dan telinga tengah, kelenjar getah bening di leher, saluran pencernaan dan lubang dubur, saluran kemih, tulang dan sendi

Menurut Anna (2012), penularan tuberkulosis paru terjadi karena bakteri dikeluarkan dengan cara batuk atau bersin oleh penderita menjadi droplet nuclei (percikan dahak) dan terhirup masuk ke pernapasan. Daya penularan ditentukan oleh banyaknya bakteri yang dikeluarkan dari parunya. Makin tinggi derajat positif (gradasi BTA), makin menular penderita tersebut. Secara epidemiologis, seorang penderita tuberkulosis paru positif dapat menularkan pada

10-15 orang setiap tahunnya. Sementara menurut Depkes RI (2011), seseorang yang tertular bakteri tuberkulosis disebabkan oleh beberapa faktor, diantaranya disebabkan oleh daya tahan tubuh yang rendah, karena gizi yang buruk dan infeksi HIV/AIDS.

2.3. K-Means

Algoritma K-means yang merupakan salah satu algoritma *clustering*. Tujuan algoritma ini yaitu untuk membagi data menjadi beberapa kelompok. Algoritma ini menerima masukan berupa data tanpa label kelas. Hal ini berbeda dengan *supervised learning* yang menerima masukan berupa vektor $(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i)$, di mana x_i merupakan data dari suatu data pelatihan dan y_i merupakan label kelas untuk x_i .

Pada algoritma pembelajaran ini, komputer mengelompokkan sendiri data-data yang menjadi masukannya tanpa mengetahui terlebih dulu target kelasnya. Pembelajaran ini termasuk dalam *unsupervised learning*. Masukan yang diterima adalah data atau objek dan k buah kelompok (*cluster*) yang diinginkan. Algoritma ini akan mengelompokkan data atau objek ke dalam k buah kelompok tersebut. Pada setiap *cluster* terdapat titik pusat (*centroid*) yang merepresentasikan *cluster* tersebut.

K-means ditemukan oleh beberapa orang yaitu Lloyd (1957, 1982), Forgey (1965), Friedman and Rubin (1967), and McQueen

(1967). Ide dari *clustering* pertama kali ditemukan oleh Lloyd pada tahun 1957, namun hal tersebut baru dipublikasi pada tahun 1982. Pada tahun 1965, Forgey juga mempublikasi teknik yang sama sehingga terkadang dikenal sebagai Lloyd-Forgey pada beberapa sumber.

Algoritma untuk melakukan K-Means *clustering* adalah sebagai berikut^[3]:

1. Pilih K buah titik *centroid* secara acak
2. Kelompokkan data sehingga terbentuk K buah *cluster* dengan titik *centroid* dari setiap *cluster* merupakan titik *centroid* yang telah dipilih sebelumnya
3. Perbaharui nilai titik *centroid*
4. Ulangi langkah 2 dan 3 sampai nilai dari titik *centroid* tidak lagi berubah

Proses pengelompokan data ke dalam suatu *cluster* dapat dilakukan dengan cara menghitung jarak terdekat dari suatu data ke sebuah titik *centroid*. Perhitungan jarak Minkowski dapat digunakan untuk menghitung jarak antar 2 buah data. Rumus untuk menghitung jarak tersebut adalah^[4]:

$$d(x_i, x_j) = (|x_{i1} - x_{j1}|^g + |x_{i2} - x_{j2}|^g + \dots + |x_{ip} - x_{jp}|^g)^{1/g}$$

Di mana:

$g = 1$, untuk menghitung jarak Manhattan

$g = 2$, untuk menghitung jarak Euclidean

$g = \infty$, untuk menghitung jarak Chebychev

x_i, x_j adalah dua buah data yang akan dihitung jaraknya

p = dimensi dari sebuah data

Pembaharuan suatu titik centroid dapat dilakukan dengan rumus berikut^[4]:

$$\mu_k = \frac{1}{N_k} \sum_{q=1}^{N_k} x_q$$

Di mana:

μ_k = titik centroid dari cluster ke-K

N_k = banyaknya data pada cluster ke-K

x_q = data ke-q pada cluster ke-K

Ada beberapa kelebihan pada algoritma k-means, yaitu ^[2]:

1. Mudah untuk diimplementasikan dan dijalankan.
2. Waktu yang dibutuhkan untuk menjalankan pembelajaran ini relatif cepat.
3. Mudah untuk diadaptasi.
4. Umum digunakan.

Algoritma k-means memiliki beberapa kelebihan, namun ada kekurangannya juga. Kekurangan dari algoritma tersebut yaitu :

1. Sebelum algoritma dijalankan, k buah titik diinisialisasi secara *random* sehingga pengelompokan data yang dihasilkan dapat berbeda-beda ^[1]. Jika nilai *random* untuk inisialisasi kurang baik, maka pengelompokan yang dihasilkan pun menjadi kurang optimal.
2. Dapat terjebak dalam masalah yang disebut *curse of dimensionality*. Hal ini dapat terjadi jika data pelatihan memiliki dimensi yang sangat tinggi (Contoh jika data pelatihan terdiri dari 2 atribut maka dimensinya adalah 2 dimensi. Namun jika ada 20 atribut, maka akan ada 20 dimensi). Salah satu cara kerja algoritma ini adalah mencari jarak terdekat antara k buah titik dengan titik lainnya. Jika mencari jarak antar titik pada 2 dimensi, masih mudah dilakukan. Namun bagaimana mencari jarak antar titik jika terdapat 20 dimensi. Hal ini akan menjadi sulit.
3. Jika hanya terdapat beberapa titik sampel data, maka cukup mudah untuk menghitung dan mencari titik terdekat dengan k titik yang diinisialisasi secara *random*. Namun jika terdapat banyak sekali titik data (misalnya satu milyar buah data), maka perhitungan dan pencarian titik terdekat akan membutuhkan waktu yang lama. Proses tersebut dapat

dipercepat, namun dibutuhkan struktur data yang lebih rumit seperti *kD-Tree* atau *hashing*.

2.4. Analisa Data Spatio-Temporal

Analisis data spatiotemporal adalah daerah penelitian yang muncul karena pengembangan dan penerapan teknik komputasi baru yang memungkinkan analisis database spatiotemporal besar. Model spatiotemporal muncul saat data dikumpulkan sepanjang waktu dan juga ruang dan setidaknya memiliki satu spasial dan satu properti temporal. Suatu peristiwa dalam kumpulan data spatiotemporal menggambarkan fenomena spasial dan temporal yang ada pada waktu dan lokasi tertentu x . Contohnya adalah pola kematian wanita kanker payudara di AS antara 1990-2010, di mana properti spasial adalah lokasi dan geometri objek - negara bagian AS dengan informasi tingkat kematian kanker payudara, dan properti temporal adalah cap waktu Atau interval waktu dimana objek spasial valid - 1990-2010 tahun kematian kanker payudara. Aplikasi lain untuk analisis spatiotemporal mencakup kasus-kasus di bidang biologi, ekologi, meteorologi, kedokteran, transportasi dan kehutanan.

Peluang dan Tantangan Analisis Spatiotemporal

Meskipun pendekatan ini dapat memberikan dimensi baru untuk interpretasi data, namun masih dalam tahap awal dan bahkan pertanyaan paling mendasar di bidang ini masih belum terjawab: pola macam apa yang dapat diambil dari lintasan dan metode dan algoritma mana yang harus diterapkan untuk mengekstraknya. mereka? Penting agar penyidik mengetahui hal ini pada awal analisis. Penting juga untuk dicatat bahwa banyak dari ini masih belum memiliki jawaban atau solusi yang solid, seperti yang diidentifikasi melalui literatur.

Analisis data spatiotemporal mensyaratkan bahwa kedua korelasi temporal dan korelasi spasial diperhitungkan. Menilai dimensi data temporal dan spasial menambahkan kompleksitas yang signifikan pada proses analisis data karena dua alasan utama: 1) Perubahan sifat spasial dan non spasial obyek spatiotemporal secara kontinyu dan diskrit dan 2) pengaruh objek spatiotemporal tetangga yang berkelompok pada satu lain.

Isu utama muncul saat memasukkan ruang dan waktu dalam satu model:

Ruang dua dimensi dan memiliki directionality tak terbatas: NSEW dan di mana-mana di antaranya sementara waktu bersifat unidimensional dan hanya bisa bergerak dalam satu arah - FORWARD, dan dengan demikian menantang cara di mana kita dapat menafsirkan hasil analisis spatiotemporal.

Masalah lain muncul dengan bagaimana data didefinisikan dan dapat memiliki dampak kuat pada pola yang ditemukan. Salah satu masalah yang terus berlanjut adalah Permasalahan Unit Areal Modifiable (MAUP):

Penyelidik bisa mendapatkan jawaban yang sama sekali berbeda tergantung pada apakah ruang dinilai oleh negara bagian atau kode pos atau saluran sensus, dan apakah waktu dinilai berdasarkan tahun atau hari atau menit. Analisis yang sama persis dapat dilakukan dengan definisi spasial / temporal yang berbeda dan hasilnya dapat menghasilkan jawaban yang sama sekali berbeda. Oleh karena itu, tergantung bagaimana penyidik mendefinisikan datanya, mereka mungkin akan menjadi menarik namun memiliki pola palsu

Masalah semacam itu dapat mempengaruhi interpretabilitas analisis. Karena kita tidak melihat dua variabel dengan arah yang sama, kita tidak bisa begitu saja menafsirkan beta seperti kita terbiasa dan harus

mempertimbangkan faktor ini. Ini bahkan lebih kompleks lagi jika Anda mulai mengubah data Anda untuk dianalisis. Konsep yang diperlukan untuk menggambarkan unit analisis dapat terjadi pada tingkat konseptual yang lebih tinggi yang lebih sesuai untuk interpretasi dan implikasi kebijakan.

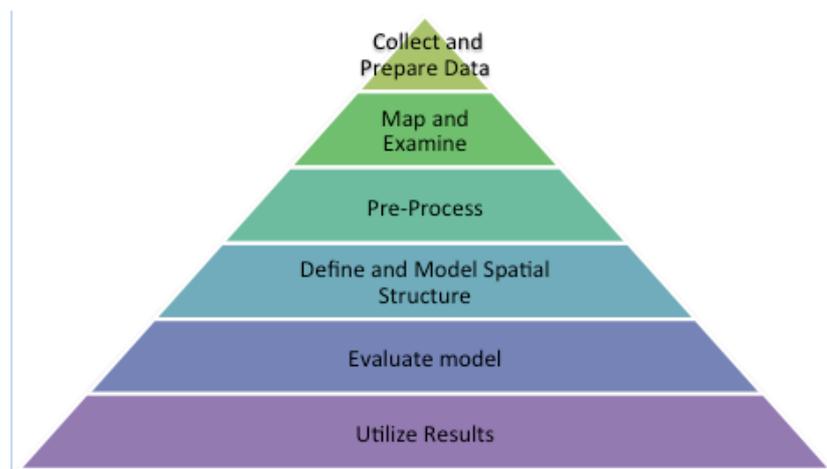
Algoritma Analisis Spatiotemporal

Dengan beberapa tantangan ini, kami memberikan panduan tentang bagaimana analisis spatiotemporal dilakukan, dengan menggunakan pendekatan yang sebagian besar bersifat umum.

Tujuannya adalah untuk memberi Anda cukup informasi untuk mengetahui apakah Anda ingin mulai bekerja dengan data spatiotemporal, bagaimana Anda bisa mulai menilai apakah data Anda sesuai, dan memberi Anda sumber daya untuk menganalisis lebih lanjut.

Dalam prakteknya, dua tujuan utama analisis spatiotemporal cenderung merupakan prediksi dan deskripsi. Untuk kesederhanaan, kita akan fokus pada alur kerja yang akan terjadi saat Anda melakukan analisis spatiotemporal deskriptif dengan fokus penyakit kronis.

Namun, pendekatan ini bisa menjadi model prediksi, yang mana kita menyediakan sumber daya untuk dijelajahi sesuai kebutuhan.



Gambar 4. Langkah 1: Kumpulkan dan Siapkan Data

Persyaratan utama adalah bahwa semua data harus dikaitkan dengan komponen spasial dan temporal. Data spasial dapat dianalisis pada banyak tingkatan, kode pos, saluran sensus, negara bagian, geokode, dll. Data temporal sering dianalisis sebagai beberapa titik data per pengamatan dari waktu ke waktu dan dapat diukur dengan sebanyak mungkin data spasial, jika tidak lebih. Kita juga bisa melihat kejadian per tahun, bulan, menit, detik, tapi ini juga harus dikaitkan dengan data spasial.

Kebanyakan pilihan untuk menentukan data spasial dan temporal Anda adalah tantangan, terutama bila Anda ingin membandingkan

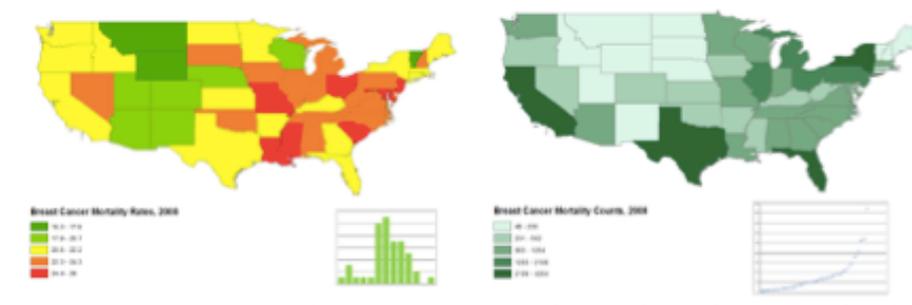
hasilnya dengan penelitian lain atau membuatnya berlaku untuk populasi lain.

Jika Anda mengumpulkan data Anda sendiri, melakukan penelitian sebelum pengumpulan data penting untuk melihat bagaimana orang lain menentukan masalahnya. Jika Anda cukup beruntung memiliki data primer yang Anda kumpulkan, Anda dapat menentukan pengamatan Anda dengan analisis Anda. Tapi seringkali, ini bukan masalahnya dan kami menganalisis data yang kami dapatkan dari sumber lain. Hal ini semakin sering untuk bisa mendapatkan informasi yang terkait dengan data spasial, yang sangat membantu saat Anda mengumpulkan data untuk analisis Anda. Saat melakukan analisis kompleks seperti analisis spatiotemporal, Anda memerlukan begitu banyak komponen data dan jarang bisa menemukan semua data dari hanya satu sumber. Seringkali Anda dapat mengumpulkan sebagian dari data itu sendiri dengan analisis khusus Anda dan Anda dapat memperoleh komponen lain dari sumber data sekunder yang berbeda. Semakin banyak database termasuk data spasial yang menyebabkan semakin banyak orang yang memiliki kemampuan untuk melakukan analisis ini.

Langkah 2: Peta dan Periksa

Begitu kita memiliki data kita, seperti dalam semua penelitian berbasis data, langkah selanjutnya adalah mulai memeriksa data. Sama seperti ketika kita menjalankan frekuensi sederhana atau tabir silang untuk mengeksplorasi bentuk data lainnya, kita melakukan pendekatan umum yang sama dengan data spatiotemporal. Selain analisis data deskriptif, kita bisa memeriksa data kita dengan menggunakan peta deskriptif sederhana. Dengan melakukan ini, kita bisa mendapatkan visualisasi yang jelas tentang karakteristik atau tren penting yang mungkin terkait dengan data spasial yang mungkin tidak kita lihat hanya dengan melihat data. Kita juga dapat menentukan outlier, data yang berpotensi salah, dan jumlah sel kecil atau besar yang mungkin menjadi masalah.

Contoh:



Gambar Peta A

Gambar Peta B

Gambar 5. Peta A dan Peta B

Tingkat kematian kanker payudara dipetakan menggunakan warna yang jelas, untuk melihat bahwa mungkin ada beberapa pengelompokan di sekitar negara bagian dengan tingkat terendah (hijau) dan tingkat tertinggi (merah). Histogram tarif juga diproduksi untuk melihat potensi outlier dan ada satu negara bagian (Mississippi) dengan tingkat yang lebih tinggi daripada negara bagian lainnya (28 kematian per 100.000). Dengan adanya temuan ini, Anda akan mengecek apakah tingkat ini konsisten dengan tahun-tahun lainnya dan di negara bagian lain.

Peta B: Di peta ini kita menggunakan jumlah kematian untuk melihat potensi jumlah sel kecil atau data yang keliru yang dapat menyebabkan masalah. Kami memiliki satu outlier besar, namun analisis lebih lanjut menunjukkan bahwa negara bagian adalah California, yang memiliki populasi yang sangat besar untuk memulai sehingga mungkin kurang memprihatinkan. Peta tersebut menunjukkan bahwa ada banyak jumlah kecil di daerah dengan populasi non-terkonsentrasi (yaitu, negara bagian barat, Maine, dll). Anda mungkin juga ingin melihat lebih jauh untuk mencari tahu mengapa - apakah jumlah ini rendah karena populasi rendah atau apakah itu sesuatu yang lain seperti masalah pengumpulan data. Selain jumlah kecil, kita bisa melihat tempat pengelompokan di peta

ini sehingga kita tidak dapat melihat dengan frekuensi tradisional.
Mengapa kedua cluster ini ada di sini?

Langkah 3: Pra Proses

Data spatiotemporal mungkin perlu ditransformasikan sebelum dianalisis. Jika perlu, gunakan teknik untuk memusatkan data dan gunakan transformasi agar data sesuai dengan distribusi normal.

Aspek penting lainnya adalah menguji observasi spasial non-independen. Perlu khawatir tentang pengelompokan, dan tergantung pada apa data Anda terlihat dan pengelompokan apa yang Anda harapkan, Anda menggunakan metode yang berbeda. Berbagai cara pengelompokan bisa terjadi adalah:

Pengelompokan spasial berdasarkan nilai atribut spasial non-spasial dari objek ST

Clustering benda bergerak

Kerapatan clustering

Jika pengelompokan ditemukan, Anda mungkin perlu mentransformasikan data menggunakan algoritma yang mengekstrak kumpulan statistik potensial

Seringkali dalam data spatiotemporal sebuah isu yang dapat menyebabkan bias adalah adanya autokorelasi. Ini berawal dari persyaratan yang telah kita bahas sebelumnya tentang model analisis bahwa semua objek spasial saling independen satu sama lain dan semua data temporal tidak bergantung padanya.

Autokorelasi Spasial:

Autokorelasi adalah mekanisme dimana subjek yang tinggal berdekatan dapat lebih mirip dari yang diharapkan dengan memberikan distribusi spasial yang benar-benar acak. Atau, berdasarkan contoh tingkat kematian akibat kanker payudara di AS, negara-negara lebih dekat satu sama lain mungkin lebih mirip daripada negara bagian yang jauh dari satu sama lain.

Dibandingkan dengan korelasi tradisional, yang melihat hubungan antara dua variabel, autokorelasi mengukur korelasi antara variabel X , dan nilai rata-rata X untuk negara tetangga (negara, kode pos, orang). Jika autokorelasi disebabkan oleh faktor yang tidak terukur yang berkorelasi spasial dengan variabel minat Anda, maka akan diperkenalkan bias terhadap hasil analisis. Kehadiran autokorelasi melanggar asumsi independensi dan model yang dihasilkan mungkin

memiliki perkiraan parameter yang tidak stabil dan nilai p yang tidak dapat diandalkan untuk analisis regresi apapun.

Metode yang paling sering digunakan untuk menilai autokorelasi adalah nilai Morans I. Ini adalah perhitungan yang paling umum karena Anda dapat menggunakan data titik atau poligon (seperti negara bagian), dan Anda juga dapat menyertakan semua tipe data, apakah Anda memiliki variabel kategoris, biner, atau kontinyu, jadi ini merupakan awal yang baik saat menilai data Anda. .

Langkah 4: Tentukan dan Model Struktur Spasial

Ada banyak model yang ditempatkan dalam kerangka spatiotemporal dan dapat digunakan untuk jenis analisis ini. Kami menemukan bahwa mengatakan 'analisis spatiotemporal' hampir sama luasnya dengan mengatakan 'regresi' yang secara bersamaan memudahkan menganalisis data Anda dalam metode ini karena Anda dapat menyesuaikan banyak model dan sulit untuk dianalisis karena jarang ada metode pemotongan yang jelas menggunakan.

Berikut adalah beberapa metode yang sering digunakan dalam literatur:

Autoregresi Bersyarat:

Autoregression sering digunakan dalam data seri longitudinal atau time series, dan memodelkan variabel hasil karena bergantung secara linear pada nilai awal itu sendiri. Ini adalah akun terbaik untuk efek lokal, jadi jika Anda berharap melihat banyak variabilitas spasial (perbedaan antar individu), ini mungkin metode yang efektif.

Space-Time Autoregressive Integrated Moving Average:

Hasil tidak hanya bergantung pada nilai sebelumnya dari waktu ke waktu, tetapi juga nilai sebelumnya di ruang angkasa. Sering digunakan untuk data dengan jarak yang jauh antara ruang dan titik waktu dan dataset yang sangat besar

Spasial Multivariate Age-Period-Cohort (APC) Efek:

Mempertimbangkan efek APC serta efek geografis yang berbeda terhadap perilaku. Sering digunakan dalam model kanker untuk menilai hubungan di mana orang tinggal, bagaimana itu mempengaruhi perilaku mereka, selain efek APC klasik yang kita lihat pada kanker.

Model P-spline:

Menyediakan estimasi parameter yang merapikan sepanjang ruang dan waktu pada skala global yang besar. Smoothing dilakukan dalam tiga dimensi (bujur, garis lintang, dan waktu). Ini bisa berguna jika Anda mengharapkan perubahan signifikan pada titik waktu yang berbeda. Misalnya, metode ini dapat digunakan jika Anda ingin melihat dampak perawatan kesehatan terhadap hasil penyakit di seluruh negara bagian sebelum dan sesudah Undang-Undang Perawatan Kesehatan Terjangkau.

Langkah 5: Evaluasi Model

Untuk mengevaluasi kualitas model, analisis kemudian meneliti residu model.

Distribusi temporal dari residual dieksplorasi dengan menggunakan grafik waktu dan distribusi spasial melalui tampilan peta. Model dianggap benar menghasilkan atau menangkap fitur umum variasi spatiotemporal bila tidak adanya pola temporal dan spasial yang jelas, atau dengan kata lain, distribusi untuk setiap dimensi tampak sebagai noise acak.

Jika distribusi acak tidak terbentuk, analis dapat memilih untuk memodifikasi model atau segmen kelompok dan merevisi analisisnya. Faktor kunci lainnya yang perlu dipertimbangkan dalam tahap evaluasi adalah dengan melihat asumsi utama model ST teoritis - semua struktur temporal harus ditangkap oleh fungsi dasar temporal yang halus dan dependensi spasial harus menunjukkan keterpaduan.

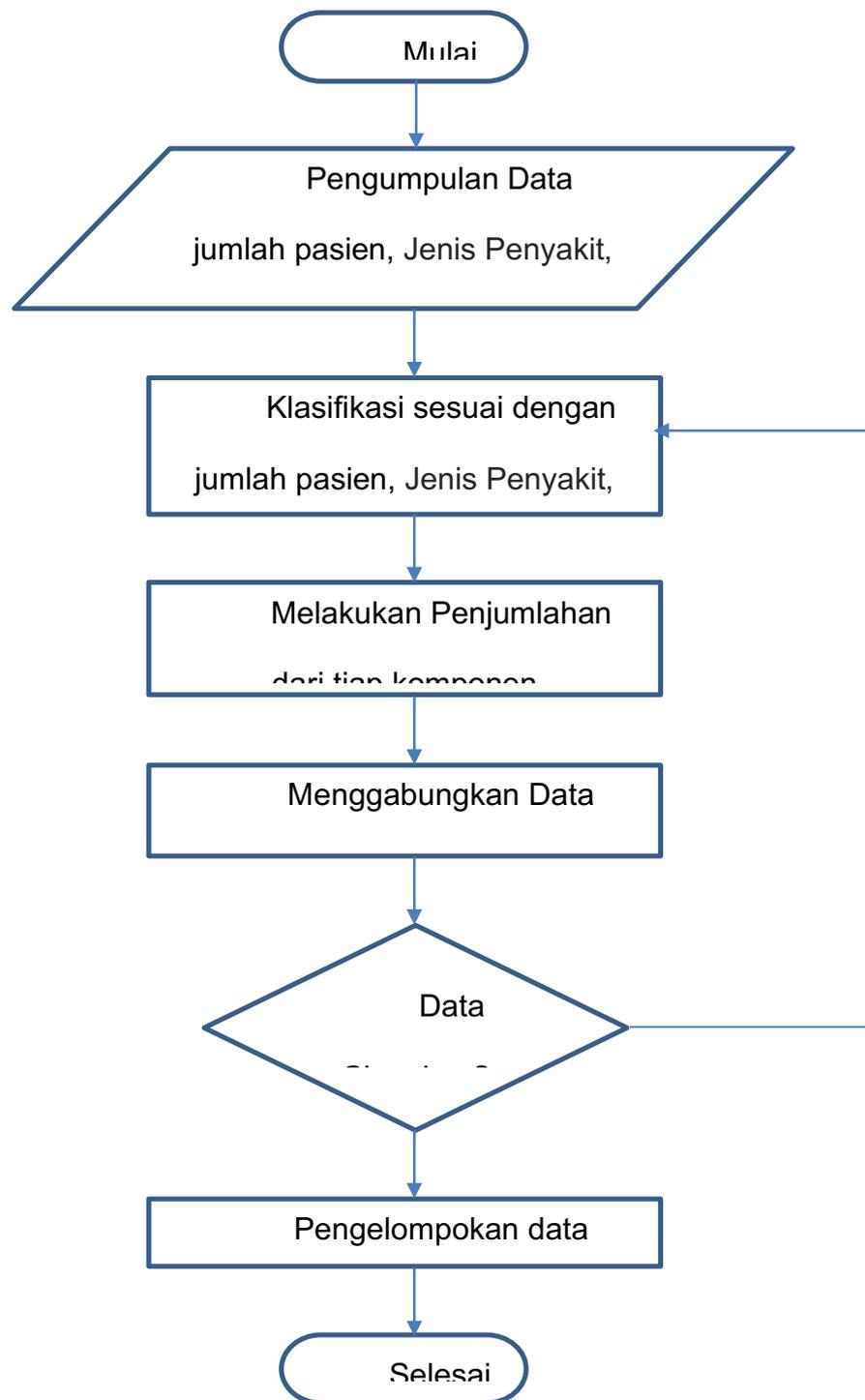
Langkah 6: Manfaatkan Hasil

Tugas terakhir dalam pendekatan analitik dengan analisis data spatiotemporal untuk memanfaatkan hasilnya.

BAB III

METODE PENELITIAN

3.1. Diagram Alir Penelitian



Gambar 6. Diagram Alir Penelitian

3.2. Kerangka Berpikir Penelitian

Sistematika berpikir penelitian ini mengikuti alur kerangka berpikir seperti diagram di bawah ini :

Kajian	Identifikasi Masalah	Pemodelan	Output
<ul style="list-style-type: none">• Pola sebaran TB• Data Cleaning• Pengelompokan• Spatio Temporal	<ul style="list-style-type: none">• Dibutuhkan pemetaan sebaran penyakit• Prediksi pola sebarang Tuberculosis	<ul style="list-style-type: none">• Mendesain Spatio-Temporal• Mendesain pemetaan data sebaran• Menggunakan Population-Based Incremental Learning	<ul style="list-style-type: none">• Model prediksi sebaran penyakit Tuberculosis secara Spatio-Temporal menggunakan Population Based incremental learning

Tabel 3. Kerangka berpikir Penelitian

3.3. Teknik Pengumpulan Data

Dalam mengumpulkan data untuk penelitian, penulis akan melakukan hal-hal sebagai berikut :

1. Data pasien Tuberculosis yang dicatat oleh petugas puskesmas di Makassar.
2. Pengambilan data Fluktuasi Cuaca, dengan Sumber dari BMKG
3. Pengambilan data Jumlah Penduduk dari Dinas Kependudukan dan Catatan Sipil
4. Pengumpulan data tingkat Sanitasi Lingkungan dari Dinas Kesehatan Kota Makassar
5. Pengumpulan Data Kualitas Udara dari Badan Lingkungan Hidup Kota Makassar

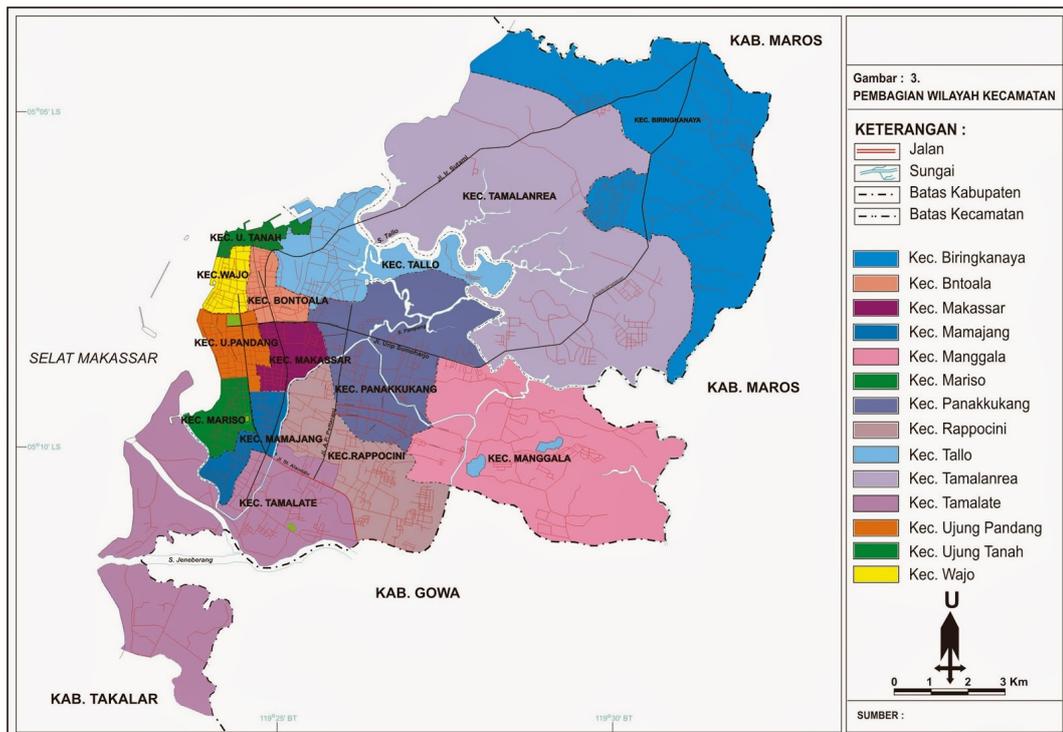
Daftar Puskesmas di Makassar

Puskesmas	Alamat	Kelurahan	Kecamatan
1. Pkm Dahlia.	Seroja No.3,	Kampung Buyang,	Mariso.
2. Pkm Panambungan.	Rajawali Lr.13B No.13,	Panambungan,	Mariso.
3. Pkm Pertiwi.	Cendrawasih III No. 2,	Mariso,	Mariso.
4. Pkm Mamajang.	Baji Minasa No.10	Mamajang Luar,	Mamajang.
5. Pkm Cendrawasih.	Cendrawasih No.404	Sambung Jawa,	Mamajang.
6. Pkm Barombong.	Perjanjian Buangaya No.13,	Barombong,	Tamalate.
7. Pkm Jongaya.	Andi Tonro No.70A,	Jongaya,	Tamalate.
8. Pkm Tamalate.	Dg.Tata I (Komp.Tabaria Blok G 8	Pr Tambung,	Tamalate.
9. Pkm Maccini Sombala.			Tamalate.
10. Pkm Kassi-Kassi.	Tamalate I No.43,	Kassi-kassi,	Rappocini.
11. Pkm Minasa Upa.	Minasa Upa Raya No.18,	Gunung Sari,	Rappocini.
12. Pkm Mangasa.	Monumen Emmy Saelan Komp BTN M 11,		Rappocini.
13. Pkm Bara-Baraya.	Abubakar Lambogo,	Bara-Baraya,	Makassar.
14. Pkm Maradekaya..	Sungai Saddang Baru Lr.5 No.27	Mandekaya,	Makassar.
15. Pkm Maccini Sawah.	Maccini Raya	Maccini,	Makassar.
16. Pkm Makkasau.	Ratulangi,		Ujung Pandang.
17. Pkm Tarakang.	Kodingareng Lr.181 No.5,	Malimongan Tua,	Wajo.
18. Pkm Andalas.	Sangir Lr.209 No.6,	Melayu,	Wajo.
19. Pkm Layang.	Tinumbu Lr.148 No.1-2,	Layang,	Bontoala.
20. Pkm Malimongan Baru.	Pontiku Lr.5 No.1,	Malimongan Baru,	Bontoala.
21. Pkm Pattingalloang.	Barukang VI No.15,	Pattingalloang Lama,	Ujung Tanah.
22. Pkm Barang Lompo.	Pulau Barraang Lompo Rw 1,		Ujung Tanah.
23. Pkm Tabaringan.	Tinumbu Lr.154	Ujung Tanah,	Ujung Tanah.
24. Pkm Pulau Kodingareng.	Pulau Kodingareng		Ujung Tanah.
25. Pkm Jumpang Baru.	Ade Irma Nasution Blok C No.1,	Rappo Jawa,	Tallo.
26. Pkm Kaluku Bodoa.	Butta Caddi No.15	Kaluku Bodoa,	Tallo.
27. Pkm Rappokalling.	Daeng Ragge	Rappokalling,	Tallo.
28. Pkm Batua.	Abd Daeng Sirua No.338	Tello Baru,	Panakkukang.
29. Pkm Karuwisi.	Urip Sumiharjo Lr.2,	Karuwisi,	Panakkukang.
30. Pkm Pampang.	Pampang III No. 28 A,		Panakkukang.

31. Pkm Tamamaung.	Abdul Daeng Sirua No.158,		Panakkukang.
32. Pkm Antang.	Antang Raya No.43,	Antang,	Manggala.
33. Pkm Antang Perumnas.	Lasuloro	Manggalo,	Manggala.
34. Pkm Tamangapa.	Tamangapa Raya No.26 H,		Manggala.
35. Pkm Bangkala.	Kec. Manggala.		
36. Pkm Sudiang Raya.	Makasar Raya,		Biring Kanaya.
37. Pkm Sudiang.	Gowa Raya,		Biring Kanaya.
38. Pkm Kapasa.	Komp. BTN Angkatan Laut		Biringkanaya.
39. Pkm Antara.	Komp.BTN Antara Blok B No.6,		Tamalanrea.
41. Pkm Tamalanrea.	Kesejahteraan Timur I BTP Blok B.		Tamalanrea.
42. Pkm Bulurokeng.			Tamalanrea.
43. Pkm Paccerakkang.			Tamalanrea.

Tabel 4. Daftar Puskesmas di Makassar

Peta Makassar



Gambar 7. Peta Makassar

3.4. Jenis dan Sumber Data

Pada penelitian ini penulis menggunakan data kuantitatif berupa angka yang diolah menggunakan model matematis statistikal, ditunjang dengan teori-teori dan hipotesis. Data yang digunakan disesuaikan dengan tujuan penelitian, sebagai berikut :

- Semua Pasien TB Paru berusia 15 – 50 tahun
- Pasien yang tercatat pada buku register TB Paru pada pemeriksaan sputum ditemukan kuman Mycobacterium Tuberculosis BTA Positif maupun Negatif
- Pasien yang tercatat pada Sistem Informasi Laboratium dengan hasil pemeriksaan Positif dan Negatif

3.5. Waktu dan Lokasi Penelitian

Adapun waktu Penelitian kurang lebih selama 3 semester dimulai pada periode Semester Ganjil 2017/2018 hingga Semester Ganjil 2018/2019.

Lokasi penelitian akan memilih area seluruh kecamatan di Kota Makassar. Data Puskesmas akan mewakili seluruh area kota makassar. Basis pengujian dan simulasi program akan bertempat di Laboratorium Jurusan Teknik Elektro Universitas Hasanuddin.

3.6. Bahan dan Alat yang digunakan

Penelitian ini menggunakan beberapa perangkat hardware dan software untuk simulasi dan analisa data sebagai berikut :

- Perangkat Keras/Hardware berupa Laptop Macintosh
- Perangkat Lunak/Software : Bahasa Pemrograman Phyton, MatLab,

DAFTAR PUSTAKA

- [1] S. Han, Y. Dang, S. Ge, and D. Zhang, "Frequent Pattern Mining," *US Pat. 20,120,278,346*, pp. 105–134, 2012.
- [2] S. Agarwal, D. T. Nguyen, L. D. Teeter, and E. A. Graviss, "Spatial-temporal distribution of genotyped tuberculosis cases in a county with active transmission," *BMC Infect. Dis.*, vol. 17, no. 1, p. 378, 2017.
- [3] A. Zaragoza Bastida *et al.*, "Spatial and temporal distribution of tuberculosis in the State of Mexico, Mexico.," *ScientificWorldJournal.*, vol. 2012, no. JULY, p. 570278, 2012.
- [4] K. Touray *et al.*, "Spatial analysis of tuberculosis in an Urban West African setting: Is there evidence of clustering?," *Trop. Med. Int. Heal.*, vol. 15, no. 6, pp. 664–672, 2010.
- [5] A. R. Abdul Rasam, N. M. Shariff, and J. F. Dony, "Identifying High-Risk Populations of Tuberculosis Using Environmental Factors and Gis Based Multi-Criteria Decision Making Method," *ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. XLII-4/W1, no. October, pp. 9–13, 2016.
- [6] D. W. Wardani, L. Lazuardi, Y. Mahendradhata, and H. Kusnanto, "Clustered Tuberculosis Incidence in Bandar Lampung, Indonesia," *WHO South-East Asia J. Public Heal.*, vol. 3, no. 2, pp. 179–185, 2014.
- [7] G. Harling, A. S. Lima Neto, G. S. Sousa, M. M. T. Machado, and M. C. Castro, "Determinants of tuberculosis transmission and treatment abandonment in Fortaleza, Brazil," *BMC Public Health*, vol. 17, no. 1, p. 508, 2017.
- [8] R. V Rendremanana *et al.*, "To cite this version : Bayesian mapping of pulmonary tuberculosis in Antananarivo , Madagascar," 2013.
- [9] S. Cao *et al.*, "A hybrid seasonal prediction model for tuberculosis incidence in China," *BMC Med. Inform. Decis. Mak.*, vol. 13, no. 56, pp. 1–7, 2013.
- [10] A. S. Unit, "Spatio Temporal Mapping of Tuberculosis using Log-Empirical Bayesian Kriging Methods Kasturi Basu," no. March, 2016.
- [11] D. Onozuka and A. Hagihara, "Geographic prediction of tuberculosis clusters in Fukuoka, Japan, using the space-time scan statistic.," *BMC Infect. Dis.*, vol. 7, p. 26, 2007.

- [12] U. Gurjav, P. Jelfs, G. A. Hill-Cawthorne, B. J. Marais, and V. Sintchenko, "Genotype heterogeneity of *Mycobacterium tuberculosis* within geospatial hotspots suggests foci of imported infection in Sydney, Australia," *Infect. Genet. Evol.*, vol. 40, pp. 346–351, 2016.
- [13] M. Gonzalez-juarrero *et al.*, "Temporal and Spatial Arrangement of Lymphocytes within Lung Granulomas Induced by Aerosol Infection with *Mycobacterium tuberculosis* Temporal and Spatial Arrangement of Lymphocytes within Lung Granulomas Induced by Aerosol Infection with *Mycobacteri*," *Infect. Immun.*, vol. 69, no. 3, pp. 1722–1728, 2001.
- [14] K. Cao *et al.*, "Spatial-temporal epidemiology of tuberculosis in mainland China: An analysis based on Bayesian theory," *Int. J. Environ. Res. Public Health*, vol. 13, no. 5, pp. 4–8, 2016.
- [15] L. Huang *et al.*, "Spatial-temporal analysis of pulmonary tuberculosis in the northeast of the Yunnan province, People's Republic of China," *Infect. Dis. Poverty*, vol. 6, no. 1, p. 53, 2017.
- [16] D. L. Da Roza, M. D. C. G. G. Caccia-Bava, and E. Z. Martinez, "Spatio-temporal patterns of tuberculosis incidence in Ribeirão Preto, State of São Paulo, southeast Brazil, and their relationship with social vulnerability: a Bayesian analysis," *Rev Soc Bras Med Trop*, vol. 45, no. 5, pp. 607–615, 2012.
- [17] J. Barat, M. Rahmaniati, D. Biostatistik, and F. K. Masyarakat, "Pola Penyebaran Kasus Tuberkulosis Dengan Pendekatan Spasial-Statistik (Studi kasus di Kota Depok, Jawa Barat)," pp. 1–15, 2012.
- [18] G. D. Alvina Karolina Bagah, Kandou and H. Palandeng, "Pemetaan kasus tuberkulosis paru di kecamatan tuminting tahun 2013," pp. 34–40, 2013.

