KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET DAN TEKNOLOGI
UNIVERSITAS MULAWARMAN
LEMBAGA PENELITIAN DAN PENGABDIAN KEPADA MASYARAKAT
Alamat :Jl.Kerayan No. 1 Gedung A8, Kampus Gunung Kelua,Samarinda 75119
Laman : http://lp2m.unmul.ac.id Surel: lppm@unmul.ac.id

KONTRAK PENELITIAN TAHUN JAMAK
PENELITIAN DASAR DAN PEMBINAAN/KAPASITAS
TAHUN ANGGARAN 2021
Nomor: 597/UN17.L1/PG/2021

Pada hari ini **Kamis** tanggal **Delapan belas** bulan **Maret** tahun **Dua Ribu Dua Puluh Satu**, kami yang bertandatangan dibawah ini :

1. ANTON RAHMADI	:	Ketua Lembaga Penelitian dan Pengabdian kepada Masyarakat, Universitas Mulawarman, dalam hal ini bertindak untuk dan atas nama Universitas Mulawarman, yang berkedudukan di Jalan Kerayan no. 1 Kampus Gn. Kelua Samarinda, untuk selanjutnya disebut **PIHAK PERTAMA;**
2. SIFRIYANI	:	Dosen Fakultas MIPA Universitas Mulawarman, dalam hal ini bertindak sebagai pengusul dan Ketua Pelaksana Penelitian Tahun Anggaran 2021 untuk selanjutnya disebut **PIHAK KEDUA**.

PIHAK PERTAMA dan **PIHAK KEDUA**, secara bersama-sama bersepakat mengikatkan diri dalam suatu Kontrak Penelitian dengan ketentuan dan syarat sebagai berikut:

PASAL 1
DASAR HUKUM

Kontrak penelitian ini berdasarkan kepada :
1. Undang-Undang Republik Indonesia Nomor 17 tahun 2003 tentang Keuangan Negara;
2. Undang-Undang Republik Indonesia Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional;
3. Undang-Undang Republik Indonesia Nomor 01 Tahun 2004 tentang Perbendaharaan Negara;
4. Undang-Undang Republik Indonesia Nomor 15 Tahun 2004 tentang Pemeriksaan Pengelolaan dan Tanggung Jawab Keuangan Negara;
5. Undang-Undang Republik Indonesia Nomor 12 Tahun 2012 tentang Pendidikan Tinggi;
6. Undang-Undang Nomor 11 Tahun 2019 tentang Sistem Nasional Ilmu Pengetahuan dan Teknologi;
7. Peraturan Pemerintah Nomor 26 Tahun 2015 tentang Bentuk dan Mekanisme Perguruan Tinggi Negeri Badan Hukum sebagaimana telah diubah dengan Peraturan Pemerintah Nomor 8 Tahun 2020 tentang Perubahan Atas Peraturan Pemerintah Nomor 26 Tahun 2015 tentang Bentuk dan Mekanisme Pendanaan Perguruan Tinggi Negeri Badan Hukum;
8. Peraturan Presiden Nomor 16 Tahun 2018 tentang Pengadaan Barang dan Jasa Pemerintah;
9. Peraturan Presiden Nomor 50 Tahun 2020 tentang Kementerian Riset dan Teknologi;

10. Keputusan Presiden Nomor 113/P Tahun 2019 tentang Pembentukan Kementerian dan Pengangkatan Menteri Kabinet Kerja Periode Tahun 2019-2024;
11. Peraturan Menteri Keuangan Nomor 119/PMK.02/2020 Standar Biaya Masukan Tahun Anggaran 2021;
12. Peraturan Menteri Keuangan Nomor 112/PMK.02/2020 tentang Standar Biaya Keluaran Tahun Anggaran 2021;
13. Peraturan Menteri Keuangan Nomor 203/PMK.05/2020 tentang Tata Cara Pembayaran dan Pertanggungjawaban Anggaran Penelitian Atas Beban Anggaran Pendapatan dan Belanja Negara;
14. Peraturan Menteri Riset, Teknologi dan Pendidikan Tinggi Republik Indonesia Nomor 69 tahun 2016 tentang Tata Cara Pembentukan Komite Penilaian dan/atau Reviewer Penelitian sebagaimana telah diubah dengan Peraturan Menteri Riset, Teknologi dan Pendidikan Tinggi Republik Indonesia Nomor 27 tahun 2019 tentang Perubahan atas Peraturan Menteri Riset, Teknologi dan Pendidikan Tinggi Republik Indonesia Nomor 69 tahun 2016 tentang Pedoman Pembentukan Komite Peneilaian dan/atau Reviewer dan Tata Cara Pelaksanaan Penilaian Penelitian dengan Menggunakan Standar Biaya Keluaran;
15. Peraturan Menteri Riset, Teknologi dan Pendidikan Tinggi Republik Indonesia Nomor 20 tahun 2018 tentang Penelitian;
16. Peraturan Menteri Riset, Teknologi dan Pendidikan Tinggi Republik Indonesia Nomor 12 tahun 2019 tentang Bantuan Operasional Perguruan Tinggi Negeri;
17. Peraturan Menteri Riset, Teknologi dan Pendidikan Tinggi Nomor 38 Tahun 2019 tentang Prioritas Riset Nasional Tahun 2020-2024;
18. Keputusan Menteri Riset, Teknologi dan Pendidikan Tinggi Nomor 105/M/KPT/2019 tentang Penggunaan Bantuan Operasional Perguruan Tinggi Negeri Penelitian dan Pengabdian kepada Masyarakat Tahun 2019;
19. Keputusan Menteri Riset dan Teknologi/ Kepala Badan Riset dan Inovasi Nasional Nomor 2/M/KPT/2021 tentang Pejabat Perbendaharaan pada Satuan Kerja Deputi Bidang Penguatan Riset dan Pengembangan Kementerian Riset dan Teknologi/ Badan Riset dan Inovasi Nasional;
20. Keputusan Kuasa Pengguna Anggaran Deputi Bidang Penguatan Riset dan Pengembangan Kementerian Riset, Teknologi, dan Pendidikan Tinggi Nomor 1/E1/KPT/2021 tentang tentang Pejabat Perbendaharaan pada Deputi Bidang Penguatan Riset dan Pengembangan Kementerian Riset dan Teknologi / Badan Riset dan Inovasi Nasional Tahun Anggaran 2021;
21. Keputusan Kuasa Pengguna Anggaran Deputi Bidang Penguatan Riset dan Pengembangan Kementerian Riset dan Teknologi/ Badan Riset dan Inovasi Nasional Nomor 9/E1/KPT/ 2021 tentang Penetapan Pendanaan Penelitian Terapan di Perguruan Tinggi Tahun Anggaran 2021;
22. Kontrak Penelitian Tahun Anggaran 2021 antara Direktorat Riset dan Pengabdian Masyarakat dengan Universitas Mualwarman Nomor 123/SP2H/LT/DRPM/2021 tanggal 18 Maret 2021.

PASAL 2
RUANG LINGKUP

Ruang Lingkup **Kontrak Penelitian** ini meliputi Pelaksanaan Penelitian Dasar dengan judul "**Pengembangan Model Nonparametric Geographically Weighted Timeseries Regression dan Sistem Aplikasi Komputasi untuk Deteksi Penyebaran Covid-19 dalam Konteks Geospasial dan Waktu**" dibebankan pada DIPA (Daftar Isian Pelaksanaan Anggaran) Deputi Bidang Penguatan Riset dan Pengembangan Kementerian Riset dan Teknologi/Badan Riset dan Inovasi Nasional Tahun 2021.

PASAL 3
JANGKA WAKTU

(1) Penelitian dengan judul sebagaimana dimaksud pada Pasal 2 dilaksanakan dalam jangka waktu **2 (Dua) tahun** yang mulai berlaku sejak tahun 2021;

(2) Keberlanjutan penelitian sebagaimana tercantum dalam Pasal 2 ditentukan berdasarkan hasil penilaian atas capaian tahun berjalan yang dilakukan oleh Komite Penilaian Keluaran Penelitian dan/atau Reviewer Keluaran Penelitian.

PASAL 4
HAK DAN KEWAJIBAN

(1) **PIHAK PERTAMA** mempunyai kewajiban:
  a. Memberikan pendanaan penelitian kepada **PIHAK KEDUA**
  b. Melakukan pemantauan dan evaluasi
  c. Mengkoordinir dan bertanggungjawab atas terlaksananya kontrak penelitian yang dilakukan oleh para peneliti
  d. Memantau pengunggahan ke laman SIMLITABMAS dokumen sebagai berikut :
    1. Revisi proposal penelitian
    2. Surat pernyataan kesanggupan penyusunan laporan penelitian;
    3. Catatan harian pelaksanaan penelitian
    4. Laporan kemajuan pelaksanaan penelitian
    5. Surat Pernyataan Tanggungjawab Belanja (SPTB) atas dana yang telah ditetapkan
    6. Laporan akhir penelitian
    7. Luaran penelitian

  paling lambat tanggal 16 November tiap tahun Anggaran berjalan.

(2) **PIHAK PERTAMA** mempunyai hak menerima dokumen dalam bentuk ***soft copy*** hasil unggahan di laman SIMLITABMAS sebagai berikut:
  a. Revisi proposal penelitian
  b. Surat pernyataan kesanggupan penyusunan laporan penelitian;
  c. Catatan harian pelaksanaan penelitian
  d. Laporan kemajuan pelaksanaan penelitian
  e. Surat Pernyataan Tanggungjawab Belanja (SPTB) atas dana yang telah ditetapkan
  f. Laporan akhir penelitian
  g. Luaran penelitian
  h. Profil penelitian, abstrak dan poster sesuai format yang ditentukan.

(3) **PIHAK KEDUA** mempunyai kewajiban:
  a. Mengunggah ke laman SIMLITABMAS dokumen sebagai berikut :
    1. Revisi proposal penelitian
    2. Surat pernyataan kesanggupan penyusunan laporan penelitian;
    3. Catatan harian pelaksanaan penelitian
    4. Laporan kemajuan pelaksanaan penelitian
    5. Surat Pernyataan Tanggungjawab Belanja (SPTB) atas dana yang telah ditetapkan
    6. Laporan akhir penelitian

7. Luaran penelitian

b. Menyerahkan kepada **PIHAK PERTAMA** Profil penelitian, abstrak dan poster sesuai format yang ditentukan dalam bentuk *soft copy*
c. Bertanggungjawab sepenuhnya dalam penggunaan dana penelitian dengan membuat laporan penggunaan anggaran sesuai dana yang telah ditetapkan dan disimpan oleh masing-masing peneliti.
d. Menyerahkan rekapitulasi anggaran sebayak 2 (dua) rangkap disertai dengan fotocopy bukti setor pajak dan peralatan (aset) yang dihasilkan dari kegiatan penelitian kepada **PIHAK PERTAMA**.
e. Menyerahkan hasil penelitian kepada **PIHAK PERTAMA** melalui Berita Acara Serah Terima (BAST)

(4) **PIHAK KEDUA** mempunyai hak memperoleh dana penelitian dari **PIHAK PERTAMA**.

PASAL 5
TATA CARA PEMBAYARAN DANA PENELITIAN

(1) **PIHAK PERTAMA** memberikan pendanaan penelitian sebesar: **Rp 166.380.000,- (*Seratus Enam Puluh Enam Juta Tiga Ratus Delapan Puluh Ribu Rupiah*)** yang dibebankan kepada DIPA Deputi Bidang Penguatan Riset dan Pengembangan Kementerian Riset dan Teknologi/Badan Riset dan Inovasi Nasional.

(2) Pendanaan penelitian sebagaimana dimaksud pada ayat (1) dibayarkan oleh **PIHAK PERTAMA** kepada **PIHAK KEDUA** secara bertahap:

a. **Rp. 79.480.000,- (*Tujuh Puluh Sembilan Juta Empat Ratus Delapan Puluh Ribu Rupiah*)** untuk dana penelitian tahun pertama. **Rp. 0 (*Nol Rupiah*)** untuk luaran tambahan pada tahun pertama;
b. **Rp. 86.900.000,- (*Delapan Puluh Enam Juta Sembilan Ratus Ribu Rupiah*)** untuk dana penelitian tahun kedua. **Rp. 0 (*Nol Rupiah*)** untuk luaran tambahan pada tahun kedua;
c. **Rp. ,- ()** untuk dana penelitian. **Rp. 0 (*Nol Rupiah*)** untuk luaran tambahan pada tahun ketiga.

(3) Pendanaan penelitian sebagaimana dimaksud pada ayat (2) huruf a, diberikan dengan ketentuan apabila revisi proposal penelitian telah diunggah ke laman SIMLITABMAS.

(4) Pendanaan penelitian sebagaimana dimaksud pada ayat (2) huruf b dan c, diberikan berdasarkan hasil penilaian atas capaian tahun sebelumnya yang dilakukan oleh Komite Penilaian Keluaran Penelitian dan/atau Reviewer Keluaran Penelitian.

(5) Biaya luaran tambahan dibayarkan kepada **PIHAK KEDUA** pada bulan **Oktober** tiap tahun

(6) Apabila luaran tambahan dinyatakan tidak valid oleh Deputi Bidang Penguatan Riset dan Pengembangan Kementerian Riset dan Teknologi/Badan Riset dan Inovasi Nasional sebagaimana dimaksud pasal 4 ayat (1), maka dana luaran tambahan yang sudah diterima harus disetorkan kembali ke kas negara.

(7) Dana Penelitian sebagaimana dimaksud pada ayat (1) akan dibayarkan oleh **PIHAK PERTAMA** kepada **PIHAK KEDUA** ke rekening sebagai berikut:

Nama	:	SIFRIYANI
Nomor Rekening	:	0174275433
Nama Bank	:	BNI

(8) **PIHAK PERTAMA** tidak bertanggung jawab atas keterlambatan dan/atau tidak terbayarnya sejumlah dana sebagaimana dimaksud pada ayat (1) yang disebabkan karena kesalahan **PIHAK KEDUA** dalam menyampaikan data peneliti, nama bank, nomor rekening, dan persyaratan lainnya yang tidak sesuai dengan ketentuan.

PASAL 6
PENGGANTIAN KEANGGOTAAN

(1) Perubahan terhadap susunan tim pelaksana dan substansi penelitian dapat dibenarkan apabila telah mendapat persetujuan dari Direktur Riset dan Pengabdian Masyarakat Deputi Bidang Penguatan Riset dan Pengembangan.

(2) Apabila Ketua tim pelaksana penelitian tidak dapat menyelesaikan penelitian atau mengundurkan diri, maka **PIHAK KEDUA** wajib menunjuk pengganti Ketua Tim Pelaksana penelitian yang merupakan salah satu anggota tim setelah mendapat persetujuan dari Direktur Riset dan Pengabdian Masyarakat Deputi Bidang Penguatan Riset dan Pengembangan.

(3) Dalam hal tidak adanya pengganti ketua tim pelaksana penelitian sesuai dengan syarat ketentuan yang ada, maka penelitian dibatalkan dan dana dikembalikan ke Kas Negara.

PASAL 7
PAJAK

PIHAK KEDUA berkewajiban memungut dan menyetor pajak ke kantor pelayanan pajak setempat yang berkenaan dengan kewajiban pajak berupa:

1. pembelian barang dan jasa dikenai PPN sebesar 10% dan PPh 22 sebesar 1,5%;
2. pajak–pajak lain sesuai ketentuan

PASAL 8
KEKAYAAN INTELEKTUAL

(1) Hak Kekayaan Intelektual yang dihasilkan dari pelaksanaan penelitian diatur dan dikelola sesuai dengan peraturan dan perundang-undangan.

(2) Setiap publikasi, makalah, dan/atau ekspos dalam bentuk apapun yang berkaitan dengan hasil penelitian ini wajib mencantumkan **Direktorat Riset dan Pengabdian Masyarakat Deputi Bidang Penguatan Riset dan Pengembangan Kementerian Riset dan Teknologi/Badan Riset dan Inovasi** sebagai pemberi dana.

(3) Hasil penelitian berupa peralatan adalah milik negara dan dapat dihibahkan kepada institusi/lembaga melalui Berita Acara Serah Terima (BAST).

PASAL 9
KEADAAN MEMAKSA

(1) **PARA PIHAK** dibebaskan dari tanggung jawab atas keterlambatan atau kegagalan dalam memenuhi kewajiban yang dimaksud dalam **Kontrak Penelitian** disebabkan atau diakibatkan oleh peristiwa atau kejadian diluar kekuasaan **PARA PIHAK** yang dapat digolongkan sebagai keadaan memaksa (*force majeure*).

(2) Peristiwa atau kejadian yang dapat digolongkan keadaan memaksa (*force majeure*) dalam **Kontrak Penelitian** ini adalah bencana alam, wabah penyakit, kebakaran, perang, blokade, peledakan, sabotase, revolusi, pemberontakan, huru-hara, serta

adanya tindakan pemerintah dalam bidang ekonomi dan moneter yang secara nyata berpengaruh terhadap pelaksanaan **Kontrak Penelitian** ini.

(3) Apabila terjadi keadaan memaksa (*force majeure*) maka pihak yang mengalami wajib memberitahukan kepada pihak lainnya secara tertulis, selambat-lambatnya dalam waktu 7 (tujuh) hari kerja sejak terjadinya keadaan memaksa (*force majeure*), disertai dengan bukti-bukti yang sah dari pihak yang berwajib, dan **PARA PIHAK** dengan itikad baik akan segera membicarakan penyelesaiannya.

PASAL 10
PENYELESAIAN PERSELISIHAN

(1) Apabila terjadi perselisihan antara **PIHAK PERTAMA** dan **PIHAK KEDUA** dalam pelaksanaan **Kontrak Penelitian** ini akan dilakukan penyelesaian secara musyawarah dan mufakat

(2) Dalam hal tidak tercapai penyelesaian secara musyawarah dan mufakat sebagaimana dimaksud pada ayat (1) maka penyelesaian dilakukan melalui proses hukum yang berlaku dengan memilih domisili hukum di Pengadilan Negeri Jakarta Pusat.

PASAL 11
AMANDEMEN KONTRAK

Apabila terdapat hal lain yang belum diatur atau terjadi perubahan dalam **Kontrak Penelitian** ini, maka akan dilakukan **Amandemen Kontrak Penelitian**.

PASAL 12
SANKSI

(1) Apabila sampai dengan batas waktu yang telah ditetapkan untuk melaksanakan Kontrak Penelitian telah berakhir, **PIHAK KEDUA** tidak melaksanakan kewajiban sebagaimana dimaksud dalam Pasal 4 ayat (2), maka **PIHAK KEDUA** dikenai sanksi administratif.

(2) Sanksi administratif sebagaimana dimaksud pada ayat (1) dapat berupa penghentian pembayaran dan tidak dapat mengajukan proposal penelitian dalam kurun waktu dua tahun berturut-turut.

PASAL 13
LAIN-LAIN

Dalam hal **PIHAK KEDUA** berhenti dari jabatannya sebelum **Kontrak Penelitian** ini selesai, maka **PIHAK KEDUA** wajib melakukan serah terima tanggung jawabnya kepada pejabat baru yang menggantikannya.

PASAL 14
PENUTUP

Surat Perjanjian Kontrak Penelitian ini dibuat rangkap 3 (tiga) bermaterai cukup sesuai dengan ketentuan yang berlaku, dan biaya materai dibebankan kepada **PIHAK KEDUA.**

PIHAK PERTAMA	PIHAK KEDUA
ANTON RAHMADI	**SIFRIYANI**
NIDN. 0001048001	**NIDN: 0023118203**

Pengisian poin C sampai dengan poin H mengikuti template berikut dan tidak dibatasi jumlah kata atau halaman namun disarankan seringkas mungkin. Dilarang menghapus/memodifikasi template ataupun menghapus penjelasan di setiap poin.

C. **HASIL PELAKSANAAN PENELITIAN:** Tuliskan secara ringkas hasil pelaksanaan penelitian yang telah dicapai sesuai tahun pelaksanaan penelitian. Penyajian meliputi data, hasil analisis, dan capaian luaran (wajib dan atau tambahan). Seluruh hasil atau capaian yang dilaporkan harus berkaitan dengan tahapan pelaksanaan penelitian sebagaimana direncanakan pada proposal. Penyajian data dapat berupa gambar, tabel, grafik, dan sejenisnya, serta analisis didukung dengan sumber pustaka primer yang relevan dan terkini.

Hasil Penelitian:

1. Inovasi metode statistika di bidang **Statistika Spasial** membangun model Nonparametric-Geographically Weighted Time series Regression (NGWTR) dan Implementasi model NGWTR di aplikasikan pada Data Kasus Angka COVID-19. Terdapat 2 basis penting yang harus diselesaikan yakni nonparametrik-spasial dan Timeseries-spasial.
2. Hasil Penelitian pertama, penulis berhasil mendapatkan Pengembangan model nonparametrik-spasial yakni Nonparametric-Geographically Weighted Spline Regression (NGWSR) dengan improve GCV pada estimasi model. Implementasi model NGWSR di aplikasikan pada Data Kasus Angka COVID-19 untuk Skala Provinsi di Indonesia.
3. Hasil Penelitian dilanjutkan basis Timeseries-spasial yakni model Geographically Temporally Weighted Regression (GTWR) dengan inovasi improve fungsi jarak Interaksi antara Spasial dan temporal. Implementasi model GTWR diaplikasikan pada data jumlah kasus positif COVID-19 di skala 56 Kabupaten/Kota di Pulau Kalimantan Indonesia pada periode Januari 2020 hingga Agustus 2021.
4. Penelitian selanjutnya kombinasi antara model model Nonparametric-Geographically Weighted Spline Regression dan Geographically Temporally Weighted Regression.

Capaian Luaran :

1. Artikel Pertama yang berjudul "**A Geospatial Analysis Using The Development Nonparametric Geographically Weighted Spline Regression and Its Application for COVID-19 Modeling**" pada Journal International **Spatial and Spatio-temporal Epidemiology,** Publisher Elsevier Ltd. Jurnal Q1 SJR 0,73. Status Review (Terlampir)

   Penelitian ini tentang pengembangan model Nonparametric-Geographically Weighted Spline Regression (NGWSR) yang diaplikasikan untuk deteksi penyebaran COVID-19 dalam konteks spasial. kajian epidemiologi spasial yang membahas bagaimana pengaruh bidang kesehatan, kependudukan dan ekonomi mempengaruhi akumulatif penderita COVID -19. Penelitian dengan menggunakan data Covid-19 skala nasional di 34 provinsi di Indonesia, data dari Bulan Desember 2019 hingga Bulan April Tahun 2021. Memberikan metodologi baru dibidang Sistem Informasi Geografis (SIG) dan statistika spasial. Kebaharuan dalam penelitian ini, digunakan pengembangan model NGWSR dengan optimasi GCV-NGWSR yang menggunakan dua basis matriks polynomial dan truncated. Selanjutnya diperoleh Uji Signifikansi Simultan dan Parsial yang mendukung model NGWSR. Hasil penelitian diperoleh 6 kelompok dari 34 Provinsi berdasarkan factor-faktor yang mempengaruhi akumulatif COVID -19. dihasilkan Pemetaan sebaran COVID -19 berdasarkan Jumlah dokter dan atau tenaga kesehatan, jumlah objek wisata, angka kasus TBC, persentase penduduk lansia, persentase rumah tangga yang memiliki akses air minum layak, kepadatan penduduk, Produk Domestik Regional Bruto (PDRB) yang mendukung sektor perekonomian di suatu wilayah, jumlah rumah sakit, jumlah desa/kelurahan yang memiliki puskesmas. Penelitian ini dapat memberikan konstribusi di bidang kesehatan dan pemerintah untuk dapat melakukan tracing berdasarkan hasil pemetaan spasial penyebaran COVID -19.

2. Artikel Kedua berjudul "**Tuberculosis Cases and Elderly Population Percentage Affect The Increase in COVID-19 Outbreak: Spatial-Temporal analysis of Geographically and Temporally Weighted Regression**" pada Journal International **Spatial Statistics,** Publisher Elsevier Ltd. Jurnal Q1 SJR 0,89. Status Under Review (Terlampir)

   Penelitian Spatio Temporal dengan menggunakan model Geographically Temporally Weighted Regression (GTWR) dengan pengembangan fungsi jarak Interaksi antara Spasial dan temporal.

Implementasi model GTWR diaplikasikan pada data jumlah kasus positif COVID-19 di skala 56 Kabupaten/Kota di Pulau Kalimantan Negara Indonesia pada periode Januari 2020 hingga Agustus 2021. Memberikan metodologi baru di bidang Statistika Spasial. Tujuan Penelitian adalah menentukan faktor-faktor yang mempengaruhi peningkatan akumulatif COVID-19 di Kalimantan dan dilanjutkan Pemetaan distribusi spasial untuk 56 Kabupaten/Kota berdasarkan variabel prediktor yang signifikan. Novelty penelitian adalah pengembangan model GTWR pada fungsi jarak dan implementasi model GTWR pada data COVID-19 di Kalimantan, Indonesia. Variabel penelitian yang digunakan adalah Akumulatif kasus COVID-19 di Kalimantan, Jumlah Dokter, Angka Kasus TBC, Presentase Penduduk Lansia, Kepadatan Penduduk, Produk Domestik Regional Bruto, Jumlah Rumah Sakit, Jumlah Puskesmas, Presentase Penduduk Miskin. Tahapan penelitian dimulai Statistika Deskriptif dengan pemetaan distribusi spasial untuk variabel penelitian, Estimasi model GTWR, Nilai Fungsi Jarak, Nilai Fungsi Pembobot Geografis, Uji signifikansi parameter model GTWR secara Simultan dan Parsial, Ukuran Kebaikan Model dan pemetaan berdasarkan signifikansi parameter model GTWR. Hasil Penelitian Model GTWR dengan Pengembangan fungsi jarak interaksi antara Spasial dan temporal menggunakan fungsi kernel Gaussian dengan Fixed bandwidth. faktor-faktor yang mempengaruhi kasus positif COVID-19 di Kalimantan untuk masing-masing Kabupaten/Kota berbeda, variable yang signifikan diantaranya Jumlah Dokter, Angka Kasus TBC, Presentase Penduduk Lansia, PDRB, dan Jumlah Rumah Sakit. Pemetaan berdasarkan variable yang signifikan ditampilkan pada Gambar 12. Kebaikan model GTWR dengan nilai koefisien determinasi $R2$ = 0,957, Adjusted R2 = 0,928, Akaike Information Criterion AIC = 1900,76 dan Root Mean Square Error RMSE = 1302,99. Hasil penelitian dapat dijadikan dijadikan rekomendasi buat pengambilan keputusan untuk pemerintah daerah dalam mengatasi masalah COVID-19 di daerah masing-masing.

3. Buku Ajar. Judul Buku **Statistika Spasial : Geographically Weighted Model,** No ISBN 978-623-7480-55-6.

D. **STATUS LUARAN**: Tuliskan jenis, identitas dan status ketercapaian setiap luaran wajib dan luaran tambahan (jika ada) yang dijanjikan. Jenis luaran dapat berupa publikasi, perolehan kekayaan intelektual, hasil pengujian atau luaran lainnya yang telah dijanjikan pada proposal. Uraian status luaran harus didukung dengan bukti kemajuan ketercapaian luaran sesuai dengan luaran yang dijanjikan. Lengkapi isian jenis luaran yang dijanjikan serta mengunggah bukti dokumen ketercapaian luaran wajib dan luaran tambahan melalui Simlitabmas.

LUARAN PENELITIAN

1. Artikel Pertama yang berjudul "**A Geospatial Analysis Using The Development Nonparametric Geographically Weighted Spline Regression and Its Application for COVID-19 Modeling**" pada Journal International **Spatial and Spatio-temporal Epidemiology,** Publisher Elsevier Ltd. Jurnal Q1 SJR 0,73.
   Status Under Review (Terlampir)

2. Artikel Kedua berjudul "**Tuberculosis Cases and Elderly Population Percentage Affect The Increase in COVID-19 Outbreak: Spatial-Temporal analysis of Geographically and Temporally Weighted Regression,** Publisher Elsevier Ltd. Jurnal Q1 SJR 0,89.
   Status Under Review (Terlampir)

3. Buku Ajar. Judul Buku **Statistika Spasial : Geographically Weighted Model,** No ISBN 978-623-7480-55-6.

E**. PERAN MITRA:** Tuliskan realisasi kerjasama dan kontribusi Mitra baik *in-kind* maupun *in-cash* (untuk Penelitian Terapan, Penelitian Pengembangan, PTUPT, PPUPT serta KRUPT). Bukti pendukung realisasi kerjasama dan realisasi kontribusi mitra dilaporkan sesuai dengan kondisi yang sebenarnya. Bukti dokumen realisasi kerjasama dengan Mitra diunggah melalui Simlitabmas.

Penelitian Dasar

F. **KENDALA PELAKSANAAN PENELITIAN**: Tuliskan kesulitan atau hambatan yang dihadapi selama melakukan penelitian dan mencapai luaran yang dijanjikan, termasuk penjelasan jika pelaksanaan penelitian dan luaran penelitian tidak sesuai dengan yang direncanakan atau dijanjikan.

1. Proses publikasi Jurnal yang membutuhkan waktu sekurang-kurangnya 6 Bulan.
2. Membutuhkan penelitian lanjutan dan diberikan waktu serta kesempatan untuk penelitian lanjutan tahun 2022 agar penelitian dapat diselesaikan dan Luaran Penelitian Tercapai.

G**. RENCANA TAHAPAN SELANJUTNYA:** Tuliskan dan uraikan rencana penelitian di tahun berikutnya berdasarkan indikator luaran yang telah dicapai, rencana realisasi luaran wajib yang dijanjikan dan tambahan (jika ada) di tahun berikutnya serta *roadmap* penelitian keseluruhan. Pada bagian ini diperbolehkan untuk melengkapi penjelasan dari setiap tahapan dalam metoda yang akan direncanakan termasuk jadwal berkaitan dengan strategi untuk mencapai luaran seperti yang telah dijanjikan dalam proposal. Jika diperlukan, penjelasan dapat juga dilengkapi dengan gambar, tabel, diagram, serta pustaka yang relevan. Jika laporan kemajuan merupakan laporan pelaksanaan tahun terakhir, pada bagian ini dapat dituliskan rencana penyelesaian target yang belum tercapai.

Rencana Tahapan Selanjutnya :

Penelitian selanjutnya kombinasi antara model model Nonparametric-Geographically Weighted Spline Regression dan Geographically Temporally Weighted Regression. Tujuan Inovasi di bidang Statistika Spatial yang menggabungkan regresi nonparametrik dan Spatial yang di pengaruhi oleh waktu. konstribusi dibidang Sistem Informasi Geografis (SIG). Dapat diaplikasikan pada berbagai fenomena alam yang terjadi. tidak hanya persoalan penataan kota dan perencanaan wilayah, tetapi juga dapat memetakan sebaran penyakit karena infeksi virus seperti COVID-19.

Target Luaran Wajib:

- Jurnal International Bereputasi Terindeks Scopus pada Journal International **Indonesian Journal of Science and Technology,** SJR 0,57.
  Isi Artikel terkait : Estimasi Model Nonparametric-Geographically Timeseries Regression dan Aplikasi di Bidang Kesehatan. Novelty Inovasi di Bidang Statistika Spasial.

Target Luaran Tambahan:

- Jurnal Nasional Terakreditasi SINTA 2, Nama Jurnal **Media Statistika**, ISSN: 24770647 | PISSN : 24770647
  Isi artikel tentang: Penerapan Model Nonparametric-Geographically Timeseries Regression dengan Pembobot Geografis Fungsi Kernel Gaussian
- **Proseding International AIP/IOP**
  Isi artikel tentang: Pemetaan GIS Berdasarkan Model Nonparametric-Geographically Timeseries Regression dengan Pembobot Geografis Eksponensial.

State of The Art

Salah satu sifat epidemi yang paling penting adalah penyebaran spasialnya, “suatu karakteristik yang terutama bergantung pada mekanisme epidemi, mobilitas manusia, dan strategi pengendalian” [1]. Kita dapat menggunakan GIS dan statistika spasial untuk menanggapi hal ini, dan juga untuk membantu mengurangi epidemi melalui informasi ilmiah, menemukan korelasi spasial dengan variabel lain, dan mengidentifikasi dinamika transmisi [2]. Salah satu karya pertama, serta salah satu yang paling banyak dikutip, yang menggunakan GIS untuk analisis spasial COVID-19 adalah karya Guan et al. [19]. Mereka mengekstrak data 1099 pasien dengan kasus COVID-19 yang dikonfirmasi laboratorium dari rumah sakit China hingga 29 Januari 2020, dan melanjutkan untuk mengkarakterisasi profil pasien rata-rata: usia rata-rata, jenis kelamin, gejala dan karakteristik *Spatiotemporal* yaitu identifikasi penyebaran penyakit yang cepat di seluruh daratan Cina, distribusi pasien berdasarkan provinsi, karakteristik antara penduduk Wuhan dan bukan penduduk, riwayat kontak langsung dengan satwa liar dan bukan penduduk Wuhan yang mengunjungi kota atau yang memiliki kontak dengan warga. Pekerjaan ini penting karena tanggal awal publikasi memungkinkan karakteristik klinis dari pasien yang terkena untuk lebih tepat didefinisikan. Juga di Cina, pada waktu yang hampir bersamaan, Chen et al. [3], menggunakan model spasial-temporal Bayesian, menentukan distribusi kasus COVID-19 dan korelasinya dengan migrasi penduduk Wuhan pada tahap awal epidemi, yang sangat penting untuk peringatan dini dan pencegahan masa depan. wabah. Huang et al. [5], berminggu-minggu kemudian COVID-19 bergerak sangat cepat, sementara timbulnya gejala tertunda rata-rata dua minggu, memungkinkan banyak hal terjadi dalam waktu singkat, menganalisis karakteristik epidemiologis COVID-19, tindakan pengendalian diambil, efeknya terhadap pandemi, dan distribusi *Spatiotemporal*nya. Analisis geografis pertama ini juga memetakan informasi ke tingkat administrasi yang lebih tepat yaitu kabupaten dan provinsi, hal ini dikerjakan untuk negara lain seperti Amerika Serikat [6], Iran [7], kemudian Korea Selatan [8], Brasil [9], Israel [10], Italia [11], Spanyol [12]. Tabel 1 memberikan informasi pengelompokan studi yang ditinjau berdasarkan tema geospasial, negara atau wilayah yang diteliti, asal informasi yang digunakan dan deskripsi singkat.

Tabel 1. *State of The Art* Penelitian *Spatial Statistics* pada data COVID-19

Peneliti	Tahun	*Region or Country of Study*	Data Yang Digunakan	Review Penelitian
Guan et al. [3]	2020	Cina	Kasus terkonfirmasi COVID-19	*Geographical* pada kasus yang terinfeksi COVID-19
Chen et al. [4]	2020	Cina	Kasus terkonfirmasi COVID-19	Distribusi penularan kasus COVID-19 dan korelasi imigrasi dari populasi Wuhan di tahap awal epidemi
Chen et al [4] dan Huang et al. [5]	2020	Cina	Kasus terkonfirmasi COVID-19	*Spatiotemporal* analysis of COVID-19 and its relationship with epidemiological characteristics, control of measures taken and their effects
Arab Mazar et al. [7]	2020	Iran	Kasus terkonfirmasi COVID-19	Analisis *Spatiotemporal* dari COVID-19 di tingkat nasional dan provinsi
Giuliani et al. [11]	2020	Italia	Kasus terkonfirmasi COVID-19	Analisis *Spatiotemporal* dari COVID-19 di tingkat nasional dan provinsi
Zhou et al. [13]	2020	Cina	Data Mining dan Kasus terkonfirmasi COVID-19	Reflections on the use of GIS with big data and *Spatiotemporal* analysis of COVID-19
Rezaei et al . [8]	2020	Korea Selatan	Kasus terkonfirmasi COVID-19	Analisis *Spatiotemporal* dari COVID-19 di tingkat nasional dan provinsi
Zhang et al. [14]	2020	Cina	Kasus terkonfirmasi COVID-19	Comparison of *Spatiotemporal* evolution between COVID-19 and SARS 2003
Dangnino et al. [15]	2020	Brazil	Kasus terkonfirmasi COVID-19	Analisis *Spatiotemporal* pada data COVID-19
Ahmadi et al. [16] dan Ahmadi et al. [17]	2020	Iran	Kasus terkonfirmasi COVID-19	Analisis *Spatiotemporal* dan GIS
Xiong et al. [2]	2020	Cina	Kasus terkonfirmasi COVID-19	Metode korelasi Pearson untuk analisis *Spatiotemporal*
Kearns et al. [18]	2020	Cina	Kasus terkonfirmasi COVID-19	Levy's flight to explain the *Spatiotemporal* dynamics of the pandemic
Desjardin et al. [19]	2020	Amerika Serikat	Kasus terkonfirmasi COVID-19	Prospective space-time statistics to identify active and emerging COVID-19 groups at the county level
Rossman et al. [20]	2020	Israel	Jajak pendapat/ Polls dan Kasus terkonfirmasi COVID-19	Kuesioner online yang ditujukan untuk identifikasi kemungkinan gejala secara geographical
Ángel Sola et al. [21]	2020	Cekungan K aribia	Kasus terkonfirmasi COVID-19	Memprediksi penyebaran COVID-19 berdasarkan data geografis dan iklim
Tim CC-R. [22]	2020	Amerika Serikat	Kasus terkonfirmasi COVID-19	karakteristik Geografis dan analisis *Spatiotemporal* dari infeksi COVID-19
Orea dan Alvarez. [23]	2020	Spain	Kasus terkonfirmasi COVID-19	Analysis by provinces of the effectiveness of quarantine on the spread of the pandemic
Murugesan et al. [24]	2020	India	Kasus terkonfirmasi COVID-19	Analisis *Spatiotemporal* COVID-19 tingkat nasional dan provinsi

Tang et al. [25]	2020	Cina	Kasus terkonfirmasi COVID-19	Poisson segmented model for the analysis of changing patterns in different geographic areas
Kayu dan al. [26]	2020	Brazil	Kasus terkonfirmasi COVID-19	Analisis *Spatiotemporal* COVID-19 di Estado da Bahia
Buzai [27]	2020	Argentina	Kasus terkonfirmasi COVID-19	Analisis *Spatiotemporal* dan refleksi pada geografi kesehatan
Santana Juarez [28]	2020	Meksiko	Kasus terkonfirmasi COVID-19	Analisis *Spatiotemporal* dari COVID-19 tingkat nasional dan provinsi
Saha et al. [29]	2020	Dunia	Kasus terkonfirmasi COVID-19	Analisis *Spatiotemporal* dan refleksi pada kegunaan GIS di dalam pandemi
Penelitian yang akan dilaksanakan	2022-2023	Indonesia	Kasus terkonfirmasi COVID-19	Pengembangan model *Nonparametric-Spatiotemporal* : *Nonparametric Geographically Temporally Weighted Regression* Dengan Fungsi Jarak *Improved Spatial-timeseries* Dan Aplikasi Pemodelan Covid-19

Road Map Penelitian

Keterangan/ Waktu	IKU	Tahun 2018 - 2020	Tahun 2021-2023	Tahun 2024 -2026	Tahun 2027 - 2030
Riset Pengembangan	IKU 9 IKU 8 IKU 7	**Industri** →			**2027 - 2030 Integrasi Sistem yang diterapkan pada Bidang Industri, Kesehatan dan Lingkungan dan Bidang lain (Big Data)**
Riset Terapan	IKU 6 IKU 5 IKU 4	**Tekhnologi** →		**2024 - 2026 Membangun Sistem Komputasi Deteksi Penyakit berbasis Lingkungan/ Spasial dan Runtun Waktu** **(Pengembangan Tekhnologi Bidang Komputasi untuk Informasi Geospasial)**	
Riset Dasar yang diusulkan	IKU 3 IKU 2 IKU 1	**Riset Dasar** →	2021-2023 Membangun Model NGTWR, Estimasi model, Komputasi Model dan Aplikasinya untuk kajian data Covid-19 Model Nonparametric GTWR Extended, estimasi komputasi dan aplikasinya Kajian Asimtotikteoritik dan revalidasi model serta Aplikasinya		
Riset yang pernah dilakukan	**Spatial: GW Model**	➢ Membangun Model *Geographically Weighted Rehression* (NGWR) dengan Pendekatan *Spline Truncated* [30], [31] ➢ Metode Baru untuk Uji Hipotesis untuk *Truncated Spline Nonparametric Regression* dipengaruhi *Spatial Heterogeneity* dan Aplikasinya [32] ➢ Simultaneous Hypothesis Testing of Multivariable Nonparametric Spline Regression in the GWR Model [33] ➢ Evaluation of Factors Affecting Increased Unemployment in East Java using NGWR-TS Method [34] ➢ Evaluation of Forest Productivity and Governance on the Preservation of Tropical Rain Forests in Kalimantan Using The NGWR-TS Nonparametric Geospatial Method [35], [36].			
	COVID -19	Pemodelan Epidemiologi SIR Data Covid-19 di Kalimantan TImur [37]			
Penelitian yang Pernah dilakukan			**Penelitian yang akan diusulkan**	**Penelitian Selanjutnya**	

Hasil Penelitian Tahun 2021:

1. A Geospatial Analysis Using The Development Nonparametric Geographically Weighted Spline Regression and Its Application for COVID-19 Modeling
2. Tuberculosis Cases and Elderly Population Percentage Affect The Increase in COVID-19 Outbreak: Spatial-Temporal analysis of Geographically and Temporally Weighted Regression

Penelitian Tahun ke-2 Tahun 2022

1. Jurnal International Bereputasi Terindeks Scopus pada Journal International **Indonesian Journal of Science and Technology,** SJR 0,57.

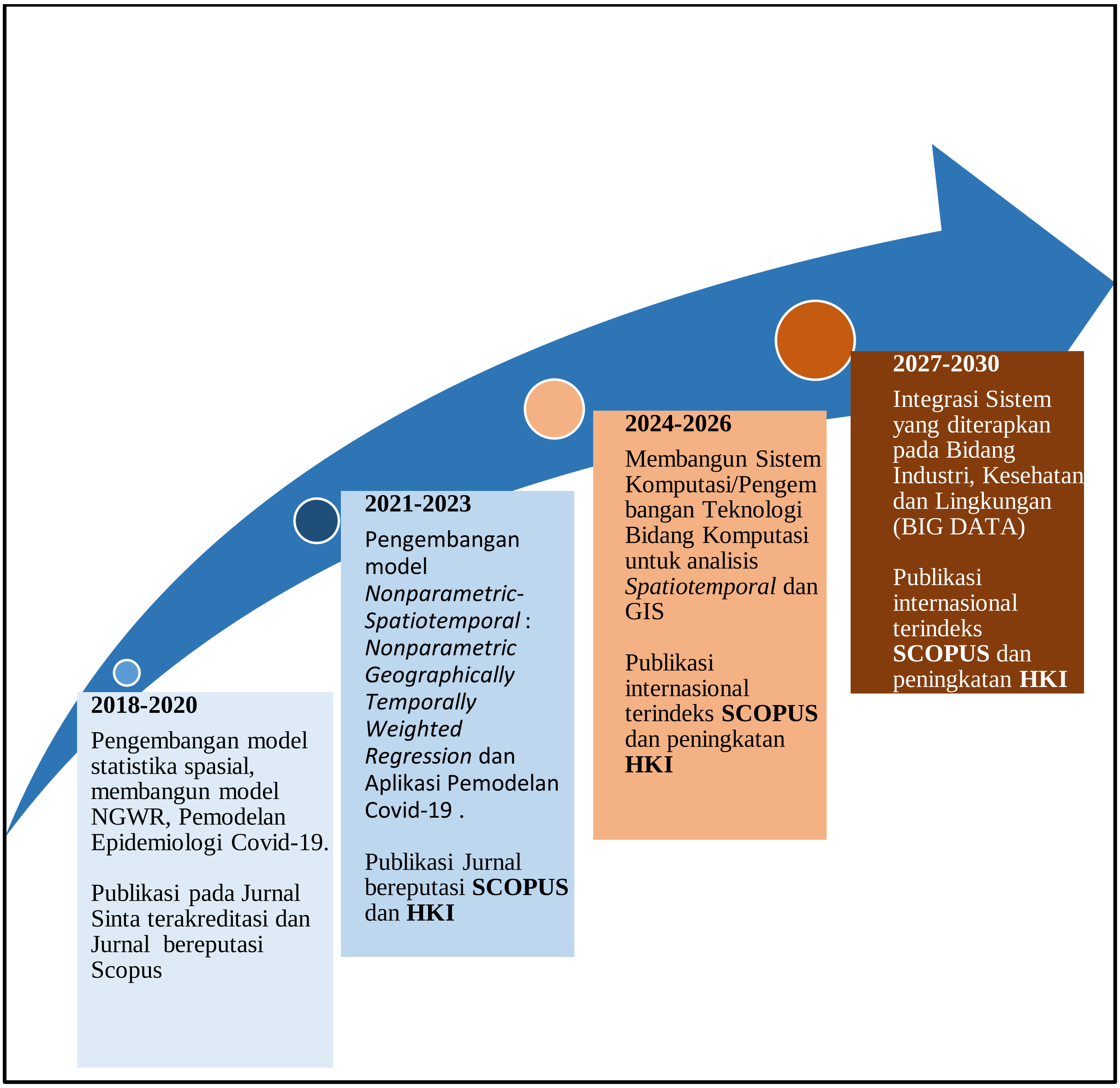
Isi Artikel terkait : Estimasi Model Nonparametric-Geographically Timeseries Regression dan Aplikasi di Bidang Kesehatan. Novelty Inovasi di Bidang Statistika Spasial.

2. Jurnal Nasional Terakreditasi SINTA 2, Nama Jurnal **Media Statistika**, ISSN: 24770647 | PISSN : 24770647

Isi artikel tentang: Penerapan Model Nonparametric-Geographically Timeseries Regression dengan Pembobot Geografis Fungsi Kernel Gaussian

3. **Proseding International AIP/IOP**

Isi artikel tentang: Pemetaan GIS Berdasarkan Model Nonparametric-Geographically Timeseries Regression dengan Pembobot Geografis Eksponensial.



Gambar 1. Road Map Penelitian

H. DAFTAR PUSTAKA: Penyusunan Daftar Pustaka berdasarkan sistem nomor sesuai dengan urutan pengutipan. Hanya pustaka yang disitasi pada laporan kemajuan yang dicantumkan dalam Daftar Pustaka.

[1] Gross, B., Zheng, Z., Liu, S., Chen, X., Sela, A., Li, J. Havlin, S., 2020. Spatio-temporal prop- agation of COVID-19 pandemics. medRxiv 2020.03.23.20041517.

[2] Xiong, Y., Guang, Y., Chen, F., Zhu, F., 2020. Spatial statistics and influencing factors of the novel coronavirus pneumonia 2019 epidemic in Hubei Province, China. ResearchSquare https://doi.org/10.21203/rs.3.rs-16858/v2 Preprint.

[3] Guan, W.J., Ni, Z.Y., Hu, Y., Liang, W.H., Ou, C.Q., He, J.X., Du, B., 2020. Clinical character- istics of coronavirus disease 2019 in China. New England Journal of Medicine 382 (18), 1708–1720

[4] Chen, Z.L., Zhang, Q., Lu, Y., Guo, Z.M., Zhang, X., Zhang, W.J. Lu, J.H., 2020b. Distribution of the COVID-19 epidemic and correlation with population emigration from wuhan, China. Chinese medical journal https://doi.org/10.1097/CM9.0000000000000782.

[5] Huang, H., Wang, Y., Wang, Z., Liang, Z., Qu, S., Ma, S., Liu, X., 2020. Epidemic Features and Control of 2019 Novel Coronavirus Pneumonia in Wenzhou, China China 3/3/ 2020).

[6] Dong, E., Du, H., Gardner, L., 2020. An interactive web-based dashboard to track COVID-19 in real time.

Lancet Infect. Dis. 20, 533–534.
[7] Arab-Mazar, Z., Sah, R., Rabaan, A.A., Dhama, K., Rodriguez-Morales, A.J., 2020. Mapping the incidence of the COVID-19 hotspot in Iran – implications for travellers. Travel Med. Infect. Dis. https://doi.org/10.1016/j.tmaid.2020.101630.
[8] Rezaei, M., Nouri, A.A., Park, G.S., Kim, D.H., 2020. Application of geographic information system in monitoring and detecting the COVID-19 outbreak. Iran. J. Public Health 49, 114–116.
[9] Dagnino, R., Weber, E.J., Panitz, L.M., 2020. Monitoramento do Coronavírus (Covid-19) nos municípios do Rio Grande do Sul, Brasil. SocArXiv https://doi.org/10.31235/osf. io/3uqn5.
[10] Rossman, H., Keshet, A., Shilo, S., Gavrieli, A., Bauman, T., Cohen, O. Segal, E., 2020. A framework for identifying regional outbreak and spread of COVID-19 from one- minute population-wide surveys. Nature Medicine 1–4.
[11] Giuliani, D., Dickson, M.M., Espa, G., Santi, F., 2020. Modelling and Predicting the Spatio- Temporal Spread of Coronavirus Disease 2019 (COVID-19) in Italy (3/20/2020). SSRN. https://doi.org/10.2139/ssrn.3559569.
[12] Orea, L., Álvarez, I.C., 2020. How Effective Has the Spanish Lockdown Been to Battle COVID-19? A Spatial Analysis of the Coronavirus Propagation across Provinces. Documento de Trabajo, 03.
[13] Zhou, C., Su, F., Pei, T., Zhang, A., Du, Y., Luo, B., Song, C., 2020. COVID-19: challenges to GIS with big data. Geography and Sustainability 1 (1), 77–87.
[14] Zhang, X., Rao, H.X., Wu, Y., Huang, Y., Dai, H., 2020. Comparison of the spatiotemporal characteristics of the COVID-19 and SARS outbreaks in mainland China. medRxiv2020.03.23.20034058.
[15] Dagnino, R., Weber, E.J., Panitz, L.M., 2020. Monitoramento do Coronavírus (Covid-19) nos municípios do Rio Grande do Sul, Brasil. SocArXiv https://doi.org/10.31235/osf. io/3uqn5 (March, 28).
[16] Ahmadi, A., Fadai, Y., Shirani, M., Rahmani, F., 2020a. Modeling and forecasting trend of COVID-19 epidemic in Iran until May 13, 2020. Medical Journal of The Islamic Repub- lic of Iran (MJIRI) 34 (1), 183–195.
[17] Ahmadi, M., Sharifi, A., Dorosti, S., Ghoushchi, S.J., Ghanbari, N., 2020b. Investigation of ef- fective climatology parameters on COVID-19 outbreak in Iran. Sci. Total Environ. 729, 138705 10 August 2020.
[18] Kearns, R., Moon, G., 2002. From medical to health geography: novelty, place and theory after a decade of change. Prog. Hum. Geogr. 26 (5), 605–625.
[19] Desjardins, M.R., Hohl, A., Delmelle, E.M., 2020. Rapid surveillance of COVID-19 in the United States using a prospective space-time scan statistic: detecting and evaluating emerging clusters. Appl. Geogr. 102202.
[20] Rossman, H., Keshet, A., Shilo, S., Gavrieli, A., Bauman, T., Cohen, O. Segal, E., 2020. A framework for identifying regional outbreak and spread of COVID-19 from one- minute population-wide surveys. Nature Medicine 1–4.
[21] Allcott, H., Boxell, L., Conway, J., Gentzkow, M., Thaler, M., Yang, D.Y., 2020. Polarization and public health: partisan differences in social distancing during the coronavirus pandemic. NBER Working Paper (w26946).
[22] Team CC-R, 2020. Geographic differences in COVID-19 cases, deaths, and incidence - United States, February 12-April 7, 2020. MMWR Morb. Mortal. Wkly Rep. 69 (15).
[23] Orea, L., Álvarez, I.C., 2020. How Effective Has the Spanish Lockdown Been to Battle COVID-19? A Spatial Analysis of the Coronavirus Propagation across Provinces. Documento de Trabajo, 03.
[24] Murugesan, B., Karuppannan, S., Mengistie, A.T., Ranganathan, M., Gopalakrishnan, G., 2020. Distribution and TrendAnalysis of COVID-19 in India: geospatial approach. Journal of Geographical Studies 4 (1), 1–9. https://doi.org/10.21523/gcj5.20040101.
[25] Tang, T., Huipeng, L., Gifty, M., Zaisheng, W., Weibin, C., Dan, W., Rongbin, Y., 2020. The changing patter of COVID-19 in China: a tempo-geographic analysis of the SARS- CoV-2 epidemic. Clin. Infect. Dis., ciaa423 https://doi.org/10.1093/cid/ciaa423.
[26] Kearns, R., Moon, G., 2002. From medical to health geography: novelty, place and theory after a decade of change. Prog. Hum. Geogr. 26 (5), 605–625.
[27] Buzai, G.D., 2020. De Wuhan a Luján. Evolución espacial del COVID-19. Posición 3, 2683–8915.
[28] Santana Juárez, M.V., 2020. COVID-19 en México: comportamiento espacio temporal y condicionantes socioespaciales, febrero y marzo de 2020. Posición 3, 2683–8915.
[29] Saha, A., Gupta, K., Patil, M., 2020. Monitoring and Epidemiological Trends of Coronavirus Disease (COVID-19) around the World. Osfpreprints.

[30] Sifriyani, S. H. Kartiko, I. N. Budiantara and Gunardi., 2018a, Development Of Nonparametric Geographically Weighted Regression Using Truncated Spline Approach, Songklanakarin Journal of Science and Technology, 40(4), (2018), 909-920.

[31] Sifriyani, S.H. Kartiko, I.N. Budiantara, and Gunardi., 2017, Geographically Weighted Regression with Spline Approach. *Far East Journal of Mathematical Sciences*, 101(6), 1183-1196.

[32] Sifriyani, I.N. Budiantara, S.H. Kartiko and Gunardi. 2018b. A New Method of Hypothesis Test for Truncated Spline Nonparametric Regression Influenced by Spatial Heterogeneity and Application. Abstract and Applied Analysis. 2018. https://doi.org/10.1155/2018/9769150

[33] Sifriyani. 2019. Simultaneous Hypothesis Testing of Multivariable Nonparametric Spline Regression in the GWR Model. International Journal of Statistics and Probability, Vol 8(4).

[34] Sifriyani, Budiantara, I. N., Kartiko, S.H., and Gunardi, 2019, Evaluation of Factors Affecting Increased Unemployment in East Java Using NGWR-TS Method, International Journal of Sciences: Basic and Applied Research. 49(1), (2019) 123-142

[35] Sifriyani, Ruslan, F.H. Susanty. "Evaluation of Forest Productivity and Governance on the Preservation of Tropical Rain Forests in Kalimantan Using The NGWR-TS Nonparametric Geospatial Method". EurAsian Journal of BioSciences, Vol 13 Issue 2, pp. 2373-2379, 2019.

[36] Sifriyani, Ruslan, F.H. Susanty. Mapping and Analysis Factors of Affecting Productivity Tropical Rain Forests in East Kalimantan, Modern Applied Science, Vol 13 No 10, 2019; DOI: 10.5539/mas.v13n10p112

[37] Sifriyani, Dedi Rosadi, "Pemodelan *Susceptible Infected Recovered* (SIR) untuk Estimasi Angka Reproduksi Covid-19 di Kalimantan Timur dan Samarinda", *Media Statistika*, Vol 13 (2), 2020

Statistika Spasial
Geographically Weighted Models

Buku Statistika Spasial dengan Judul Geographically Weighted Models merupakan salah satu buku yang wajib dimiliki oleh mahasiswa maupun peneliti pada bidang statistika khususnya yang ingin menguasai regresi spasial berbasis titik dan menggunakan analisis spasial pada penelitian.

Geographically Weighted Models adalah analisis regresi berbasis data titik yang merupakan bentuk lokal dari regresi klasik yang memperhatikan aspek spasial atau lokasi geografis yang berupa koordinat titik, nilai taksiran atau estimator parameter regresi yang diperoleh untuk setiap lokasi pengamatan akan berbeda-beda.

Kebutuhan akan eksplorasi dan analisis data dibidang spasial semakin meningkat, sehingga dibutuhkan pengetahuan tentang analisis spasial secara teori dan aplikasi komputasi. Buku ini secara komprehensif membahas teori dan program komputasi untuk mengimplementasikan model.

1. Pengenalan Data Spasial dan Efek Spasial
2. Autokorelasi Spasial dan Heterogenitas Spasial
3. Geographically Weighted Regression
4. Geographically Weighted Nonparametric Regression
5. Geographically Weighted Poisson Regression
6. Geographically Weighted Logistic Regression
7. Mix Geographically Weighted Regression
8. Regresi Spasial
9. Implementasi Model Geographically Weighted Nonparametric Regression

Biodata Penulis

Dr. Sifriyani, S.Pd., M.Si. adalah Dosen Program Studi Statistika Jurusan Matematika FMIPA Universitas Mulawarman. Pendidikan Penulis Sarjana di Jurusan Pendidikan Matematika Universitas Negeri Makassar, Pendidikan Magister di Jurusan Statistika Institut Teknologi Sepuluh Nopember, Pendidikan Doktor di Jurusan Matematika Bidang Keahlian Statistika di Universitas Gadjah Mada dan Kunjungan Penelitian *Spatial Data Science* di *Faculty of Science and Technology di Lancaster University United Kingdom.*

Dr. Sifriyani, M.Si

Statistika Spasial
Geographically Weighted

MULAWARMAN UNIVERSITY PRESS

Dr. Sifriyani, M.Si

Statistika Spasial

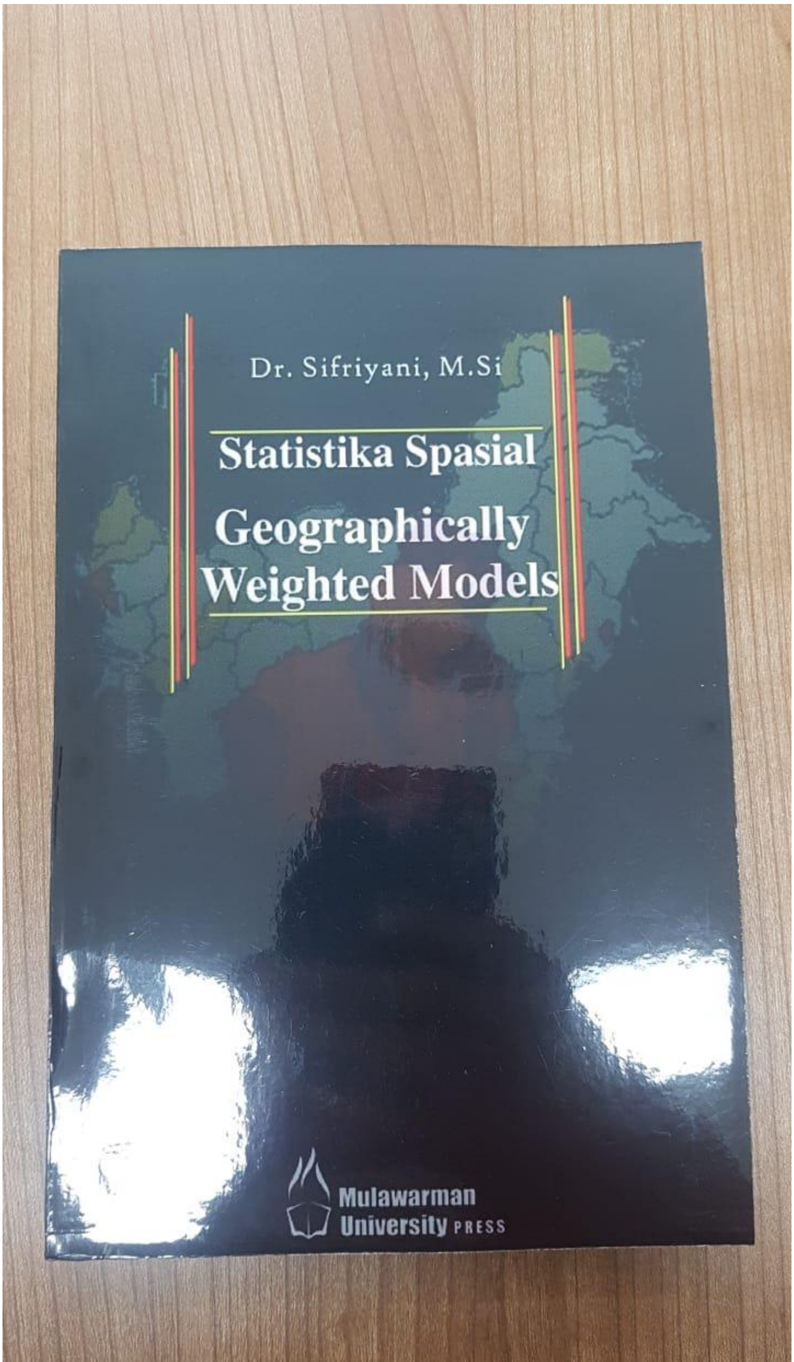
Geographically Weighted Models

MULAWARMAN UNIVERSITY PRESS

ISBN

ISBN 978-623-7480-55-6

9 786237 480556


Dr. Sifriyani, M.Si
Statistika Spasial
Geographically
Weighted Models
Mulawarman
University PRESS

Statistika Spasial
Geographically Weighted Models

Buku Statistika Spasial dengan Judul Geographically Weighted Models merupakan salah satu buku yang wajib dimiliki oleh mahasiswa maupun peneliti pada bidang statistika khususnya yang ingin menguasai regresi spasial berbasis titik dan menggunakan analisis spasial pada penelitian.

Geographically Weighted Models adalah analisis regresi berbasis data titik yang merupakan bentuk lokal dari regresi klasik yang memperhatikan aspek spasial atau lokasi geografis yang berupa koordinat titik, nilai taksiran atau estimator parameter regresi yang diperoleh untuk setiap lokasi pengamatan akan berbeda-beda.

Kebutuhan akan eksplorasi dan analisis data dibidang spasial semakin meningkat, sehingga dibutuhkan pengetahuan tentang analisis spasial secara teori dan aplikasi komputasi. Buku ini secara komprehensif membahas teori dan program komputasi untuk mengimplementasikan model.

1. Pengenalan Data Spasial dan Efek Spasial
2. Autokorelasi Spasial dan Heterogenitas Spasial
3. Geographically Weighted Regression
4. Geographically Weighted Nonparametric Regression
5. Geographically Weighted Poisson Regression
6. Geographically Weighted Logistic Regression
7. Mix Geographically Weighted Regression
8. Regresi Spasial
9. Implementasi Model Geographically Weighted Nonparametric Regression

Biodata Penulis

Dr. Sifriyani, S.Pd., M.Si. adalah Dosen Program Studi Statistika Jurusan Matematika FMIPA Universitas Mulawarman. Pendidikan Penulis Sarjana di Jurusan Pendidikan Matematika Universitas Negeri Makassar, Pendidikan Magister di Jurusan Statistika Institut Teknologi Sepuluh Nopember, Pendidikan Doktor di Jurusan Matematika Bidang Keahlian Statistika di Universitas Gadjah Mada dan Kunjungan Penelitian *Spatial Data Science di Faculty of Science and Technology di Lancaster University United Kingdom.*

Penerbit
Mulawarman University PRESS

ISBN 978-623-7480-55-6
9 786237 480556

Spatial and Spatio-temporal Epidemiology

A Geospatial Analysis Using The Development Nonparametric Geographically-Weighted Spline Regression and Its Application for COVID-19 Modeling

--Manuscript Draft--

Manuscript Number:	SSTE-D-21-00140
Article Type:	Research Paper
Keywords:	COVID-19 Geographically Nonparametric Spline Spatial GIS
Corresponding Author:	sifriyani sifriyani, M.Si Mulawarman University Samarinda, Kalimantan Timur INDONESIA
First Author:	sifriyani sifriyani, M.Si
Order of Authors:	sifriyani sifriyani, M.Si
	Dedi Rosadi, Professor
	Mariani Rasjid, Pulmonologist
Abstract:	Highlights A Geospatial Analysis Using The Development Nonparametric Geographically-Weighted Spline Regression and Its Application for COVID-19 Modeling
Suggested Reviewers:	

Title Page (with Author Details)

Highlights

A Geospatial Analysis Using The Development Nonparametric Geographically-Weighted Spline Regression and Its Application for COVID-19 Modeling

Sifriyani,Dedi Rosadi,Mariani Rasjid

- The novelty of this study lies in the development of a NGWSR model with GCV-NGWSR optimization using two polynomial and truncated matrix bases.
- The present study develops a new model in spatial statistics called the Nonparametric Geographically-Weighted Spline Regression model.
- Program syntax is created based on the developed NGWSR model to support the computing system so that it can be used to process data based on the NGWSR model.
- The study offers a new methodology in the fields of Geographic Information Systems (GIS) and spatial statistics.
- Implementation of the NGWSR Model using the nationwide COVID-19 data in 34 provinces in Indonesia.
- This study used the nationwide COVID-19 data of the 34 provinces from December 2019 to April 2021.

A Geospatial Analysis Using The Development Nonparametric Geographically-Weighted Spline Regression and Its Application for COVID-19 Modeling

Sifriyani[a,*], Dedi Rosadi[b] and Mariani Rasjid[c,d]

[a]Statistics Study Program, Department of Mathematics, Faculty of Mathematics and Natural Sciences, Mulawarman University, Samarinda, Indonesia

[b]Department of Mathematics, Faculty of Mathematics and Natural Sciences, Gadjah Mada University, Yogyakarta, Indonesia

[c]Pulmonologist, Tadulako General Hospital, Palu, Indonesia

[d]Faculty of Medicine, Tadulako University, Palu, Indonesia

ARTICLE INFO

Keywords:
COVID-19
Geographically
Nonparametric
Spline
Spatial
GIS

ABSTRACT

The present study deals with the development of a Nonparametric Geographically-Weighted Spline Regression (NGWSR) model which is applied to detect the spread of COVID-19 (Coronavirus disease 2019) in a spatial context. It is a spatial epidemiology discussing how the health, population and economic sectors affect the cumulative number of COVID-19 patients. The present study used COVID-19 nationwide data of 34 provinces in Indonesia from December 2019 to April 2021. The study offers a new methodology in the fields of Geographic Information Systems (GIS) and spatial statistics. The novelty of this study lies in the development of a NGWSR model with GCV-NGWSR optimization using two bases of polynomial and truncated matrices. Furthermore, the results of the simultaneous and partial significance tests supported the NGWSR model. The present study produced 6 groups from 34 provinces based on the factors affecting the cumulative number of COVID-19 patients. A map of COVID-19 spread was made on the basis of the total number of doctors and/or health professionals, the number of tourist attractions, the number of Tuberculosis (TB) cases, the percentage of the elderly population, the percentage of households with access to proper drinking water, population density, Gross Regional Domestic Product (GRDP) supporting the economic sector of a region, the number of hospitals, and the number of villages/urban villages having a public health center (puskesmas). The present study contributes to the health sector and the government with regard to tracing based on the spatial mapping results of the spread of COVID-19.

1. Introduction

The regression analysis used for spatial data modeling is known as the geographically-weighted regression (GWR) model. GWR was first introduced by Fotheringham in 1967. The response variable in the GWR model contains predictor variables, in which each of its regression coefficient depends on the location where the data are observed. GWR takes the point approach. Each parameter value is estimated at each observation location, so that each point of observation location has a different parameter value (Fotheringham, Brundson and Charlton, 2002). If each parameter has a constant value in each geographic location, the GWR model will be the same as the linear regression model. This means that each geographic location has the same model. Studies undertaken using the GWR theory include Brunsdon and Fotheringham (1999), Crespo, Fotheringham and Charlton (2007), Leung, Mei and Zhang (2000a), and Leung, Mei and Zhang (2000b).

Those statisticians who carried out studies on the GWR are including Nakaya, Fotheringham, Brunsdon and Charltron (2005) who developed the Geographically Weighted Poisson Regression Models, and Mei, Wang and Zhang (2006) who produced a model that combines the global regression model and GWR known as the Mixed Geographically-Weighted Regression models. The GWR model is also developed in the time-series-related spatial field known as spatio-temporal studies, in which exploratory spatio-temporal data are analyzed using the GWR and geo-visual analytics, as undertaken by Demsar, Fotheringham and Charlton (2008). The spatio-temporal study was also used by Huang, Wu and Barry (2010) in the Geographically- and Temporally-Weighted Regression Model. Furthermore,

ORCID(s): 0000-0002-4616-775X (Sifriyani); 0000-0003-2689-253X (D. Rosadi); 0000-0002-1724-3565 (M. Rasjid)

Yu (2010) studies spatial panel data using Geographically-Weighted Panel Regression and Wrenn and Sam (2014) in their study used Geographically- and Temporally-Weighted Likelihood Regression models.

The developed GWR model remains in a linear form. In some cases of regression analysis, the question is that whether all relationships between predictor variables and response variable will form a known regression curve with a linear pattern. In fact, indeed, not all data relationship patterns are linear and have a known regression curve. Given the rapid development of science and technology and natural phenomena leading to unusual patterns, it is very difficult to predict the behavior of nature. For example, in the past decade, one was able to predict exactly when the dry and rainy seasons began and ended in various geographic areas, making farmers capable of preparing themselves for planting rice and harvesting it. However, currently it is difficult to do so. Other problems, such as the percentage of poverty, underdevelopment, literacy rates, increasing ignorance and uneven development in each region and their causative variables constitute examples of events with unclear patterns, not following a certain pattern, having no linear pattern and as if irregular. Thus, it is necessary to develop nonparametric regressions in the GWR model to overcome and provide solutions to these problems.

A good regression model should be viewed from a variety of aspects and a modeling problem should be placed proportionally. Differences in environmental and geographical characteristics among observation locations result in observations with different variations or differences in the effect of predictor variables on response variable for each observation location. The GWR model is not able to resolve the problem of the effect of predictor variables on the response variable not following a specific pattern and varying patterns in certain sub-interval. Thus, it is highly necessary to develop studies in the field of nonparametric regression and spatial statistics. Previous studies developed nonparametric regression models in GWR (Sifriyani, Kartiko, Budiantara and Gunardi, 2018) and continued. The present study develops a new model in spatial statistics called the Nonparametric Geographically-Weighted Spline Regression model.

The present study would contribute to the Geographic Information Systems (GIS) and spatial statistics fields. It can be applied to a variety of natural phenomena. It does not only address the problems of city planning and regional planning, but also mapping the distribution of diseases in an area, one of which being viral infections, such as COVID-19, the outbreaks of which hitting numerous countries, including Indonesia. The COVID-19 situation in Indonesia has not subsided. The COVID-19 Task Force reported that the number of positive cases of COVID-19 on 26 June 2021 was 21,095 people. This figure again broke the highest daily record in Indonesia of 20,547 positive people for the COVID-19 virus (Satuan Gugus Tugas Penanganan COVID-19, 2021), (Sifriyani and Rosadi, 2020). Multiple causes lead to this increase. Therefore, the present study investigated the causes of the cumulative increase in COVID-19 patients in Indonesia based on factors related to the health, population and economic sectors using the Nonparametric Geographically-Weighted Spline Regression (NGWSR) model.

Research that studies the modeling of COVID-19 using spatial statistics is carried out by several researchers. Spatial statistical modeling research for confirmed cases of COVID-19 in China was conducted by Guan, Ni, Hu, Liang, Ou, He and Du (2020), Chen, Zhang, Lu, Guo, Zhang, Zhang and Lu (2020b), Huang, Wang, Wang, Liang, Qu, Ma and Liu (2020), Zhou, Su, Pei, Zhang, Du, Luo and Song (2020), Zhang, Rao, Wu, Huang and Dai (2020), and Xiong, Guang, Chen and Zhu (2020). A spatial statistical modeling study for confirmed cases of COVID-19 in the USA was conducted by Desjardins, Hohl and Delmelle (2020). The research Spatiotemporal analysis of COVID-19 at the national and provincial levels was conducted by Arab-Mazar, Sah, Rabaan, Dhama and Rodriguez-Morales (2020), Giuliani, Dickson, Espa and Santi (2020), Saha, Gupta and Patil (2020) and Buzai (2020). This research is about geospatial models and its application for cumulative number cases of COVID-19 in Indonesia.

The NGWSR model is applied to detect the spread of COVID-19 in a spatial context. A spatial epidemiological study deals with the effect of the health, population and economy sectors on the cumulative increase in COVID-19 patients. The present study is a nationwide analysis of COVID-19 expected to help formulate regional policy for the 34 provinces to restrain the continuous spread of COVID-19 and strengthen the understanding of the spatial epidemiology of the disease. It addresses the identification of factors affecting the cumulative number of COVID-19 patients in Indonesia based on the health, population and economic factors. The present study used the cumulative data of COVID-19 patients in 34 provinces in Indonesia from December 2019 to April 2021.

The novelty of the present study is the development of a Nonparametric Geographically-Weighted Spline Regression (NGWSR) model with the optimization of GCV-NGWSR using two base matrices of polynomial and truncated functions. Furthermore, the results of the simultaneous and partial significance tests supported the NGWSR model. The present study produced 6 groups from 34 provinces based on the factors affecting the cumulative number of COVID-19 patients. The results of the NGWSR model showed that each group has different influencing factors, leading to a different handling of COVID-19 for each province. This produced a map of the spread of COVID-19 based factors affecting the cumulative number of COVID-19 patients, including the number of doctors, the number tourist of attractions, the number of Tuberculosis (TB) cases, the percentage of the elderly population, the percentage of households with access to proper drinking water, population density, Gross Regional Domestic Product (GRDP), the number of hospitals, and the number of villages/urban villages having a public health center. The present study contributes to the health sector and the government with regard to tracing based on the spatial mapping results of the spread of COVID-19.

2. Nonparametric Geographically-Weighted Spline Regression

The Nonparametric Geographically-Weighted Spline Regression (NGWSR) model is a development of the nonparametric regression for spatial data with local parameter estimators for each observation location. The spline approach is used to solve the problem of spatial analysis in which the shape of the regression curve is unknown. The NGWSR model assumes a normally distributed error with a mean of zero and a variance of $\sigma^2(u_i, v_i)$ at each location. Location coordinates (u_i, v_i) are among the important factors in determining the weights used to estimate the parameters of the model. The response variable and predictor variable data are given at $\left(x_{1i}, x_{2i}, \dots, x_{li}, y_i\right)$ and the relationship between $\left(x_{1i}, x_{2i}, \dots, x_{li}\right)$ and y_i is assumed to follow the NGWSR model. Mathematically, the form of the relationship between the response variable y_i and the predictor variables $\left(x_{1i}, x_{2i}, \dots, x_{li}\right)$ at the i-th location for the NGWSR model can be expressed as follows:

$$y_i = \beta_0(u_i, v_i) + \sum_{p=1}^{l}\sum_{k=1}^{m} \beta_{pk}(u_i, v_i)x_{pi}^k + \sum_{p=1}^{l}\sum_{h=1}^{r} \delta_{p,m+h}(u_i, v_i)K_{ph}\left(x_{pi} - K_{ph}\right)_+^m + \varepsilon_i \tag{1}$$

Equation (1) is a NGWSR model consisting of two bases, namely the polynomial basis of order m and a truncated basis function with the addition of knot points K_{ph}, distinguishing it from other nonparametric GWR, such as those of the previous study by Sifriyani, Kartiko, Budiantara and Gunardi (2017). Equation (1) has a polynomial at m degree with n area. The components of Equation (1) described as y_i are the response variable at the i-th location, where $i = 1, 2, \dots, n$. Furthermore, x_{pi} is the p-th predictor variables at the i-th area with $p = 1, 2, \dots, l$. K_{ph} is the h-th knot point on the component of the p-th predictor variable with $h = 1, 2, \dots, r$. $\beta_{pk}(u_i, v_i)$ is the polynomial component parameter of the multivariable NGWSR model. $\beta_{pk}(u_i, v_i)$ is the k-th parameter of the p-th predictor variable on the i-th area. $\delta_{p,m+h}(u_i, v_i)$ is the truncated component parameter of the NGWSR model. $\delta_{p,m+h}(u_i, v_i)$ is the l-th $+h$ parameter, on the h-th knot point and the p-th predictor variable on the i-th area.

The breakdown of the NGWSR model in Equation (1) can be expressed in Equation (2).

$$\begin{aligned} y_i = {} & \big(\beta_0(u_i, v_i) + \beta_{11}(u_i, v_i)x_{1i} + \beta_{12}(u_i, v_i)x_{1i}^2 + \cdots + \beta_{1m}(u_i, v_i)x_{1i}^m + \\ & + \beta_{21}(u_i, v_i)x_{2i} + \beta_{22}(u_i, v_i)x_{2i}^2 + \beta_{23}(u_i, v_i)x_{2i}^3 + \cdots + \beta_{2m}(u_i, v_i)x_{2i}^m + \cdots + \\ & + \beta_{l1}(u_i, v_i)x_{li} + \beta_{l2}(u_i, v_i)x_{li}^2 + \beta_{l3}(u_i, v_i)x_{li}^3 + \cdots + \beta_{lm}(u_i, v_i)x_{li}^m\big) + \\ & + \delta_{1,m+1}(u_i, v_i)K_{11}(x_{1i} - K_{11})_+^m + \cdots + \delta_{1,m+r}(u_i, v_i)K_{1r}(x_{1i} - K_{1r})_+^m + \\ & + \delta_{2,m+1}(u_i, v_i)K_{21}(x_{2i} - K_{21})_+^m + \cdots + \delta_{2,m+r}(u_i, v_i)K_{2r}(x_{1i} - K_{2r})_+^m + \\ & + \cdots + \delta_{l,m+1}(u_i, v_i)K_{l1}(x_{li} - K_{l1})_+^m + \cdots + \delta_{l,m+r}(u_i, v_i)K_{lr}(x_{li} - K_{lr})_+^m + \varepsilon_i \end{aligned} \tag{2}$$

Equations (1) and (2) for the n-th area can be written according to Equation (3):

$$y_n = \beta_0(u_n, v_n) + \sum_{p=1}^{l}\sum_{k=1}^{m} \beta_{pk}(u_n, v_n)x_{pn}^k + \sum_{p=1}^{l}\sum_{h=1}^{r} \delta_{p,m+h}(u_n, v_n)K_{ph}(x_{pn} - K_{ph})_+^m + \varepsilon_n \tag{3}$$

Equations (1) and (2) can be expressed as Equation (4):

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \beta_0(u_1,v_1) + \sum_{p=1}^{l}\sum_{k=1}^{m} \beta_{pk}(u_1,v_1)x_{pn}^k \\ \beta_0(u_2,v_2) + \sum_{p=1}^{l}\sum_{k=1}^{m} \beta_{pk}(u_2,v_2)x_{pn}^k \\ \vdots \\ \beta_0(u_n,v_n) + \sum_{p=1}^{l}\sum_{k=1}^{m} \beta_{pk}(u_n,v_n)x_{pn}^k \end{bmatrix} + \begin{bmatrix} \sum_{p=1}^{l}\sum_{h=1}^{r} \delta_{p,m+h}(u_1,v_1)K_{ph}\left(x_{pn}-K_{ph}\right)_+^m \\ \sum_{p=1}^{l}\sum_{h=1}^{r} \delta_{p,m+h}(u_2,v_2)K_{ph}\left(x_{pn}-K_{ph}\right)_+^m \\ \vdots \\ \sum_{p=1}^{l}\sum_{h=1}^{r} \delta_{p,m+h}(u_n,v_n)K_{ph}\left(x_{pn}-K_{ph}\right)_+^m \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \tag{4}$$

Thus, Equation (4) can be expressed as Equation (5):

$$\tilde{\mathbf{Y}} = \tilde{f} + \tilde{\varepsilon} = \mathbf{X}\tilde{\beta}(u_i,v_i) + \mathbf{P}\tilde{\delta}(u_i,v_i) + \tilde{\varepsilon}. \tag{5}$$

2.1. Estimation of NGWSR Models

The estimator parameters $\tilde{\beta}(u_i,v_i)$ and $\tilde{\delta}(u_i,v_i)$ of the NGWSR model is presented in Theorem 1.

Theorem 1. *If a regression model (1) is given with a normally distributed error with a mean of zero and a variance of $\sigma^2(u_i,v_i)$ and the weighted likelihood function is given by () then, the MLE estimator for $\hat{\tilde{\beta}}(u_i,v_i)$ and $\hat{\tilde{\delta}}(u_i,v_i)$ is given by*

$$\hat{\tilde{\beta}}(u_i,v_i) = \mathbf{A}(\mathbf{K})\tilde{\mathbf{Y}}$$
$$\hat{\tilde{\delta}}(u_i,v_i) = \mathbf{B}(\mathbf{K})\tilde{\mathbf{Y}}$$

where

$$\mathbf{A}(\mathbf{K}) = \mathbf{S}\left(\mathbf{X}^T\mathbf{W}(u_i,v_i)\mathbf{X}\right)^{-1}\left[\mathbf{X}^T - \mathbf{X}^T\mathbf{W}(u_i,v_i)\mathbf{P}\left(\mathbf{P}^T\mathbf{W}(u_i,v_i)\mathbf{P}\right)^{-1}\mathbf{P}^T\right]\mathbf{W}(u_i,v_i)$$
$$\mathbf{B}(\mathbf{K}) = \mathbf{R}\left(\mathbf{P}^T\mathbf{W}(u_i,v_i)\mathbf{P}\right)^{-1}\left[\mathbf{P}^T - \mathbf{P}^T\mathbf{W}(u_i,v_i)\mathbf{X}\left(\mathbf{X}^T\mathbf{W}(u_i,v_i)\mathbf{X}\right)^{-1}\mathbf{X}^T\right]\mathbf{W}(u_i,v_i)$$

The regression curve estimator $\hat{\tilde{f}}$ contains a polynomial component represented by the **X** matrix and a *truncated* component represented by the **P** matrix. If the matrix is $\mathbf{P} = \mathbf{0}$, then the NGWSR regression curve estimator $\hat{\tilde{f}}$ will go to the polynomial parametric regression curve estimator in the GWR model. Furthermore, if $\mathbf{P} = \mathbf{0}$ and the **X** matrix contains a linear function, the NGWSR regression curve estimator $\hat{\tilde{f}}$ will become a linear parametric regression curve estimator in the GWR model (multiple linear regression in the GWR model) developed by many researchers, such as Fotheringham et al. (2002), Brunsdon and Fotheringham (1999), Demsar et al. (2008), Li, Jiao and Browder (2016), shan Wu, Yang, Guo and Han (2017) and Benassi and Naccarato (2017).

The best NGWSR estimator is obtained by using optimal knot points. The knot point is a joint point where there is a change in the behavior pattern of the function or curve. The optimal knot point in the NGWSR model of Equation (1) is obtained using the Generalized Cross-Validation Development method, namely GCV-NGWSR of Equation (6)

$$GCV(K_1,K_2,\ldots,K_r) = \frac{MSE(K_1,K_2,\ldots,K_r)}{\left(\frac{1}{n}\text{ trace }[\mathbf{I} - \mathbf{A}(K)B(K)]\right)^2} \tag{6}$$

by having two hat matrices of $\mathbf{A}(\mathbf{K})$ and $\mathbf{B}(\mathbf{K})$, respectively. The $\mathbf{A}(\mathbf{K})$ matrix is polynomial-based and the $(\mathbf{K})$ matrix is function-based that contains knots.

$$\mathbf{A}(\mathbf{K}) = \mathbf{S}\left(\mathbf{X}^T\mathbf{W}(u_i,v_i)\mathbf{X}\right)^{-1}\left[\mathbf{X}^T - \mathbf{X}^T\mathbf{W}(u_i,v_i)\mathbf{P}\left(\mathbf{P}^T\mathbf{W}(u_i,v_i)\mathbf{P}\right)^{-1}\mathbf{P}^T\right]\mathbf{W}(u_i,v_i)$$
$$\mathbf{B}(\mathbf{K}) = \mathbf{R}\left(\mathbf{P}^T\mathbf{W}(u_i,v_i)\mathbf{P}\right)^{-1}\left[\mathbf{P}^T - \mathbf{P}^T\mathbf{W}(u_i,v_i)\mathbf{X}\left(\mathbf{X}^T\mathbf{W}(u_i,v_i)\mathbf{X}\right)^{-1}\mathbf{X}^T\right]\mathbf{W}(u_i,v_i)$$

with **I** as the identity matrix, n number of observations, and $MSE(K_1,K_2,\ldots,K_r)$ the mean squared error value of NGWSR. Thus, Equation (6) is different from the GCV as said by Hardle (1990), Utami, Haris, Prahutama and Purnomo (2020), Budiantara and Purnomo (2011), and Yuniartika, Kusnandar and Mara (2013).

2.2. Simultaneous Parameter Significance Test of NGWSR Model

The form of the simultaneous hypothesis testing is as follows:

$$H_0: \beta_{11}(u_i, v_i) = \beta_{12}(u_i, v_i) = \dots = \beta_{pm}(u_i, v_i) = \delta_{1,m+1}(u_i, v_i) = \delta_{1,m+2}(u_i, v_i) = \dots = \delta_{i,m+r}(u_i, v_i) = 0,$$
$$i = 1, 2, \dots, n$$
$$H_1: \text{There is at least one } \beta_{kj}(u_i, v_i) \neq 0 \text{ or } \delta_{k,m+h}(u_i, v_i) \neq 0, \quad k = 1, 2, \dots, p; \; j = 1, 2, \dots, m;$$
$$h = 1, 2, \dots, r; i = 1, 2, \dots, n$$

and the statistic test for the simultaneous test is given by:

$$V^* = \frac{\left(\dfrac{\mathbf{y}^T \mathbf{M}(u_i, v_i)\mathbf{y}}{tr((\mathbf{I} - \mathbf{B}_\omega)^T(\mathbf{I} - \mathbf{B}_\omega))}\right)}{\left(\dfrac{\mathbf{y}^T \mathbf{D}(u_i, v_i)\mathbf{y}}{tr((\mathbf{I} - \xi)^T(\mathbf{I} - \xi))}\right)} \tag{7}$$

with

$$\mathbf{M}(u_i, v_i) = (\mathbf{I} - \mathbf{B}_\omega)^T(\mathbf{I} - \mathbf{B}_\omega)$$

$$\mathbf{B}_\omega = \begin{bmatrix} \dfrac{W_{1(1)}}{\sum_{j=1}^{n} w_{1(j)}} & \dfrac{W_{2(1)}}{\sum_{j=1}^{n} w_{1(j)}} & \cdots & \dfrac{W_{n(1)}}{\sum_{j=1}^{n} w_{1(j)}} \\ \dfrac{W_{1(2)}}{\sum_{j=1}^{n} w_{2(j)}} & \dfrac{W_{2(2)}}{\sum_{j=1}^{n} w_{2(j)}} & \cdots & \dfrac{W_{n(2)}}{\sum_{j=1}^{n} w_{2(j)}} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{W_{1(n)}}{\sum_{j=1}^{n} w_{n(j)}} & \dfrac{W_{2(n)}}{\sum_{j=1}^{n} w_{n(j)}} & \cdots & \dfrac{W_{n(n)}}{\sum_{j=1}^{n} w_{n(j)}} \end{bmatrix}$$

The statistic test of the hypothesis testis is simultaneous, with **y** as the vector of response variable and **I** as the identity matrix. H_0 is rejected if $F_{count}(V^*) > F_{\alpha,db_1,db_2}$ or $p - value < \alpha$.

2.3. Significance Test of Partial Parameters of the NGWSR Model

The partial hypothesis testing is formulated as follows.

$$H_0: \beta_{pj}(u_i, v_i) = 0 \text{ and } \delta_{k,m+h}(u_i, v_i) = 0$$
$$\text{with } k = 1, 2, \dots, p; j = 1, 2, \dots, m; h = 1, 2, \dots, r; i = 1, 2, \dots, n$$
$$H_1: \text{There is at least one} \beta_{kj}(u_i, v_i) \neq 0 \text{ or } \delta_{k,m+h}(u_i, v_i) \neq 0$$
$$k = 1, 2, \dots, p; j = 1, 2, \dots, m; h = 1, 2, \dots, r; i = 1, 2, \dots, n$$

The statistic tests for the partial test of the NGWSR model are as follows:

$$t = \frac{\eta(u_i, v_i)}{SE(\eta(u_i, v_i))} \tag{8}$$

where $\eta(u_i, v_i)$ is the estimator of $\eta(u_i, v_i)$ which is a size vector of $[1 + (l \times m) + (l \times r)] \times 1$ where the $\eta(u_i, v_i)$ vector makes up the parameters $\beta(u_i, v_i)$ and $\delta(u_i, v_i)$, $SE(\eta(u_i, v_i)) = \sqrt{g_{kk}}$. g_{kk} is the k-th $+1$ diagonal element of the $(\mathbf{Q}^T\mathbf{W}(u_i, v_i)\mathbf{Q})^{-1}\sigma^2(u_i, v_i)$ matrix. **Q** is the size matrix $n \times [1 + (l \times m) + (l \times r)]$ containing predictor variables and predictor variables with truncated spline components. The statistical test follows the t distribution with degrees of freedom at a significance level of α; thus, the decision is to reject H_0 if the value of $|t| > t_{\left(\frac{\alpha}{2},(n-1)\right)}$.

Table 1
Description of Variables and Data Sources of Observations

Variables	Symbols	Description of Variables	Data Source of Observations	Unit
Response	y	Cumulative number of COVID-19 patients in Indonesia	Official website www.COVID19.go.id	People
Predictor	x_1	Number of Doctors	Statistics Indonesia 2019-2021 ([BPS], 2019a), ([BPS], 2020a), ([BPS], 2021a)	People
	x_2	Number of Tourist Attractions	Statistics Indonesia 2019-2021 ([BPS], 2019a), ([BPS], 2020a), ([BPS], 2021a)	Place
	x_3	Number of TB Cases	Statistics Indonesia 2019-2021 ([BPS], 2019a), ([BPS], 2020a), ([BPS], 2021a)	Cases
	x_4	Percentage of Elderly Population	Statistics Indonesia 2019-2021 ([BPS], 2019b), ([BPS], 2020b), ([BPS], 2021b)	%
	x_5	Percentage of Households with access to proper drinking water	Statistics Indonesia 2019-2021 ([BPS], 2019a), ([BPS], 2020a), ([BPS], 2021a)	%
	x_6	Population density	Statistics Indonesia 2019-2021 ([BPS], 2019a), ([BPS], 2020a), ([BPS], 2021a)	People / Km^2
	x_7	Gross Regional Domestic Product at Market Price	Statistics Indonesia 2019-2021 ([BPS], 2019a), ([BPS], 2020a), ([BPS], 2021a)	Billion Rupiah
	x_8	Number of Hospitals	Statistics Indonesia 2019-2021 ([BPS], 2019a), ([BPS], 2020a), ([BPS], 2021a)	Units
	x_9	Number of Villages/Urban villages having Public Health Center	Statistics Indonesia 2019-2021 ([BPS], 2019a), ([BPS], 2020a), ([BPS], 2021a)	Units

3. Methodology

3.1. Data and Data Sources

The present study used the 2019, 2020, and 2021 cross-sectional data sourced from the Statistics Indonesia and the official website www.COVID19.go.id. The variable y originated from the cumulative number of COVID-19 patients in Indonesia from December 2019 to April 2021. The predictor variable x consisted of several sectors of observation, including the health, population sector and economic sectors. The variables consisted of the number of doctors and health professionals, the number of tourist attractions, the number of TB cases, the percentage of the elderly population, the percentage of households with access to proper drinking water, population density, Gross Regional Domestic Product (GRDP) at market price, which is the amount of gross added value arising from all economic sectors in a region, the number of hospitals, and the number of villages/urban villages having public health center. Descriptions and data sources of observations in the present study are shown in Table 1.

3.2. Data Analysis Procedure

Analysis and modeling were conducted using the statistical program R-3.4.2. The stages of analysis in the present study were as follows:

1. Conduct data exploration to obtain an overview and information about the response variable and explanatory variables used.

2. Make a descriptive statistical analysis of each variable consisting of area mapping, average concentration size, standard deviation, minimum value and maximum value.
3. Analyze the linear pattern between the response variable and each predictor variable.
4. Identify the spatial diverts of the tests for spatial heterogeneity using the Breusch-Pagan method.
   The hypothesis used in the Breusch-Pagan test:
   H_0 : The variance among regions is the same.
   H_1 : There is a difference in variance among regions.
   Test statistics:

$$BP = \frac{1}{2}\mathbf{f}^T\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{f}, \tag{9}$$

   with $\mathbf{f} = (f_1, f_2, \ldots, f_n)^T$
   with $\mathbf{f} = \left(\frac{\varepsilon_i^2}{\sigma^2} - 1\right)$, $\varepsilon_i = y_i - \hat{y}_i$
   $\mathbf{Z}$ is a matrix containing sectors that have been normalized and standardized for each observation. H_0 is rejected if $BP > \chi^2_{(p)}$ or $p - value < \alpha$ with p being the number of predictors (Kosfeld, 2006), (Anselin, 2003).
5. Calculate the Euclidean distance (Fotheringham et al., 2002), (Sifriyani, Ruslan and Susanty, 2019b) between the i-th location located at coordinates (u_i, v_i) to the j-th location located at coordinates (u_i, v_i) using Equation (10).

$$d_{ij} = \sqrt{\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2} \tag{10}$$

6. Calculate the geographical weighting matrix by using the geographic weighting function of the Gaussian kernels (Fotheringham et al., 2002), (Sifriyani, Budiantara, Kartiko and Gunardi, 2019a), (Sifriyani et al., 2019b).

$$w_j(u_i, v_i) = \exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{b_i}\right)^2\right) \tag{11}$$

   and the Bisquare kernels (Fotheringham et al., 2002), (Sifriyani, Ruslan and Susanty, 2019c), (Sifriyani et al., 2019b).

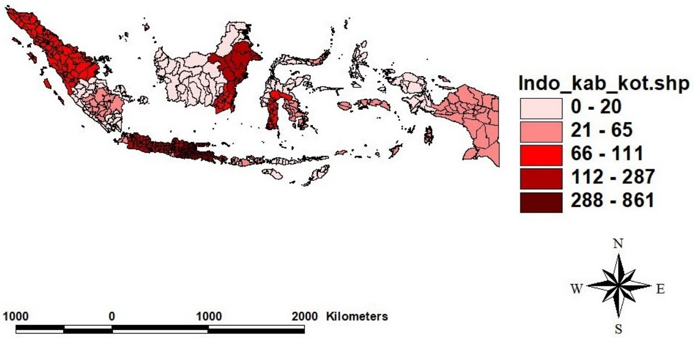
$$w_j(u_i, v_i) = \begin{cases} \left(1 - \left(\frac{d_{ij}}{b_i}\right)^2\right)^2, & \text{for } d_{ij} \le b_i \\ 0, & \text{for } d_{ij} > b_i \end{cases} \tag{12}$$

7. Determine the optimal knot point using the Generalized Cross-Validation (GCV-NGWSR) development method of Equation (6).
8. Obtain the estimator of nonparametric geographically-weighted spline regression model by using Theorem 1.
9. Test the simultaneous parameter significance according to Equation (7).
10. Test the hypothesis partially or test the location effect partially for each predictor variable for the NGWSR model according to Equation (8).
11. Interpret the NGWSR model to obtain the factors affecting the cumulative number of COVID-19 patients in 34 provinces in Indonesia.
12. Map the areas of Kalimantan based on the significant predictor variables.

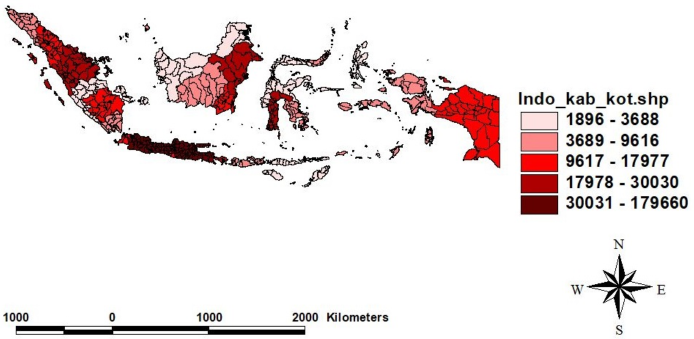
4. Results and Discussion

4.1. Data Exploration

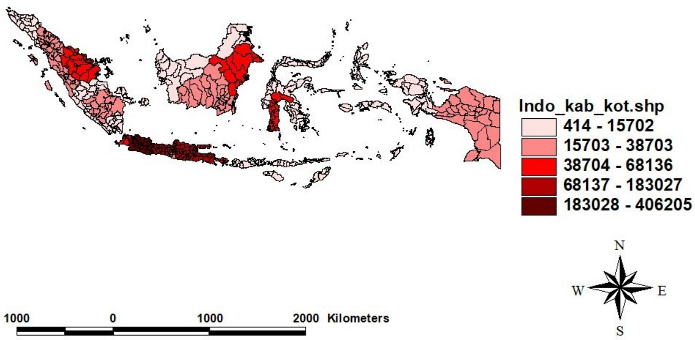
A NGWSR model was developed by using of observational data of the cumulative number of COVID-19 in April 2021, taking into account the previous data of December 2019 to March 2021. The spatial distribution and mapping for observational data of cumulative number of COVID19 patients for 2019 to 2021 in Indonesia are shown in Figure 1.



(a) Number of Positive Cases for COVID-19 in 2019



(b) Number of Positive Cases for COVID-19 in 2020



(c) Number of Positive Cases for COVID-19 in 2021

Figure 1: Spatial distribution of cumulative data of COVID-19 patients in Indonesia

Figure 1 shows that the spatial distribution of the cumulative data of COVID-19 patients in Indonesia is divided into 4 groups. The green color category represents the cumulative number of COVID-19 patients of 20,308 people. This category includes such provinces as Papua, West Papua, Maluku, Central Sulawesi, North Sulawesi, West Sulawesi, Southeast Sulawesi, Central Kalimantan, West Kalimantan, North Kalimantan and Aceh. The blue color category represents the cumulative number of COVID-19 patients with the highest figure relative to the yellow and brown categories. The highest figure with the cumulative number of COVID-19 patients of 183,028 to 406,205 occurred in the provinces of West Java and DKI Jakarta. Figure 2 shows a bar chart of comparison of the cumulative number of COVID-19 patients in several provinces. DKI Jakarta and West Java have the highest cumulative number of COVID-19 patients in Indonesia for the period of April 2021.

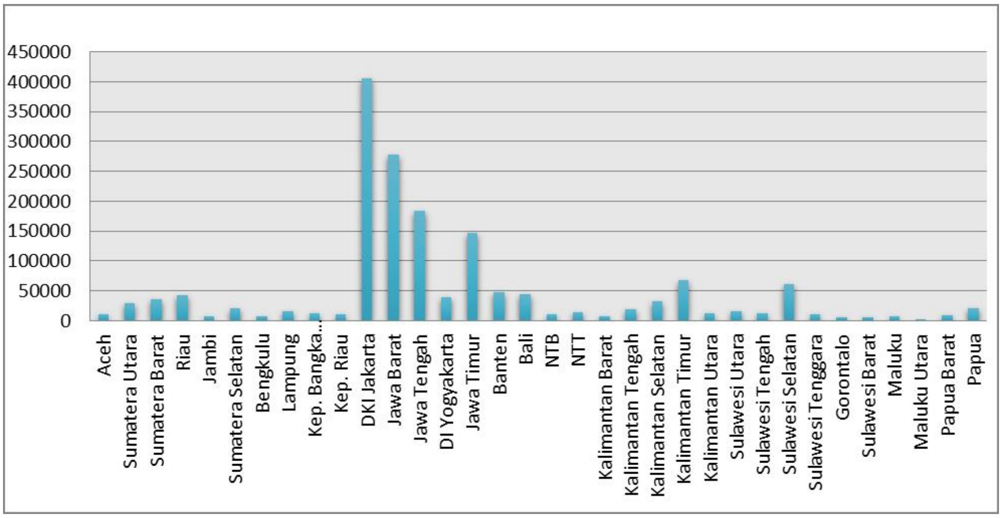


Figure 2: Bar chart of cumulative number of COVID-19 Patients in Indonesia

Table 2
Description of Variables and Data Sources of Observations

Observation data	Variables	Number of Observation Data (n Province)	Minimum	Maximum	Mean	Standard Deviation
Cumulative number of COVID-19 Patients in Indonesia	y	34	414	406205	48464,53	85194,51
Health sector	x_1	34	302	11365	2382,676	3051,568
	x_3	34	33,90	92,20	54,23529	16,71807
	x_4	34	3,630	14,710	8,330882	2,439596
	x_5	34	62,47	99,84	85,40529	9,590242
	x_8	34	10	296	68,94118	76,50508
	x_9	34	59	1094	308,2059	251,5978
Population and Economy sectors	x_2	34	6	529	85,17647	129,8189
	x_6	34	9	15907	739,2353	2708,964
	x_7	34	41726	2772381	464284,7	672108,7

Description of observational data comprising the response variable y and predictor variable x is presented in Table 2. Statistical description consists of mean, standard deviation, and minimum and maximum values. The descriptive statistics is calculated using the R software and the output of the software is shown in Table 2.

Table 2 shows that the average cumulative number of COVID-19 patients in Indonesia amounted to 48464.53 people. The lowest cumulative number of COVID-19 patients occurs in the North Maluku Province with 414 people, while the highest cumulative number of COVID-19 patients occurs in DKI Jakarta with 406,205 people.

4.2. Spatial Effect Tests using the Breusch-Pagan Method

A spatial heterogeneity test aims to determine whether or not the research data have a spatial effect. If so, the study proceeds with using spatial analysis modeling. A spatial effect test using the Breusch-Pagan method is expressed in Equation (5).

The results of the Breusch-Pagan test showed a spatial heterogeneity; thus, the study proceeded with using the spatial analysis modeling.

Table 3
Results of the Breusch-Pagan test analysis

Breusch-Pagan	df	$P-value$	Remarks
21.113	18	0.027373	There is a spatial heterogeneity.

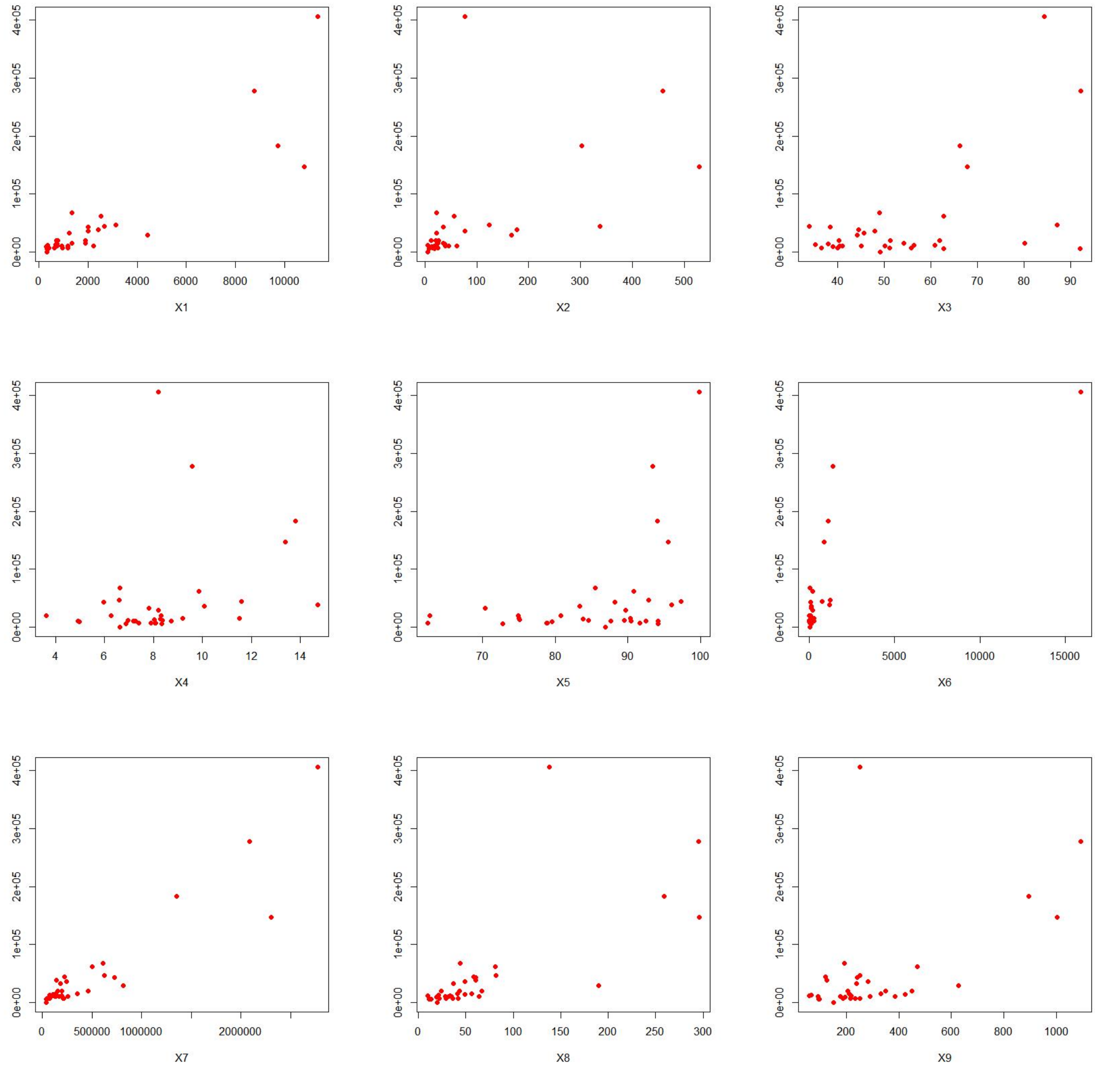


Figure 3: Patterns of relationship between the cumulative number of COVID-19 patient and each predictor variable

4.3. Patterns of Relationships between Variables

The patterns of the relationship between variables are determined for initial identification to find out whether the observation data are suitable for nonparametric regression.

As Figure 3 shows, the patterns of relationship formed between the cumulative number of COVID-19 patients and each predictor variable do not follow a certain pattern and there are changing patterns at certain sub-intervals. These changes occurred because each province used different characteristics. Therefore, nonparametric regression is

Table 4
CV values for weighting function

Weight Function	Bandwidth Value	Generalized Cross-Validation (GCV)
Gaussian kernel function	42.29	1.1311×10^{10}
Bisquare kernel function	42.199	1.3413×10^{10}

Table 5
Optimal Knot Points and Model Goodness-of-Fit Criteria

Order	Knot Points	R-squared	AIC	RMSE
1	1	81.379	11.888	5.5892×10^{-7}
	2	60.212	11.82	4.0521×10^{-5}
	3	99.298	17.433	7032.2

feasible to use. Since spatial effects and nonparametric regression were identified, the Geographically-Weighted Spline Regression (NGWSR) model was used.

4.4. Geographical Weighting

The first stage of the NGWSR model is to obtain geographic weighting. The geographic weighting of the Gaussian kernel function is presented in Equation (7) and the Bisquare kernel function is presented in Equation (8). The first step is to determine the coordinates of the observation locations and calculate the Euclidean distance among the observation locations using Equation (10). Table 4 shows the calculation results of bandwidth and GCV values.

Of the two weighting functions, the Gaussian kernel function was found to be the best with a GCV value smaller than that of the Bisquare kernel function. The NGWSR model is presented upon determination of the Gaussian kernel function as the best weighting function.

4.5. Nonparametric Geographically-Weighted Spline Regression Modeling

The mathematical model of Nonparametric Geographically-Weighted Spline Regression is presented in Equation (13).

$$y_i = \beta_0(u_i, v_i) + \sum_{p=1}^{l} \sum_{j=1}^{m} \beta_{pj}(u_i, v_i) x_{pi}^{j} + \sum_{p=1}^{l} \sum_{h=1}^{r} \delta_{p,m+h}(u_i, v_i) \left(x_{pi} - K_{ph}\right)_{+}^{m} + \varepsilon_i \tag{13}$$

With m as order and h as knot points, an optimum knot point and the values of GCV, R-Squared, AIC and RMSE would be selected in order to obtain the best model as shown in Table 5.

As Table 5 shows, the best model is the NGWSR model with order $m = 1$ and 3 knot points, according to the R^2 value of to 99.298. This value indicates that such predictor variables as the number of doctors, the number of tourist attractions, the number of TB cases, the percentage of the elderly population, the percentage of households with access to proper drinking water, population density, GRDP, the number of hospitals and the number of villages/urban villages having public health center are able to account for 99.298% of the cumulative number of COVID-19 patients in Indonesia, while the remaining 0.702% is accounted for by other variables not studied.

The mathematical model of Nonparametric Geographically-Weighted Spline Regression of Equation (14), with an order $m = 1$ and a knot point $h = 3$ is as follows:

$$y_i = \beta_0(u_i, v_i) + \sum_{p=1}^{9} \sum_{j=1}^{1} \beta_{pj}(u_i, v_i) x_{pi}^{j} + \sum_{p=1}^{9} \sum_{h=1}^{3} \delta_{p,m+h}(u_i, v_i) \left(x_{pi} - K_{ph}\right)_{+}^{m} + \varepsilon_i \tag{14}$$

With total predictor variables of $p = 9$, namely x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, and x_9.

Table 6
Estimator of cumulative number of COVID-19 patients in Indonesia

Provinces	Cumulative number of COVID-19 patients in Indonesia (y)	Cumulative number of COVID-19 patients in Indonesia ($\hat{y}$)
Aceh	10814	8177.9
North Sumatera	29198	30192
West Sumatera	36268	36404
Riau	42698	53339
Jambi	7407	16689
South Sumatera	20068	29209
Bengkulu	6630	10645
Lampung	15702	28651
Bangka Belitung Islands	12822	4908
Riau Islands	10838	3065.6
DKI Jakarta	406200	406300
West Java	277550	276510
Central Java	183030	184000
DI Yogyakarta	38703	32884
East Java	146810	146540
Banten	47101	49427
Bali	44236	48524
NTB	10291	12984
NTT	14200	13402
West Kalimantan	7503	13424
Central Kalimantan	19780	18126
South Kalimantan	32612	14746
East Kalimantan	68136	51186
North Kalimantan	11702	17766
North Sulawesi	15638	13660
Central Sulawesi	12125	20550
South Sulawesi	61419	46929
Southeast Sulawesi	10396	14923
Gorontalo	5355	8228.3
West Sulawesi	5443	5536.9
Maluku	7515	3503
North Maluku	414	-1082.6
West Papua	8879	7207.1
Papua	20308	19868

4.6. Estimation of Nonparametric Geographically-Weighted Spline Regression Parameters

Furthermore, the estimation stage of NGWSR is presented in Table 6.

Figure 4 shows that the estimator of response variable of the cumulative number of COVID-19 patients $\hat{Y}$ produced by NGWR has a value close to the original data (Y).

The red line is the estimator of the response variable of the cumulative number of COVID-19 patients in Indonesia using NGWR and the blue line is the observation data of the cumulative number of COVID-19 patients in Indonesia in 2021.

4.7. Significance Test of Simultaneous Parameters

The significance test of simultaneous parameters aims to determine the 9 predictor variables, including the number of doctors, the number of tourist attractions, the number of TB cases, the percentage of the elderly population, the percentage of households with access to proper drinking water, population density, GRDP, number of hospitals and the number of villages/urban villages having public health center which has a simultaneous effect on the cumulative number of COVID-19 patients in Indonesia. The test used Equation (11) at a significance level of $\alpha = 0.05$ and the results of ANOVA calculation are shown in Table 7.

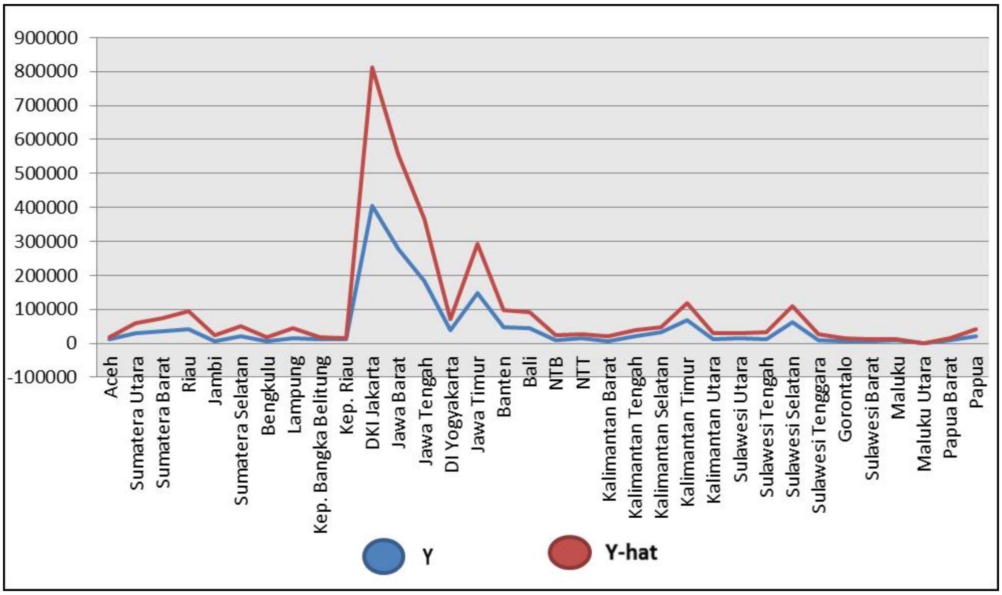


Figure 4: Distribution pattern of the cumulative number of COVID-19 patients in Indonesia

Table 7
ANOVA

Source of diversity	Number of Squares	Degree of freedom	Median square	F-count	$p - value$
Regression	2.377×10^{11}	33	7.2053×10^{9}	85.7079	1.1102×10^{-15}
Error	1.6814×10^{9}	20	8.4068×10^{7}		
Total	2.3945×10^{11}	53			

Table 7 shows that $V(85.7079) > F_{(0.05;33;20)}$; thus, the predictor variables in the NGWSR model have a significant effect on the response variable.

4.8. Significance Test of Partial Parameters

The significance test of partial parameters as expressed in Equation (11) was aimed at testing each predictor variable for each effect on the response variable of the cumulative number of COVID-19 patients in Indonesia. Results of the partial significance test and the grouping by province are presented in Table 8 and the distribution map is given in Figure 5.

As Figure 5 shows, there are 6 groups with their own influence factor. The largest number of provinces are in group 1 with factors affecting the cumulative number of COVID-19 patients include the number of doctors, the number of tourist attractions, the number of TB cases, the percentage of the elderly population, the percentage of households with access to proper drinking water, population density, GRDP, the number of hospitals, and the number of villages/urban villages having public health center. The factors affecting the cumulative number of of COVID-19 patients based on the NGWSR model for each province are presented in Table 8.

Conclusions

The model developed in the present study uses the Nonparametric Geographically-Weighted Spline Regression (NGWSR) with GCV-NGWSR optimization that uses two bases of polynomial and truncated matrices. The novelty of the model lies on the significance test of simultaneous and partial parameters that support the NGWSR model. The variables with significant effects on the cumulative number of COVID-19 patients in Indonesia for 34 provinces are the number of doctors and health professionals, the number of tourist attractions, the number of TB cases, the percentage of the elderly population, the percentage of households with access to proper drinking water, population density, Gross

Table 8
Grouping of the cumulative number of COVID-19 patients in Indonesia in 2021 based on the significant variables

Groups	Provinces	Influence factors
1	North Sumatera West Sumatera DKI Jakarta West Java Central Java East Java Banten NTT Central Kalimantan North Sulawesi West Sulawesi North Maluku West Papua Papua	Number of doctors (x_1), number of tourist attractions (x_2), numbers of TB cases (x_3), percentage of the elderly population (x_4), percentage of households with access to proper drinking water (x_5), population density (x_6), GRDP (x_7), number of hospitals (x_8) and number of villages/urban villages having public health center (x_9)
2	NTB	Number of doctors (x_1), number of tourist attractions (x_2), numbers of TB cases (x_3), percentage of the elderly population (x_4), percentage of households with access to proper drinking water (x_5), population density (x_6), number of hospitals (x_8) and number of villages/urban villages having public health center (x_9)
3	Aceh Jambi South Sumatra Bengkulu Bangka Belitung Islands Riau Islands DI Yogyakarta Bali West Kalimantan North Kalimantan Southeast Sulawesi Gorontalo	Number of doctors (x_1), number of tourist attractions (x_2), numbers of TB cases (x_3), percentage of the elderly population (x_4), percentage of households with access to proper drinking water (x_5), population density (x_6), number of hospitals (x_8)
4	Riau	Number of doctors (x_1), number of tourist attractions (x_2), numbers of TB cases (x_3), percentage of the elderly population (x_4), percentage of households with access to proper drinking water (x_5), number of hospitals (x_8)
5	Central Sulawesi Maluku	Number of doctors (x_1), number of tourist attractions (x_2), percentage of the elderly population (x_4), percentage of households with access to proper drinking water (x_5), population density (x_6), number of hospitals (x_8)
6	Lampung South Kalimantan East Kalimantan South Sulawesi	Number of doctors (x_1), number of tourist attractions (x_2), numbers of TB cases (x_3), percentage of households with access to proper drinking water (x_5), and number of hospitals (x_8)

Regional Domestic Product. (GDP) at market price, which is the amount of gross added value arising from all economic sectors in a region, the number of hospitals, and the number of villages/urban villages having public health center. The NGWSR model produces a coefficient of determination or an accurate model goodness-of-fit of 99.298%. The present study contribute to the health sector and the government with regard to tracing on the basis of results of spatial mapping of the spread of COVID-19. Additionally, it contributes to the field of Geographic Information Systems (GIS).

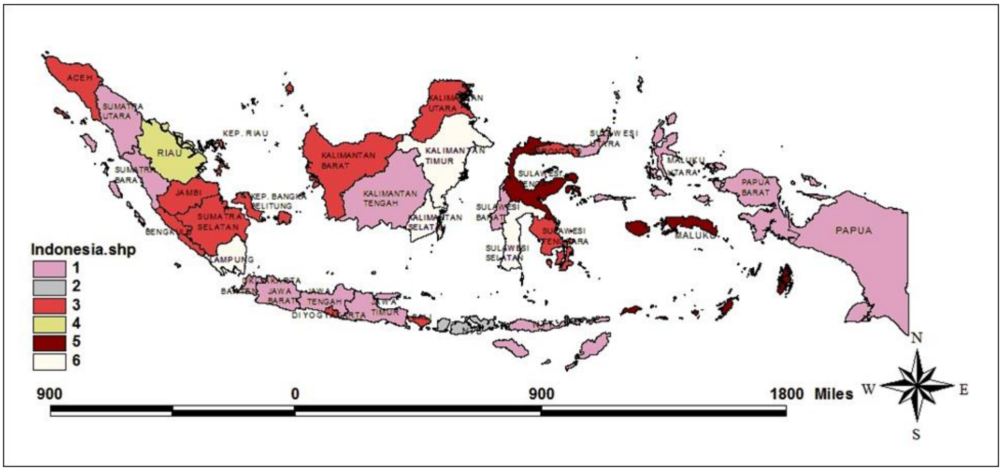


Figure 5: Map of the cumulative number of COVID-19 patients in Indonesia in 2021 based on the significant variables

Further studies are recommended to develop NGWSR models by adding up time parameters to complete local nonparametric regression analyses according to the time series for each observation location.

Acknowledgments

The research work is facilitated by the support of the Indonesian Ministry of Education, Culture, Research and Technology and Mulawarman University. Hence, the authors greatly appreciate the support.

CRediT authorship contribution statement

Sifriyani: Developing a Nonparametric-Geographically Weighted Spline Regression (NGWSR) model using GCV-NGWSR optimization and data analysis. **Dedi Rosadi:** Creating a program syntax to support the computing system thereby it can be used to process data based on the NGWSR model. **Mariani Rasjid:** Collecting the nationwide COVID-19 data of the 34 Provinces from December 2019 to April 2021.

References

Anselin, L., 2003. An Introduction to Spatial Regression Analysis in R.

Arab-Mazar, Z., Sah, R., Rabaan, A., Dhama, K., Rodriguez-Morales, A., 2020. Mapping the incidence of the covid-19 hotspot in iran âĂŞ implications for travellers. Travel Med. Infect. Dis. doi:10.1016/j.tmaid.2020.101630.

Benassi, F., Naccarato, A., 2017. Households in potential economic distress. a geographically weighted regression model for italy, 2001-2011. Spatial Statistics 21, 362–376. doi:10.1016/j.spasta.2017.03.002.

[BPS], B.P.S., 2019a. Statistik Indonesia 2019. Badan Pusat Statistik, Jakarta Pusat.

[BPS], B.P.S., 2019b. Statistik Penduduk Lanjut Usia 2019. Badan Pusat Statistik, Jakarta Pusat.

[BPS], B.P.S., 2020a. Statistik Indonesia 2020. Badan Pusat Statistik, Jakarta Pusat.

[BPS], B.P.S., 2020b. Statistik Penduduk Lanjut Usia 2020. Badan Pusat Statistik, Jakarta Pusat.

[BPS], B.P.S., 2021a. Statistik Indonesia 2021. Badan Pusat Statistik, Jakarta Pusat.

[BPS], B.P.S., 2021b. Statistik Penduduk Lanjut Usia 2021. Badan Pusat Statistik, Jakarta Pusat.

Brunsdon, Fotheringham, 1999. Some notes on parametric significance test for geographically weighted regression. Journal of Regional Science 39, 497–524. doi:10.1111/0022-4146.00146.

Budiantara, I., Purnomo, J., 2011. infants' weight growth model in surabaya (indonesia) by using weighted spline regression. International Journal of Basic, Applied Sciences 11, 119–123.

Buzai, G., 2020. De wuhan a luján. Evolución espacial del COVID-19 3.

Chen, Z., Zhang, Q., Lu, Y., Guo, Z., Zhang, X., Zhang, W., Lu, J., 2020b. Distribution of the covid-19 epidemic and correlation with population emigration from wuhan, china. Chinese medical journal doi:10.1097/CM9.0000000000000782.

Crespo, R., Fotheringham, A., Charlton, M., 2007. Application of geograhically weighted regression to a 19-year set of house price data in london to calibrate local hedonic price models, in: Proceeding of the 9th International Conference on Geocomputation.

Demsar, Fotheringham, Charlton, 2008. Testing the importance of the explanatory variables in a mixed geographically weighted regression model. Exploring the spatio-temporal dynamics of geographical processes with geograhically weighted regression, geovisual analytics 7, 181–197.

Desjardins, M., Hohl, A., Delmelle, E., 2020. Rapid surveillance of covid-19 in the united states using a prospective space-time scan statistic: detecting and evaluating emerging clusters. Appl. Geogr. .

Fotheringham, A., Brundson, C., Charlton, M., 2002. Geographically Weighted Regression : The Analysis of Spatially Varying Relationships. John Wiley Sons Ltd, England.

Giuliani, D., Dickson, M., Espa, G., Santi, F., 2020. Modelling and predicting the spatio- temporal spread of coronavirus disease 2019 (covid-19) in italy. SSRN doi:10.2139/ssrn.3559569.

Guan, W., Ni, Z., Hu, Y., Liang, W., Ou, C., He, J., Du, B., 2020. Clinical character- istics of coronavirus disease 2019 in china. New England Journal of Medicine 382, 1708âĂŞ1720.

Hardle, G., 1990. Applied Nonparametric Regression. Cambridge University Press, New York.

Huang, Wu, Barry, 2010. International journal geographical inference science. Geographically, temporally weighted regression for modeling spatio-temporal variation in house prices 24, 383–401.

Huang, H., Wang, Y., Wang, Z., Liang, Z., Qu, S., Ma, S., Liu, X., 2020. Clinical character- istics of coronavirus disease 2019 in china. China 3.

Kosfeld, R., 2006. Spatial econometric. http//www.scribd.com.

Leung, Y., Mei, C., Zhang, W., 2000a. Statistic test for spatial non stationarity based on the geographically weighted regression model. Journal Environment, Planing A 32, 9–32.

Leung, Y., Mei, C., Zhang, W., 2000b. Testing for spatial autocorrelation among the residuals of the geograhically weighted regression. Journal Environment, Planing A 32, 871–890.

Li, Y., Jiao, Y., Browder, J.A., 2016. Modeling spatially-varying ecological relationships using geographically weighted generalized linear model: A simulation study based on longline seabird bycatch. Fisheries Research 181, 14–24. doi:10.1016/j.fishres.2016.03.024.

Mei, Wang, Zhang, 2006. Testing the importance of the explanatory variables in a mixed geographically weighted regression model. Journal Environment, Planning 38, 587–598.

Nakaya, Fotheringham, Brunsdon, Charltron, 2005. Geographically weighted poisson regression for disease association mapping. Journal Statistics Medicine 24, 2695–2717.

Saha, A., Gupta, K., Patil, M., 2020. Monitoring and epidemiological trends of coronavirus disease (covid-19) around the world. Osfpreprints .

Satuan Gugus Tugas Penanganan COVID-19, 2021. Peta sebaran. www.covid19.go.id/peta-sebaran.

Sifriyani, Budiantara, I.N., Kartiko, S., Gunardi, 2019a. Evaluation of factors affecting increased unemployment in east java using ngwr-ts method. International Journal of Sciences: Basic, Applied Research 49, 123–142.

Sifriyani, Kartiko, S., Budiantara, I., Gunardi, 2017. Geographically weighted regression with spline approach. Far East Journal of Mathematical Sciences 101, 1183–1196.

Sifriyani, Kartiko, S.H., Budiantara, I.N., Gunardi, 2018. Development of nonparametric geographically weighted regression using truncated spline approach. Songklanakarin Journal of Science, Technology 40, 909–920.

Sifriyani, Rosadi, D., 2020. Pemodelan susceptible infected recovered (sir) untuk estimasi angka reproduksi covid-19 di kalimantan timur, samarinda. Media Statistika 13.

Sifriyani, Ruslan, Susanty, F., 2019b. Evaluation of forest productivity, governance on the preservation of tropical rain forests in kalimantan using the ngwr-ts nonparametric geospatial method. EurAsian Journal of BioSciences 13, 2373–2379.

Sifriyani, Ruslan, Susanty, F., 2019c. Mapping, analysis factors of affecting productivity tropical rain forests in east kalimantan. Modern Applied Science 13. doi:10.5539/mas.v13n10p112.

Utami, T.W., Haris, M.A., Prahutama, A., Purnomo, E.A., 2020. Optimal knot selection in spline regression using unbiased risk, generalized cross validation methods. Journal of Physics: Conference Series 1446, 012049.

Wrenn, D., Sam, A., 2014. Geographically, temporally weighted likelihood regression : Exploring the spatiotemporal determinants of land use change. Regional Science, Urban Economics 44, 60–74.

shan Wu, S., Yang, H., Guo, F., Han, R.M., 2017. Spatial patterns, origins of heavy metals in sheyang river catchment in jiangsu, china based on geographically weighted regression. Science of the Total Environment 580, 1518–1529. doi:10.1016/j.scitotenv.2016.12.137.

Xiong, Y., Guang, Y., Chen, F., Zhu, F., 2020. Spatial statistics and inïňĆuencing factors of the novel coronavirus pneumonia 2019 epidemic in hubei province, china. ResearchSquare doi:10.21203/rs.3.rs-16858/v2.

Yu, D., 2010. Exploring spatiotemporally varying regressed relationships: the geographically weighted panel regression analysis. proceedings of the joint international conference on theory. Data Handling, Modeling in GeoSpatial Information Science , 134–139.

Yuniartika, Y., Kusnandar, D., Mara, M.N., 2013. Penentuan generalized cross validation (gcv) sebagai kriteria dalam pemilihan model regresi b-spline terbaik. Bultin Ilmiah Matematika, Statistika, Terapannya 2, 121–126.

Zhang, X., Rao, H., Wu, Y., Huang, Y., Dai, H., 2020. Comparison of the spatiotemporal characteristics of the covid-19 and sars outbreaks in mainland china. Geography and Sustainability doi:10.1101/2020.03.23.20034058.

Zhou, C., Su, F., Pei, T., Zhang, A., Du, Y., Luo, B., Song, C., 2020. Covid-19: challenges to gis with big data. Geography and Sustainability 1, 77âĂŞ87.

https://scholar.google.co.id/citations?user=79VWksEAAAAJ
Lecturer at Statistics Study Program, Mulawarman University, Samarinda, Indonesia. Now she is an associate professor. Graduated from Master's Education in the Master's Degree in Statistics at the Sepuluh November Institute of Technology, Surabaya, Indonesia and Doctoral Education in the Doctoral Program in Statistics at Gadjah Mada University. Short-study as visitor for the Spatial Statistics Research at the Faculty of Science and Technology at Lancaster University, United Kingdom. Doing some researches since 2017 until now particularly focused on developing researches in the field of Spatial Statistics and has produced several reputable articles about spatial statistics. She earned a Research Grant for the Topic of Spatial Statistics Development from the Ministry of Education and Culture in 2019-2020 and 2021-2022. In 2020 as a member of the Task Force Team to accelerate the handling of COVID-19 at Mulawarman University, Samarinda, Indonesia.

https://acadstaff.ugm.ac.id/dedirosadi
https://scholar.google.co.id/citations?user=3WlY-YUAAAAJ
I am a full time professor in the Department of Mathematics, Faculty of Mathematics and Sciences, Universitas Gadjah Mada (UGM). I graduated from Statistics Study Program in Gadjah Mada University in February 1996 and started to work as the lecturer in UGM. In August 1997, I continue my study to obtain the master of science degree in Stochastic Modeling (Applied Statistics) in the University of Twente, The Netherland, graduated in June 1999. From September2001- September 2004 I took my doctoral study (in Econometrics) in the Institute of Econometrics and Operation Research (EOS 119), the Vienna University of Technology (TU Wien), Austria. After finished my doctoral study, i come back to Yogyakarta and continue my work in the Universitas Gadjah Mada until now. In September 2013, I receive the full professor title.

https://scholar.google.co.id/citations?user=6-mDEVEAAAAJ
Lecturer at the Faculty of Medicine, Tadulako University, Palu, Central Sulawesi, Indonesia. Now she is also a doctor at Tadulako General Hospital, Palu, Central Sulawesi, Indonesia. During the COVID-19 pandemic, the author served as the Head of the Task Force Team to accelerate the handling of COVID-19 at the Tadulako General Hospital, Indonesia, the head of the COVID-19 Task Force Team at Muhammadiyah, Central Sulawesi, Indonesia. The author has created a Pocket Book of Getting to Know COVID-19. Graduated from the Medical Faculty of Hasanuddin University in 2009 and successfully completed the Pulmonology Specialist Education and Respiratory Medicine at Medical Faculty of Airlangga University in 2017. Joined the Professional Organizations called Indonesian Medical Association (IDI) for Palu City Branch and the Indonesian Society of Respiralogy for Sulawesi Branch.

Highlights

A Geospatial Analysis Using The Development Nonparametric Geographically-Weighted Spline Regression and Its Application for COVID-19 Modeling

- The novelty of this study lies in the development of a NGWSR model with GCV-NGWSR optimization using two polynomial and truncated matrix bases.
- The present study develops a new model in spatial statistics called the Nonparametric Geographically-Weighted Spline Regression model.
- Program syntax is created based on the developed NGWSR model to support the computing system so that it can be used to process data based on the NGWSR model.
- The study offers a new methodology in the fields of Geographic Information Systems (GIS) and spatial statistics.
- Implementation of the NGWSR Model using the nationwide COVID-19 data in 34 provinces in Indonesia.
- This study used the nationwide COVID-19 data of the 34 provinces from December 2019 to April 2021.

A Geospatial Analysis Using The Development Nonparametric Geographically-Weighted Spline Regression and Its Application for COVID-19 Modeling

ARTICLE INFO

Keywords:
COVID-19
Geographically
Nonparametric
Spline
Spatial
GIS

ABSTRACT

The present study deals with the development of a Nonparametric Geographically-Weighted Spline Regression (NGWSR) model which is applied to detect the spread of COVID-19 (Coronavirus disease 2019) in a spatial context. It is a spatial epidemiology discussing how the health, population and economic sectors affect the cumulative number of COVID-19 patients. The present study used COVID-19 nationwide data of 34 provinces in Indonesia from December 2019 to April 2021. The study offers a new methodology in the fields of Geographic Information Systems (GIS) and spatial statistics. The novelty of this study lies in the development of a NGWSR model with GCV-NGWSR optimization using two bases of polynomial and truncated matrices. Furthermore, the results of the simultaneous and partial significance tests supported the NGWSR model. The present study produced 6 groups from 34 provinces based on the factors affecting the cumulative number of COVID-19 patients. A map of COVID-19 spread was made on the basis of the total number of doctors and/or health professionals, the number of tourist attractions, the number of Tuberculosis (TB) cases, the percentage of the elderly population, the percentage of households with access to proper drinking water, population density, Gross Regional Domestic Product (GRDP) supporting the economic sector of a region, the number of hospitals, and the number of villages/urban villages having a public health center (puskesmas). The present study contributes to the health sector and the government with regard to tracing based on the spatial mapping results of the spread of COVID-19.

1. Introduction

The regression analysis used for spatial data modeling is known as the geographically-weighted regression (GWR) model. GWR was first introduced by Fotheringham in 1967. The response variable in the GWR model contains predictor variables, in which each of its regression coefficient depends on the location where the data are observed. GWR takes the point approach. Each parameter value is estimated at each observation location, so that each point of observation location has a different parameter value (Fotheringham, Brundson and Charlton, 2002). If each parameter has a constant value in each geographic location, the GWR model will be the same as the linear regression model. This means that each geographic location has the same model. Studies undertaken using the GWR theory include Brunsdon and Fotheringham (1999), Crespo, Fotheringham and Charlton (2007), Leung, Mei and Zhang (2000a), and Leung, Mei and Zhang (2000b).

Those statisticians who carried out studies on the GWR are including Nakaya, Fotheringham, Brunsdon and Charltron (2005) who developed the Geographically Weighted Poisson Regression Models, and Mei, Wang and Zhang (2006) who produced a model that combines the global regression model and GWR known as the Mixed Geographically-Weighted Regression models. The GWR model is also developed in the time-series-related spatial field known as spatio-temporal studies, in which exploratory spatio-temporal data are analyzed using the GWR and geo-visual analytics, as undertaken by Demsar, Fotheringham and Charlton (2008). The spatio-temporal study was also used by Huang, Wu and Barry (2010) in the Geographically- and Temporally-Weighted Regression Model. Furthermore, Yu (2010) studies spatial panel data using Geographically-Weighted Panel Regression and Wrenn and Sam (2014) in their study used Geographically- and Temporally-Weighted Likelihood Regression models.

The developed GWR model remains in a linear form. In some cases of regression analysis, the question is that whether all relationships between predictor variables and response variable will form a known regression curve with a linear pattern. In fact, indeed, not all data relationship patterns are linear and have a known regression curve. Given the rapid development of science and technology and natural phenomena leading to unusual patterns, it is very difficult to predict the behavior of nature. For example, in the past decade, one was able to predict exactly when the dry and rainy

ORCID(s):

seasons began and ended in various geographic areas, making farmers capable of preparing themselves for planting rice and harvesting it. However, currently it is difficult to do so. Other problems, such as the percentage of poverty, underdevelopment, literacy rates, increasing ignorance and uneven development in each region and their causative variables constitute examples of events with unclear patterns, not following a certain pattern, having no linear pattern and as if irregular. Thus, it is necessary to develop nonparametric regressions in the GWR model to overcome and provide solutions to these problems.

A good regression model should be viewed from a variety of aspects and a modeling problem should be placed proportionally. Differences in environmental and geographical characteristics among observation locations result in observations with different variations or differences in the effect of predictor variables on response variable for each observation location. The GWR model is not able to resolve the problem of the effect of predictor variables on the response variable not following a specific pattern and varying patterns in certain sub-interval. Thus, it is highly necessary to develop studies in the field of nonparametric regression and spatial statistics. Previous studies developed nonparametric regression models in GWR (Sifriyani, Kartiko, Budiantara and Gunardi, 2018) and continued. The present study develops a new model in spatial statistics called the Nonparametric Geographically-Weighted Spline Regression model.

The present study would contribute to the Geographic Information Systems (GIS) and spatial statistics fields. It can be applied to a variety of natural phenomena. It does not only address the problems of city planning and regional planning, but also mapping the distribution of diseases in an area, one of which being viral infections, such as COVID-19, the outbreaks of which hitting numerous countries, including Indonesia. The COVID-19 situation in Indonesia has not subsided. The COVID-19 Task Force reported that the number of positive cases of COVID-19 on 26 June 2021 was 21,095 people. This figure again broke the highest daily record in Indonesia of 20,547 positive people for the COVID-19 virus (Satuan Gugus Tugas Penanganan COVID-19, 2021), (Sifriyani and Rosadi, 2020). Multiple causes lead to this increase. Therefore, the present study investigated the causes of the cumulative increase in COVID-19 patients in Indonesia based on factors related to the health, population and economic sectors using the Nonparametric Geographically-Weighted Spline Regression (NGWSR) model.

Research that studies the modeling of COVID-19 using spatial statistics is carried out by several researchers. Spatial statistical modeling research for confirmed cases of COVID-19 in China was conducted by Guan, Ni, Hu, Liang, Ou, He and Du (2020), Chen, Zhang, Lu, Guo, Zhang, Zhang and Lu (2020b), Huang, Wang, Wang, Liang, Qu, Ma and Liu (2020), Zhou, Su, Pei, Zhang, Du, Luo and Song (2020), Zhang, Rao, Wu, Huang and Dai (2020), and Xiong, Guang, Chen and Zhu (2020). A spatial statistical modeling study for confirmed cases of COVID-19 in the USA was conducted by Desjardins, Hohl and Delmelle (2020). The research Spatiotemporal analysis of COVID-19 at the national and provincial levels was conducted by Arab-Mazar, Sah, Rabaan, Dhama and Rodriguez-Morales (2020), Giuliani, Dickson, Espa and Santi (2020), Saha, Gupta and Patil (2020) and Buzai (2020). This research is about geospatial models and its application for cumulative number cases of COVID-19 in Indonesia.

The NGWSR model is applied to detect the spread of COVID-19 in a spatial context. A spatial epidemiological study deals with the effect of the health, population and economy sectors on the cumulative increase in COVID-19 patients. The present study is a nationwide analysis of COVID-19 expected to help formulate regional policy for the 34 provinces to restrain the continuous spread of COVID-19 and strengthen the understanding of the spatial epidemiology of the disease. It addresses the identification of factors affecting the cumulative number of COVID-19 patients in Indonesia based on the health, population and economic factors. The present study used the cumulative data of COVID-19 patients in 34 provinces in Indonesia from December 2019 to April 2021.

The novelty of the present study is the development of a Nonparametric Geographically-Weighted Spline Regression (NGWSR) model with the optimization of GCV-NGWSR using two base matrices of polynomial and truncated functions. Furthermore, the results of the simultaneous and partial significance tests supported the NGWSR model. The present study produced 6 groups from 34 provinces based on the factors affecting the cumulative number of COVID-19 patients. The results of the NGWSR model showed that each group has different influencing factors, leading to a different handling of COVID-19 for each province. This produced a map of the spread of COVID-19 based factors affecting the cumulative number of COVID-19 patients, including the number of doctors, the number tourist of attractions, the number of Tuberculosis (TB) cases, the percentage of the elderly population, the percentage of households with access to proper drinking water, population density, Gross Regional Domestic Product (GRDP), the number of hospitals, and the number of villages/urban villages having a public health center. The present study contributes to the health sector and the government with regard to tracing based on the spatial mapping results of the spread of COVID-19.

2. Nonparametric Geographically-Weighted Spline Regression

The Nonparametric Geographically-Weighted Spline Regression (NGWSR) model is a development of the nonparametric regression for spatial data with local parameter estimators for each observation location. The spline approach is used to solve the problem of spatial analysis in which the shape of the regression curve is unknown. The NGWSR model assumes a normally distributed error with a mean of zero and a variance of $\sigma^2(u_i, v_i)$ at each location. Location coordinates (u_i, v_i) are among the important factors in determining the weights used to estimate the parameters of the model. The response variable and predictor variable data are given at $\left(x_{1i}, x_{2i}, \ldots, x_{li}, y_i\right)$ and the relationship between $\left(x_{1i}, x_{2i}, \ldots, x_{li}\right)$ and y_i is assumed to follow the NGWSR model. Mathematically, the form of the relationship between the response variable y_i and the predictor variables $\left(x_{1i}, x_{2i}, \ldots, x_{li}\right)$ at the i-th location for the NGWSR model can be expressed as follows:

$$y_i = \beta_0(u_i, v_i) + \sum_{p=1}^{l} \sum_{k=1}^{m} \beta_{pk}(u_i, v_i) x_{pi}^k + \sum_{p=1}^{l} \sum_{h=1}^{r} \delta_{p,m+h}(u_i, v_i) K_{ph} \left(x_{pi} - K_{ph}\right)_+^m + \varepsilon_i \tag{1}$$

Equation (1) is a NGWSR model consisting of two bases, namely the polynomial basis of order m and a truncated basis function with the addition of knot points K_{ph}, distinguishing it from other nonparametric GWR, such as those of the previous study by Sifriyani, Kartiko, Budiantara and Gunardi (2017). Equation (1) has a polynomial at m degree with n area. The components of Equation (1) described as y_i are the response variable at the i-th location, where $i = 1, 2, \ldots, n$. Furthermore, x_{pi} is the p-th predictor variables at the i-th area with $p = 1, 2, \ldots, l$. K_{ph} is the h-th knot point on the component of the p-th predictor variable with $h = 1, 2, \ldots, r$. $\beta_{pk}(u_i, v_i)$ is the polynomial component parameter of the multivariable NGWSR model. $\beta_{pk}(u_i, v_i)$ is the k-th parameter of the p-th predictor variable on the i-th area. $\delta_{p,m+h}(u_i, v_i)$ is the truncated component parameter of the NGWSR model. $\delta_{p,m+h}(u_i, v_i)$ is the l-th $+h$ parameter, on the h-th knot point and the p-th predictor variable on the i-th area.

The breakdown of the NGWSR model in Equation (1) can be expressed in Equation (2).

$$\begin{aligned} y_i = &\big(\beta_0(u_i, v_i) + \beta_{11}(u_i, v_i)x_{1i} + \beta_{12}(u_i, v_i)x_{1i}^2 + \cdots + \beta_{1m}(u_i, v_i)x_{1i}^m + \\ &+ \beta_{21}(u_i, v_i)x_{2i} + \beta_{22}(u_i, v_i)x_{2i}^2 + \beta_{23}(u_i, v_i)x_{2i}^3 + \cdots + \beta_{2m}(u_i, v_i)x_{2i}^m + \cdots + \\ &+ \beta_{l1}(u_i, v_i)x_{li} + \beta_{l2}(u_i, v_i)x_{li}^2 + \beta_{l3}(u_i, v_i)x_{li}^3 + \cdots + \beta_{lm}(u_i, v_i)x_{li}^m\big) + \\ &+ \delta_{1,m+1}(u_i, v_i)K_{11}(x_{1i} - K_{11})_+^m + \cdots + \delta_{1,m+r}(u_i, v_i)K_{1r}(x_{1i} - K_{1r})_+^m + \\ &+ \delta_{2,m+1}(u_i, v_i)K_{21}(x_{2i} - K_{21})_+^m + \cdots + \delta_{2,m+r}(u_i, v_i)K_{2r}(x_{1i} - K_{2r})_+^m + \\ &+ \cdots + \delta_{l,m+1}(u_i, v_i)K_{l1}(x_{li} - K_{l1})_+^m + \cdots + \delta_{l,m+r}(u_i, v_i)K_{lr}(x_{li} - K_{lr})_+^m + \varepsilon_i \end{aligned} \tag{2}$$

Equations (1) and (2) for the n-th area can be written according to Equation (3):

$$y_n = \beta_0(u_n, v_n) + \sum_{p=1}^{l} \sum_{k=1}^{m} \beta_{pk}(u_n, v_n) x_{pn}^k + \sum_{p=1}^{l} \sum_{h=1}^{r} \delta_{p,m+h}(u_n, v_n) K_{ph} (x_{pn} - K_{ph})_+^m + \varepsilon_n \tag{3}$$

Equations (1) and (2) can be expressed as Equation (4):

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \beta_0(u_1, v_1) + \sum\limits_{p=1}^{l} \sum\limits_{k=1}^{m} \beta_{pk}(u_1, v_1) x_{pn}^k \\ \beta_0(u_2, v_2) + \sum\limits_{p=1}^{l} \sum\limits_{k=1}^{m} \beta_{pk}(u_2, v_2) x_{pn}^k \\ \vdots \\ \beta_0(u_n, v_n) + \sum\limits_{p=1}^{l} \sum\limits_{k=1}^{m} \beta_{pk}(u_n, v_n) x_{pn}^k \end{bmatrix} + \begin{bmatrix} \sum\limits_{p=1}^{l} \sum\limits_{h=1}^{r} \delta_{p,m+h}(u_1, v_1) K_{ph} \left(x_{pn} - K_{ph}\right)_+^m \\ \sum\limits_{p=1}^{l} \sum\limits_{h=1}^{r} \delta_{p,m+h}(u_2, v_2) K_{ph} \left(x_{pn} - K_{ph}\right)_+^m \\ \vdots \\ \sum\limits_{p=1}^{l} \sum\limits_{h=1}^{r} \delta_{p,m+h}(u_n, v_n) K_{ph} \left(x_{pn} - K_{ph}\right)_+^m \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \tag{4}$$

Thus, Equation (4) can be expressed as Equation (5):

$$\tilde{\mathbf{Y}} = \tilde{f} + \tilde{\varepsilon} = \mathbf{X}\tilde{\beta}(u_i, v_i) + \mathbf{P}\tilde{\delta}(u_i, v_i) + \tilde{\varepsilon}. \tag{5}$$

2.1. Estimation of NGWSR Models

The estimator parameters $\tilde{\beta}(u_i, v_i)$ and $\tilde{\delta}(u_i, v_i)$ of the NGWSR model is presented in Theorem 1.

Theorem 1. *If a regression model (1) is given with a normally distributed error with a mean of zero and a variance of $\sigma^2(u_i, v_i)$ and the weighted likelihood function is given by () then, the MLE estimator for $\hat{\tilde{\beta}}(u_i, v_i)$ and $\hat{\tilde{\delta}}(u_i, v_i)$ is given by*

$$\hat{\tilde{\beta}}(u_i, v_i) = \mathbf{A}(\mathbf{K})\tilde{\mathbf{Y}}$$
$$\hat{\tilde{\delta}}(u_i, v_i) = \mathbf{B}(\mathbf{K})\tilde{\mathbf{Y}}$$

where

$$\mathbf{A}(\mathbf{K}) = \mathbf{S}\left(\mathbf{X}^T\mathbf{W}(u_i, v_i)\mathbf{X}\right)^{-1}\left[\mathbf{X}^T - \mathbf{X}^T\mathbf{W}(u_i, v_i)\mathbf{P}\left(\mathbf{P}^T\mathbf{W}(u_i, v_i)\mathbf{P}\right)^{-1}\mathbf{P}^T\right]\mathbf{W}(u_i, v_i)$$
$$\mathbf{B}(\mathbf{K}) = \mathbf{R}\left(\mathbf{P}^T\mathbf{W}(u_i, v_i)\mathbf{P}\right)^{-1}\left[\mathbf{P}^T - \mathbf{P}^T\mathbf{W}(u_i, v_i)\mathbf{X}\left(\mathbf{X}^T\mathbf{W}(u_i, v_i)\mathbf{X}\right)^{-1}\mathbf{X}^T\right]\mathbf{W}(u_i, v_i)$$

The regression curve estimator $\hat{\tilde{f}}$ contains a polynomial component represented by the **X** matrix and a *truncated* component represented by the **P** matrix. If the matrix is $\mathbf{P} = \mathbf{0}$, then the NGWSR regression curve estimator $\hat{\tilde{f}}$ will go to the polynomial parametric regression curve estimator in the GWR model. Furthermore, if $\mathbf{P} = \mathbf{0}$ and the **X** matrix contains a linear function, the NGWSR regression curve estimator $\hat{\tilde{f}}$ will become a linear parametric regression curve estimator in the GWR model (multiple linear regression in the GWR model) developed by many researchers, such as Fotheringham et al. (2002), Brunsdon and Fotheringham (1999), Demsar et al. (2008), Li, Jiao and Browder (2016), shan Wu, Yang, Guo and Han (2017) and Benassi and Naccarato (2017).

The best NGWSR estimator is obtained by using optimal knot points. The knot point is a joint point where there is a change in the behavior pattern of the function or curve. The optimal knot point in the NGWSR model of Equation (1) is obtained using the Generalized Cross-Validation Development method, namely GCV-NGWSR of Equation (6)

$$GCV(K_1, K_2, \dots, K_r) = \frac{MSE(K_1, K_2, \dots, K_r)}{\left(\frac{1}{n}\text{ trace }[\mathbf{I} - \mathbf{A}(K)B(K)]\right)^2} \tag{6}$$

by having two hat matrices of $\mathbf{A}(\mathbf{K})$ and $\mathbf{B}(\mathbf{K})$, respectively. The $\mathbf{A}(\mathbf{K})$ matrix is polynomial-based and the $(\mathbf{K})$ matrix is function-based that contains knots.

$$\mathbf{A}(\mathbf{K}) = \mathbf{S}\left(\mathbf{X}^T\mathbf{W}(u_i, v_i)\mathbf{X}\right)^{-1}\left[\mathbf{X}^T - \mathbf{X}^T\mathbf{W}(u_i, v_i)\mathbf{P}\left(\mathbf{P}^T\mathbf{W}(u_i, v_i)\mathbf{P}\right)^{-1}\mathbf{P}^T\right]\mathbf{W}(u_i, v_i)$$
$$\mathbf{B}(\mathbf{K}) = \mathbf{R}\left(\mathbf{P}^T\mathbf{W}(u_i, v_i)\mathbf{P}\right)^{-1}\left[\mathbf{P}^T - \mathbf{P}^T\mathbf{W}(u_i, v_i)\mathbf{X}\left(\mathbf{X}^T\mathbf{W}(u_i, v_i)\mathbf{X}\right)^{-1}\mathbf{X}^T\right]\mathbf{W}(u_i, v_i)$$

with **I** as the identity matrix, n number of observations, and $MSE(K_1, K_2, \dots, K_r)$ the mean squared error value of NGWSR. Thus, Equation (6) is different from the GCV as said by Hardle (1990), Utami, Haris, Prahutama and Purnomo (2020), Budiantara and Purnomo (2011), and Yuniartika, Kusnandar and Mara (2013).

2.2. Simultaneous Parameter Significance Test of NGWSR Model

The form of the simultaneous hypothesis testing is as follows:

$$H_0: \beta_{11}(u_i, v_i) = \beta_{12}(u_i, v_i) = \cdots = \beta_{pm}(u_i, v_i) = \delta_{1,m+1}(u_i, v_i) = \delta_{1,m+2}(u_i, v_i) = \cdots = \delta_{i,m+r}(u_i, v_i) = 0,$$
$$i = 1, 2, \dots, n$$
$$H_1: \text{There is at least one } \beta_{kj}(u_i, v_i) \neq 0 \text{ or } \delta_{k,m+h}(u_i, v_i) \neq 0, \quad k = 1, 2, \dots, p;\ j = 1, 2, \dots, m;$$
$$h = 1, 2, \dots, r; i = 1, 2, \dots, n$$

and the statistic test for the simultaneous test is given by:

$$V^* = \frac{\left(\dfrac{\mathbf{y}^T\mathbf{M}(u_i, v_i)\mathbf{y}}{tr((\mathbf{I} - \mathbf{B}_\omega)^T(\mathbf{I} - \mathbf{B}_\omega))}\right)}{\left(\dfrac{\mathbf{y}^T\mathbf{D}(u_i, v_i)\mathbf{y}}{tr((\mathbf{I} - \xi)^T(\mathbf{I} - \xi))}\right)} \tag{7}$$

with

$$\mathbf{M}(u_i, v_i) = (\mathbf{I} - \mathbf{B}_{\omega})^T(\mathbf{I} - \mathbf{B}_{\omega})$$

$$\mathbf{B}_{\omega} = \begin{bmatrix} \frac{W_{1(1)}}{\sum_{j=1}^{n} w_{1(j)}} & \frac{W_{2(1)}}{\sum_{j=1}^{n} w_{1(j)}} & \cdots & \frac{W_{n(1)}}{\sum_{j=1}^{n} w_{1(j)}} \\ \frac{W_{1(2)}}{\sum_{j=1}^{n} w_{2(j)}} & \frac{W_{2(2)}}{\sum_{j=1}^{n} w_{2(j)}} & \cdots & \frac{W_{n(2)}}{\sum_{j=1}^{n} w_{2(j)}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{W_{1(n)}}{\sum_{j=1}^{n} w_{n(j)}} & \frac{W_{2(n)}}{\sum_{j=1}^{n} w_{n(j)}} & \cdots & \frac{W_{n(n)}}{\sum_{j=1}^{n} w_{n(j)}} \end{bmatrix}$$

The statistic test of the hypothesis testis is simultaneous, with $\mathbf{y}$ as the vector of response variable and $\mathbf{I}$ as the identity matrix. H_0 is rejected if $F_{count}(V^*) > F_{\alpha,db_1,db_2}$ or $p - value < \alpha$.

2.3. Significance Test of Partial Parameters of the NGWSR Model

The partial hypothesis testing is formulated as follows.

$$\begin{aligned} H_0 \ :& \beta_{pj}(u_i, v_i) = 0 \text{ and } \delta_{k,m+h}(u_i, v_i) = 0 \\ & \text{with } k = 1, 2, \ldots, p; j = 1, 2, \ldots, m; h = 1, 2, \ldots, r; i = 1, 2, \ldots, n \\ H_1 \ :& \text{There is at least one} \beta_{kj}(u_i, v_i) \neq 0 \text{ or } \delta_{k,m+h}(u_i, v_i) \neq 0 \\ & k = 1, 2, \ldots, p; j = 1, 2, \ldots, m; h = 1, 2, \ldots, r; i = 1, 2, \ldots, n \end{aligned}$$

The statistic tests for the partial test of the NGWSR model are as follows:

$$t = \frac{\eta(u_i, v_i)}{SE(\eta(u_i, v_i))} \tag{8}$$

where $\eta(u_i, v_i)$ is the estimator of $\eta(u_i, v_i)$ which is a size vector of $[1 + (l \times m) + (l \times r)] \times 1$ where the $\eta(u_i, v_i)$ vector makes up the parameters $\beta(u_i, v_i)$ and $\delta(u_i, v_i)$, $SE(\eta(u_i, v_i)) = \sqrt{g_{kk}}$. g_{kk} is the k-th +1 diagonal element of the $(\mathbf{Q}^T\mathbf{W}(u_i, v_i)\mathbf{Q})^{-1}\sigma^2(u_i, v_i)$ matrix. $\mathbf{Q}$ is the size matrix $n \times [1 + (l \times m) + (l \times r)]$ containing predictor variables and predictor variables with truncated spline components. The statistical test follows the t distribution with degrees of freedom at a significance level of α; thus, the decision is to reject H_0 if the value of $|t| > t_{\left(\frac{\alpha}{2},(n-1)\right)}$.

3. Methodology

3.1. Data and Data Sources

The present study used the 2019, 2020, and 2021 cross-sectional data sourced from the Statistics Indonesia and the official website www.COVID19.go.id. The variable y originated from the cumulative number of COVID-19 patients in Indonesia from December 2019 to April 2021. The predictor variable x consisted of several sectors of observation, including the health, population sector and economic sectors. The variables consisted of the number of doctors and health professionals, the number of tourist attractions, the number of TB cases, the percentage of the elderly population, the percentage of households with access to proper drinking water, population density, Gross Regional Domestic Product (GRDP) at market price, which is the amount of gross added value arising from all economic sectors in a region, the number of hospitals, and the number of villages/urban villages having public health center. Descriptions and data sources of observations in the present study are shown in Table 1.

3.2. Data Analysis Procedure

Analysis and modeling were conducted using the statistical program R-3.4.2. The stages of analysis in the present study were as follows:

Table 1
Description of Variables and Data Sources of Observations

Variables	Symbols	Description of Variables	Data Source of Observations	Unit
Response	y	Cumulative number of COVID-19 patients in Indonesia	Official website www.COVID19.go.id	People
Predictor	x_1	Number of Doctors	Statistics Indonesia 2019-2021 ([BPS], 2019a), ([BPS], 2020a), ([BPS], 2021a)	People
	x_2	Number of Tourist Attractions	Statistics Indonesia 2019-2021 ([BPS], 2019a), ([BPS], 2020a), ([BPS], 2021a)	Place
	x_3	Number of TB Cases	Statistics Indonesia 2019-2021 ([BPS], 2019a), ([BPS], 2020a), ([BPS], 2021a)	Cases
	x_4	Percentage of Elderly Population	Statistics Indonesia 2019-2021 ([BPS], 2019b), ([BPS], 2020b), ([BPS], 2021b)	%
	x_5	Percentage of Households with access to proper drinking water	Statistics Indonesia 2019-2021 ([BPS], 2019a), ([BPS], 2020a), ([BPS], 2021a)	%
	x_6	Population density	Statistics Indonesia 2019-2021 ([BPS], 2019a), ([BPS], 2020a), ([BPS], 2021a)	People / Km^2
	x_7	Gross Regional Domestic Product at Market Price	Statistics Indonesia 2019-2021 ([BPS], 2019a), ([BPS], 2020a), ([BPS], 2021a)	Billion Rupiah
	x_8	Number of Hospitals	Statistics Indonesia 2019-2021 ([BPS], 2019a), ([BPS], 2020a), ([BPS], 2021a)	Units
	x_9	Number of Villages/Urban villages having Public Health Center	Statistics Indonesia 2019-2021 ([BPS], 2019a), ([BPS], 2020a), ([BPS], 2021a)	Units

1. Conduct data exploration to obtain an overview and information about the response variable and explanatory variables used.
2. Make a descriptive statistical analysis of each variable consisting of area mapping, average concentration size, standard deviation, minimum value and maximum value.
3. Analyze the linear pattern between the response variable and each predictor variable.
4. Identify the spatial diverts of the tests for spatial heterogeneity using the Breusch-Pagan method.
   The hypothesis used in the Breusch-Pagan test:
   H_0 : The variance among regions is the same.
   H_1 : There is a difference in variance among regions.
   Test statistics:
   $$BP = \frac{1}{2}\mathbf{f}^T\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{f}, \tag{9}$$
   with $\mathbf{f} = (f_1, f_2, \dots, f_n)^T$
   with $\mathbf{f} = \left(\frac{\varepsilon_i^2}{\sigma^2} - 1\right)$, $\varepsilon_i = y_i - \hat{y}_i$
   $\mathbf{Z}$ is a matrix containing sectors that have been normalized and standardized for each observation. H_0 is rejected if $BP > \chi^2_{(p)}$ or $p - value < \alpha$ with p being the number of predictors (Kosfeld, 2006), (Anselin, 2003).
5. Calculate the Euclidean distance (Fotheringham et al., 2002), (Sifriyani, Ruslan and Susanty, 2019b) between the i-th location located at coordinates (u_i, v_i) to the j-th location located at coordinates (u_i, v_i) using Equation

(10).

$$d_{ij} = \sqrt{\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2} \tag{10}$$

6. Calculate the geographical weighting matrix by using the geographic weighting function of the Gaussian kernels (Fotheringham et al., 2002), (Sifriyani, Budiantara, Kartiko and Gunardi, 2019a), (Sifriyani et al., 2019b).

$$w_j(u_i, v_i) = \exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{b_i}\right)^2\right) \tag{11}$$

and the Bisquare kernels (Fotheringham et al., 2002), (Sifriyani, Ruslan and Susanty, 2019c), (Sifriyani et al., 2019b).

$$w_j(u_i, v_i) = \begin{cases} \left(1 - \left(\frac{d_{ij}}{b_i}\right)^2\right)^2, & \text{for } d_{ij} \le b_i \\ 0, & \text{for } d_{ij} > b_i \end{cases} \tag{12}$$

7. Determine the optimal knot point using the Generalized Cross-Validation (GCV-NGWSR) development method of Equation (6).
8. Obtain the estimator of nonparametric geographically-weighted spline regression model by using Theorem 1.
9. Test the simultaneous parameter significance according to Equation (7).
10. Test the hypothesis partially or test the location effect partially for each predictor variable for the NGWSR model according to Equation (8).
11. Interpret the NGWSR model to obtain the factors affecting the cumulative number of COVID-19 patients in 34 provinces in Indonesia.
12. Map the areas of Kalimantan based on the significant predictor variables.

4. Results and Discussion

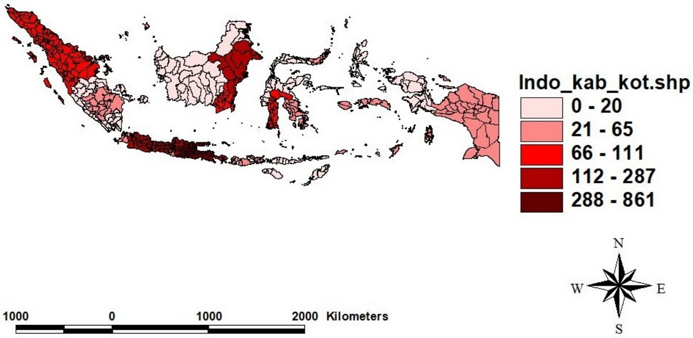
4.1. Data Exploration

A NGWSR model was developed by using of observational data of the cumulative number of COVID-19 in April 2021, taking into account the previous data of December 2019 to March 2021. The spatial distribution and mapping for observational data of cumulative number of COVID19 patients for 2019 to 2021 in Indonesia are shown in Figure 1.

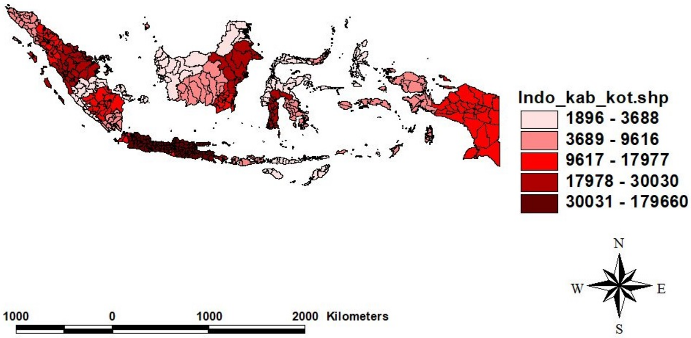
Figure 1 shows that the spatial distribution of the cumulative data of COVID-19 patients in Indonesia is divided into 4 groups. The green color category represents the cumulative number of COVID-19 patients of 20,308 people. This category includes such provinces as Papua, West Papua, Maluku, Central Sulawesi, North Sulawesi, West Sulawesi, Southeast Sulawesi, Central Kalimantan, West Kalimantan, North Kalimantan and Aceh. The blue color category represents the cumulative number of COVID-19 patients with the highest figure relative to the yellow and brown categories. The highest figure with the cumulative number of COVID-19 patients of 183,028 to 406,205 occurred in the provinces of West Java and DKI Jakarta. Figure 2 shows a bar chart of comparison of the cumulative number of COVID-19 patients in several provinces. DKI Jakarta and West Java have the highest cumulative number of COVID-19 patients in Indonesia for the period of April 2021.

Description of observational data comprising the response variable y and predictor variable x is presented in Table 2. Statistical description consists of mean, standard deviation, and minimum and maximum values. The descriptive statistics is calculated using the R software and the output of the software is shown in Table 2.

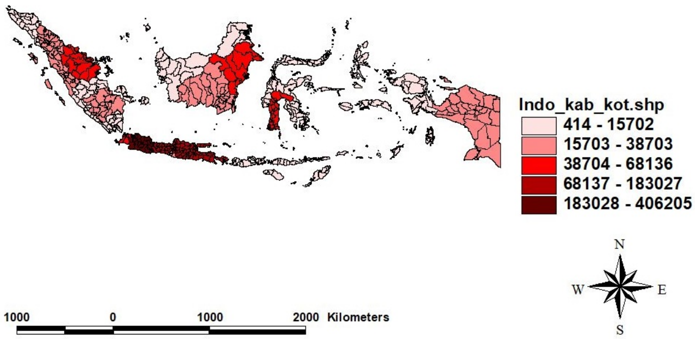
Table 2 shows that the average cumulative number of COVID-19 patients in Indonesia amounted to 48464.53 people. The lowest cumulative number of COVID-19 patients occurs in the North Maluku Province with 414 people, while the highest cumulative number of COVID-19 patients occurs in DKI Jakarta with 406,205 people.



(a) Number of Positive Cases for COVID-19 in 2019



(b) Number of Positive Cases for COVID-19 in 2020



(c) Number of Positive Cases for COVID-19 in 2021

Figure 1: Spatial distribution of cumulative data of COVID-19 patients in Indonesia

4.2. Spatial Effect Tests using the Breusch-Pagan Method

A spatial heterogeneity test aims to determine whether or not the research data have a spatial effect. If so, the study proceeds with using spatial analysis modeling. A spatial effect test using the Breusch-Pagan method is expressed in Equation (5).

The results of the Breusch-Pagan test showed a spatial heterogeneity; thus, the study proceeded with using the spatial analysis modeling.

4.3. Patterns of Relationships between Variables

The patterns of the relationship between variables are determined for initial identification to find out whether the observation data are suitable for nonparametric regression.

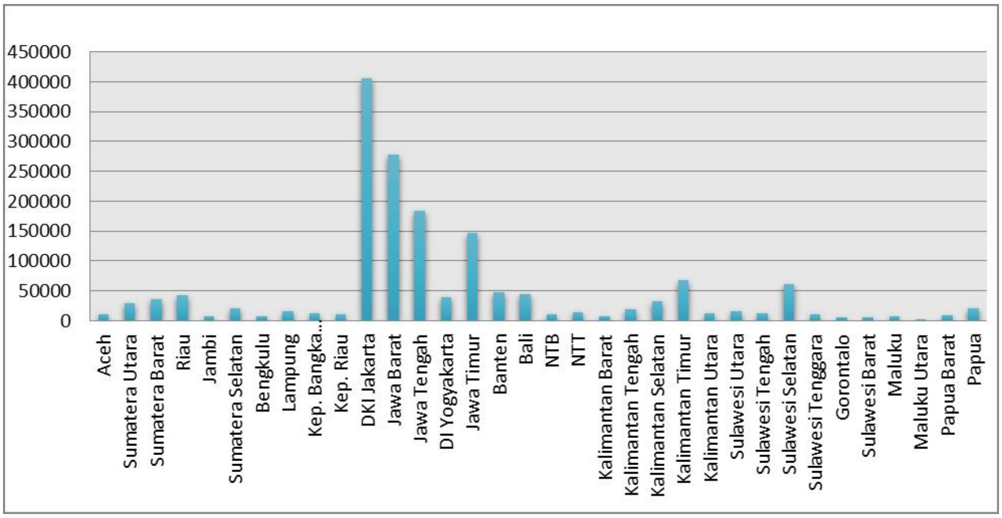


Figure 2: Bar chart of cumulative number of COVID-19 Patients in Indonesia

Table 2
Description of Variables and Data Sources of Observations

Observation data	Variables	Number of Observation Data (n Province)	Minimum	Maximum	Mean	Standard Deviation
Cumulative number of COVID-19 Patients in Indonesia	y	34	414	406205	48464,53	85194,51
Health sector	x_1	34	302	11365	2382,676	3051,568
	x_3	34	33,90	92,20	54,23529	16,71807
	x_4	34	3,630	14,710	8,330882	2,439596
	x_5	34	62,47	99,84	85,40529	9,590242
	x_8	34	10	296	68,94118	76,50508
	x_9	34	59	1094	308,2059	251,5978
Population and Economy sectors	x_2	34	6	529	85,17647	129,8189
	x_6	34	9	15907	739,2353	2708,964
	x_7	34	41726	2772381	464284,7	672108,7

Table 3
Results of the Breusch-Pagan test analysis

Breusch-Pagan	df	$P-value$	Remarks
21.113	18	0.027373	There is a spatial heterogeneity.

As Figure 3 shows, the patterns of relationship formed between the cumulative number of COVID-19 patients and each predictor variable do not follow a certain pattern and there are changing patterns at certain sub-intervals. These changes occurred because each province used different characteristics. Therefore, nonparametric regression is feasible to use. Since spatial effects and nonparametric regression were identified, the Geographically-Weighted Spline Regression (NGWSR) model was used.

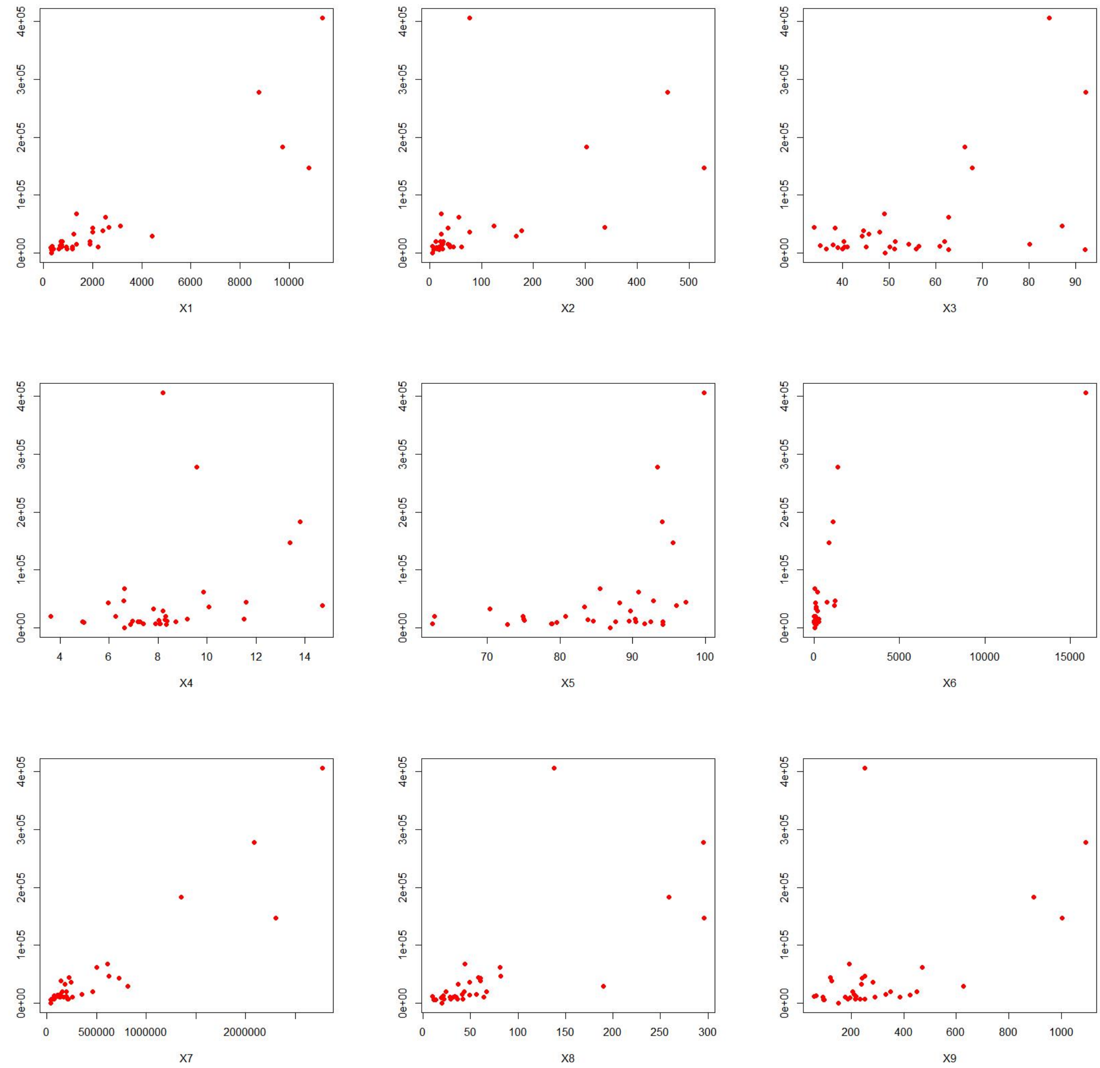


Figure 3: Patterns of relationship between the cumulative number of COVID-19 patient and each predictor variable

Table 4
CV values for weighting function

Weight Function	Bandwidth Value	Generalized Cross-Validation (GCV)
Gaussian kernel function	42.29	1.1311×10^{10}
Bisquare kernel function	42.199	1.3413×10^{10}

4.4. Geographical Weighting

The first stage of the NGWSR model is to obtain geographic weighting. The geographic weighting of the Gaussian kernel function is presented in Equation (7) and the Bisquare kernel function is presented in Equation (8). The first step is to determine the coordinates of the observation locations and calculate the Euclidean distance among the observation locations using Equation (10). Table 4 shows the calculation results of bandwidth and GCV values.

Table 5
Optimal Knot Points and Model Goodness-of-Fit Criteria

Order	Knot Points	R-squared	AIC	RMSE
1	1	81.379	11.888	5.5892×10^{-7}
	2	60.212	11.82	4.0521×10^{-5}
	3	99.298	17.433	7032.2

Of the two weighting functions, the Gaussian kernel function was found to be the best with a GCV value smaller than that of the Bisquare kernel function. The NGWSR model is presented upon determination of the Gaussian kernel function as the best weighting function.

4.5. Nonparametric Geographically-Weighted Spline Regression Modeling

The mathematical model of Nonparametric Geographically-Weighted Spline Regression is presented in Equation (13).

$$y_i = \beta_0(u_i, v_i) + \sum_{p=1}^{l}\sum_{j=1}^{m} \beta_{pj}(u_i, v_i) x_{pi}^{j} + \sum_{p=1}^{l}\sum_{h=1}^{r} \delta_{p,m+h}(u_i, v_i)\left(x_{pi} - K_{ph}\right)_{+}^{m} + \varepsilon_i \tag{13}$$

With m as order and h as knot points, an optimum knot point and the values of GCV, R-Squared, AIC and RMSE would be selected in order to obtain the best model as shown in Table 5.

As Table 5 shows, the best model is the NGWSR model with order $m = 1$ and 3 knot points, according to the R^2 value of to 99.298. This value indicates that such predictor variables as the number of doctors, the number of tourist attractions, the number of TB cases, the percentage of the elderly population, the percentage of households with access to proper drinking water, population density, GRDP, the number of hospitals and the number of villages/urban villages having public health center are able to account for 99.298% of the cumulative number of COVID-19 patients in Indonesia, while the remaining 0.702% is accounted for by other variables not studied.

The mathematical model of Nonparametric Geographically-Weighted Spline Regression of Equation (14), with an order $m = 1$ and a knot point $h = 3$ is as follows:

$$y_i = \beta_0(u_i, v_i) + \sum_{p=1}^{9}\sum_{j=1}^{1} \beta_{pj}(u_i, v_i) x_{pi}^{j} + \sum_{p=1}^{9}\sum_{h=1}^{3} \delta_{p,m+h}(u_i, v_i)\left(x_{pi} - K_{ph}\right)_{+}^{m} + \varepsilon_i \tag{14}$$

With total predictor variables of $p = 9$, namely x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, and x_9.

4.6. Estimation of Nonparametric Geographically-Weighted Spline Regression Parameters

Furthermore, the estimation stage of NGWSR is presented in Table 6.

Figure 4 shows that the estimator of response variable of the cumulative number of COVID-19 patients $\hat{Y}$ produced by NGWR has a value close to the original data (Y).

The red line is the estimator of the response variable of the cumulative number of COVID-19 patients in Indonesia using NGWR and the blue line is the observation data of the cumulative number of COVID-19 patients in Indonesia in 2021.

4.7. Significance Test of Simultaneous Parameters

The significance test of simultaneous parameters aims to determine the 9 predictor variables, including the number of doctors, the number of tourist attractions, the number of TB cases, the percentage of the elderly population, the percentage of households with access to proper drinking water, population density, GRDP, number of hospitals and the number of villages/urban villages having public health center which has a simultaneous effect on the cumulative number of COVID-19 patients in Indonesia. The test used Equation (11) at a significance level of $\alpha = 0.05$ and the results of ANOVA calculation are shown in Table 7.

Table 7 shows that $V(85.7079) > F_{(0.05;33;20)}$; thus, the predictor variables in the NGWSR model have a significant effect on the response variable.

Table 6
Estimator of cumulative number of COVID-19 patients in Indonesia

Provinces	Cumulative number of COVID-19 patients in Indonesia (y)	Cumulative number of COVID-19 patients in Indonesia ($\hat{y}$)
Aceh	10814	8177.9
North Sumatera	29198	30192
West Sumatera	36268	36404
Riau	42698	53339
Jambi	7407	16689
South Sumatera	20068	29209
Bengkulu	6630	10645
Lampung	15702	28651
Bangka Belitung Islands	12822	4908
Riau Islands	10838	3065.6
DKI Jakarta	406200	406300
West Java	277550	276510
Central Java	183030	184000
DI Yogyakarta	38703	32884
East Java	146810	146540
Banten	47101	49427
Bali	44236	48524
NTB	10291	12984
NTT	14200	13402
West Kalimantan	7503	13424
Central Kalimantan	19780	18126
South Kalimantan	32612	14746
East Kalimantan	68136	51186
North Kalimantan	11702	17766
North Sulawesi	15638	13660
Central Sulawesi	12125	20550
South Sulawesi	61419	46929
Southeast Sulawesi	10396	14923
Gorontalo	5355	8228.3
West Sulawesi	5443	5536.9
Maluku	7515	3503
North Maluku	414	-1082.6
West Papua	8879	7207.1
Papua	20308	19868

Table 7
ANOVA

Source of diversity	Number of Squares	Degree of freedom	Median square	F-count	$p - value$
Regression	2.377×10^{11}	33	7.2053×10^{9}	85.7079	1.1102×10^{-15}
Error	1.6814×10^{9}	20	8.4068×10^{7}		
Total	2.3945×10^{11}	53			

4.8. Significance Test of Partial Parameters

The significance test of partial parameters as expressed in Equation (11) was aimed at testing each predictor variable for each effect on the response variable of the cumulative number of COVID-19 patients in Indonesia. Results of the partial significance test and the grouping by province are presented in Table 8 and the distribution map is given in Figure 5.

As Figure 5 shows, there are 6 groups with their own influence factor. The largest number of provinces are in group 1 with factors affecting the cumulative number of COVID-19 patients include the number of doctors, the number of

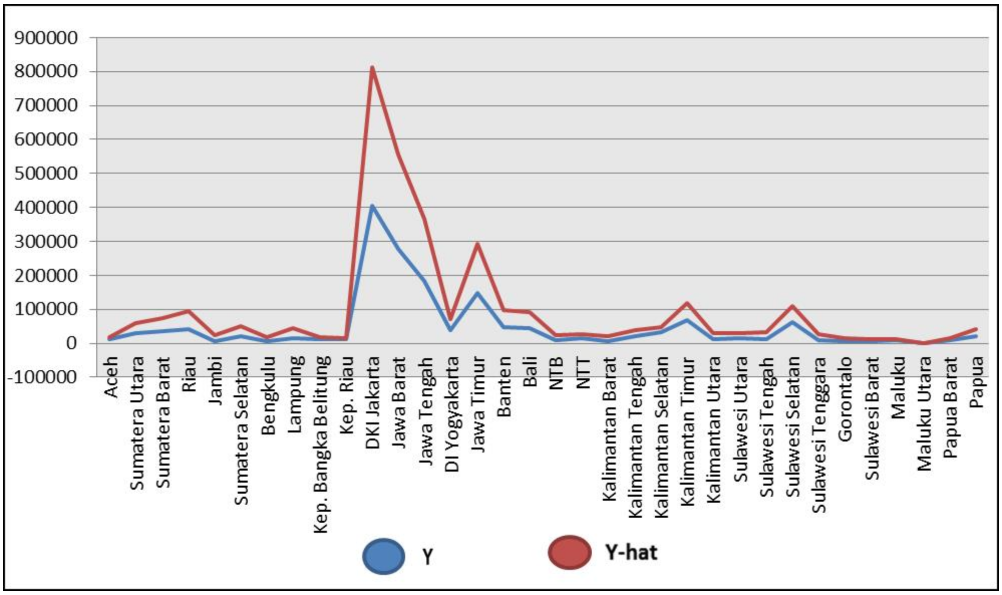


Figure 4: Distribution pattern of the cumulative number of COVID-19 patients in Indonesia

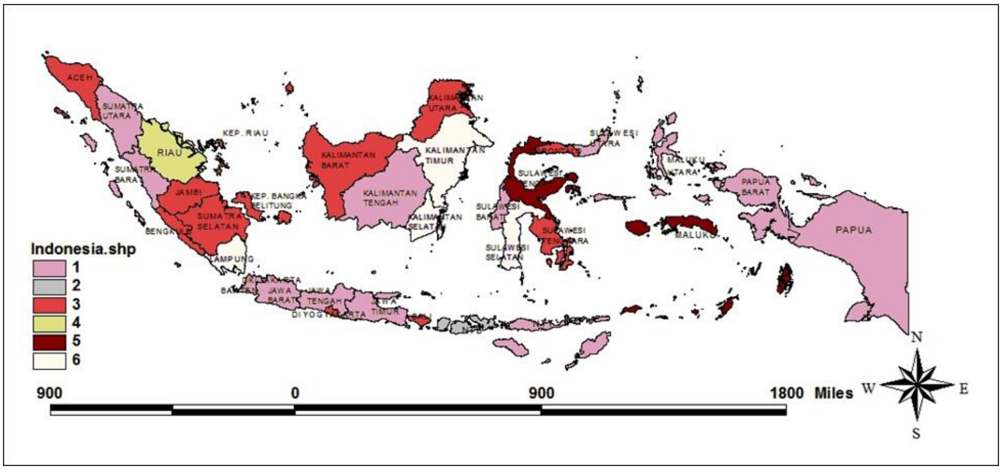


Figure 5: Map of the cumulative number of COVID-19 patients in Indonesia in 2021 based on the significant variables

tourist attractions, the number of TB cases, the percentage of the elderly population, the percentage of households with access to proper drinking water, population density, GRDP, the number of hospitals, and the number of villages/urban villages having public health center. The factors affecting the cumulative number of of COVID-19 patients based on the NGWSR model for each province are presented in Table 8.

Conclusions

The model developed in the present study uses the Nonparametric Geographically-Weighted Spline Regression (NGWSR) with GCV-NGWSR optimization that uses two bases of polynomial and truncated matrices. The novelty of the model lies on the significance test of simultaneous and partial parameters that support the NGWSR model. The variables with significant effects on the cumulative number of COVID-19 patients in Indonesia for 34 provinces are the number of doctors and health professionals, the number of tourist attractions, the number of TB cases, the percentage of the elderly population, the percentage of households with access to proper drinking water, population density, Gross

Table 8
Grouping of the cumulative number of COVID-19 patients in Indonesia in 2021 based on the significant variables

Groups	Provinces	Influence factors
1	North Sumatera West Sumatera DKI Jakarta West Java Central Java East Java Banten NTT Central Kalimantan North Sulawesi West Sulawesi North Maluku West Papua Papua	Number of doctors (x_1), number of tourist attractions (x_2), numbers of TB cases (x_3), percentage of the elderly population (x_4), percentage of households with access to proper drinking water (x_5), population density (x_6), GRDP (x_7), number of hospitals (x_8) and number of villages/urban villages having public health center (x_9)
2	NTB	Number of doctors (x_1), number of tourist attractions (x_2), numbers of TB cases (x_3), percentage of the elderly population (x_4), percentage of households with access to proper drinking water (x_5), population density (x_6), number of hospitals (x_8) and number of villages/urban villages having public health center (x_9)
3	Aceh Jambi South Sumatra Bengkulu Bangka Belitung Islands Riau Islands DI Yogyakarta Bali West Kalimantan North Kalimantan Southeast Sulawesi Gorontalo	Number of doctors (x_1), number of tourist attractions (x_2), numbers of TB cases (x_3), percentage of the elderly population (x_4), percentage of households with access to proper drinking water (x_5), population density (x_6), number of hospitals (x_8)
4	Riau	Number of doctors (x_1), number of tourist attractions (x_2), numbers of TB cases (x_3), percentage of the elderly population (x_4), percentage of households with access to proper drinking water (x_5), number of hospitals (x_8)
5	Central Sulawesi Maluku	Number of doctors (x_1), number of tourist attractions (x_2), percentage of the elderly population (x_4), percentage of households with access to proper drinking water (x_5), population density (x_6), number of hospitals (x_8)
6	Lampung South Kalimantan East Kalimantan South Sulawesi	Number of doctors (x_1), number of tourist attractions (x_2), numbers of TB cases (x_3), percentage of households with access to proper drinking water (x_5), and number of hospitals (x_8)

Regional Domestic Product. (GDP) at market price, which is the amount of gross added value arising from all economic sectors in a region, the number of hospitals, and the number of villages/urban villages having public health center. The NGWSR model produces a coefficient of determination or an accurate model goodness-of-fit of 99.298%. The present study contribute to the health sector and the government with regard to tracing on the basis of results of spatial mapping of the spread of COVID-19. Additionally, it contributes to the field of Geographic Information Systems (GIS).

Further studies are recommended to develop NGWSR models by adding up time parameters to complete local nonparametric regression analyses according to the time series for each observation location.

References

Anselin, L., 2003. An Introduction to Spatial Regression Analysis in R.

Arab-Mazar, Z., Sah, R., Rabaan, A., Dhama, K., Rodriguez-Morales, A., 2020. Mapping the incidence of the covid-19 hotspot in iran implications for travellers. Travel Med. Infect. Dis. doi:10.1016/j.tmaid.2020.101630.

Benassi, F., Naccarato, A., 2017. Households in potential economic distress. a geographically weighted regression model for italy, 2001-2011. Spatial Statistics 21, 362–376. doi:10.1016/j.spasta.2017.03.002.

[BPS], B.P.S., 2019a. Statistik Indonesia 2019. Badan Pusat Statistik, Jakarta Pusat.

[BPS], B.P.S., 2019b. Statistik Penduduk Lanjut Usia 2019. Badan Pusat Statistik, Jakarta Pusat.

[BPS], B.P.S., 2020a. Statistik Indonesia 2020. Badan Pusat Statistik, Jakarta Pusat.

[BPS], B.P.S., 2020b. Statistik Penduduk Lanjut Usia 2020. Badan Pusat Statistik, Jakarta Pusat.

[BPS], B.P.S., 2021a. Statistik Indonesia 2021. Badan Pusat Statistik, Jakarta Pusat.

[BPS], B.P.S., 2021b. Statistik Penduduk Lanjut Usia 2021. Badan Pusat Statistik, Jakarta Pusat.

Brunsdon, Fotheringham, 1999. Some notes on parametric significance test for geographically weighted regression. Journal of Regional Science 39, 497–524. doi:10.1111/0022-4146.00146.

Budiantara, I., Purnomo, J., 2011. infants' weight growth model in surabaya (indonesia) by using weighted spline regression. International Journal of Basic, Applied Sciences 11, 119–123.

Buzai, G., 2020. De wuhan a luján. Evolución espacial del COVID-19 3.

Chen, Z., Zhang, Q., Lu, Y., Guo, Z., Zhang, X., Zhang, W., Lu, J., 2020b. Distribution of the covid-19 epidemic and correlation with population emigration from wuhan, china. Chinese medical journal doi:10.1097/CM9.0000000000000782.

Crespo, R., Fotheringham, A., Charlton, M., 2007. Application of geograhically weighted regression to a 19-year set of house price data in london to calibrate local hedonic price models, in: Proceeding of the 9th International Conference on Geocomputation.

Demsar, Fotheringham, Charlton, 2008. Testing the importance of the explanatory variables in a mixed geographically weighted regression model. Exploring the spatio-temporal dynamics of geographical processes with geograhically weighted regression, geovisual analytics 7, 181–197.

Desjardins, M., Hohl, A., Delmelle, E., 2020. Rapid surveillance of covid-19 in the united states using a prospective space-time scan statistic: detecting and evaluating emerging clusters. Appl. Geogr. .

Fotheringham, A., Brundson, C., Charlton, M., 2002. Geographically Weighted Regression : The Analysis of Spatially Varying Relationships. John Wiley Sons Ltd, England.

Giuliani, D., Dickson, M., Espa, G., Santi, F., 2020. Modelling and predicting the spatio- temporal spread of coronavirus disease 2019 (covid-19) in italy. SSRN doi:10.2139/ssrn.3559569.

Guan, W., Ni, Z., Hu, Y., Liang, W., Ou, C., He, J., Du, B., 2020. Clinical characteristics of coronavirus disease 2019 in china. New England Journal of Medicine 382, 1708–1720.

Hardle, G., 1990. Applied Nonparametric Regression. Cambridge University Press, New York.

Huang, Wu, Barry, 2010. International journal geographical inference science. Geographically, temporally weighted regression for modeling spatio-temporal variation in house prices 24, 383–401.

Huang, H., Wang, Y., Wang, Z., Liang, Z., Qu, S., Ma, S., Liu, X., 2020. Clinical character- istics of coronavirus disease 2019 in china. China 3.

Kosfeld, R., 2006. Spatial econometric. http//www.scribd.com.

Leung, Y., Mei, C., Zhang, W., 2000a. Statistic test for spatial non stationarity based on the geographically weighted regression model. Journal Environment, Planing A 32, 9–32.

Leung, Y., Mei, C., Zhang, W., 2000b. Testing for spatial autocorrelation among the residuals of the geograhically weighted regression. Journal Environment, Planing A 32, 871–890.

Li, Y., Jiao, Y., Browder, J.A., 2016. Modeling spatially-varying ecological relationships using geographically weighted generalized linear model: A simulation study based on longline seabird bycatch. Fisheries Research 181, 14–24. doi:10.1016/j.fishres.2016.03.024.

Mei, Wang, Zhang, 2006. Testing the importance of the explanatory variables in a mixed geographically weighted regression model. Journal Environment, Planning 38, 587–598.

Nakaya, Fotheringham, Brunsdon, Charltron, 2005. Geographically weighted poisson regression for disease association mapping. Journal Statistics Medicine 24, 2695–2717.

Saha, A., Gupta, K., Patil, M., 2020. Monitoring and epidemiological trends of coronavirus disease (covid-19) around the world. Osfpreprints .

Satuan Gugus Tugas Penanganan COVID-19, 2021. Peta sebaran. www.covid19.go.id/peta-sebaran.

Sifriyani, Budiantara, I.N., Kartiko, S., Gunardi, 2019a. Evaluation of factors affecting increased unemployment in east java using ngwr-ts method. International Journal of Sciences: Basic, Applied Research 49, 123–142.

Sifriyani, Kartiko, S., Budiantara, I., Gunardi, 2017. Geographically weighted regression with spline approach. Far East Journal of Mathematical Sciences 101, 1183–1196.

Sifriyani, Kartiko, S.H., Budiantara, I.N., Gunardi, 2018. Development of nonparametric geographically weighted regression using truncated spline approach. Songklanakarin Journal of Science, Technology 40, 909–920.

Sifriyani, Rosadi, D., 2020. Pemodelan susceptible infected recovered (sir) untuk estimasi angka reproduksi covid-19 di kalimantan timur, samarinda. Media Statistika 13.

Sifriyani, Ruslan, Susanty, F., 2019b. Evaluation of forest productivity, governance on the preservation of tropical rain forests in kalimantan using the ngwr-ts nonparametric geospatial method. EurAsian Journal of BioSciences 13, 2373–2379.

Sifriyani, Ruslan, Susanty, F., 2019c. Mapping, analysis factors of affecting productivity tropical rain forests in east kalimantan. Modern Applied Science 13. doi:10.5539/mas.v13n10p112.

Utami, T.W., Haris, M.A., Prahutama, A., Purnomo, E.A., 2020. Optimal knot selection in spline regression using unbiased risk, generalized cross validation methods. Journal of Physics: Conference Series 1446, 012049.

Wrenn, D., Sam, A., 2014. Geographically, temporally weighted likelihood regression : Exploring the spatiotemporal determinants of land use change. Regional Science, Urban Economics 44, 60–74.

shan Wu, S., Yang, H., Guo, F., Han, R.M., 2017. Spatial patterns, origins of heavy metals in sheyang river catchment in jiangsu, china based on geographically weighted regression. Science of the Total Environment 580, 1518–1529. doi:10.1016/j.scitotenv.2016.12.137.

Xiong, Y., Guang, Y., Chen, F., Zhu, F., 2020. Spatial statistics and influencing factors of the novel coronavirus pneumonia 2019 epidemic in hubei province, china. ResearchSquare doi:10.21203/rs.3.rs-16858/v2.

Yu, D., 2010. Exploring spatiotemporally varying regressed relationships: the geographically weighted panel regression analysis. proceedings of the joint international conference on theory. Data Handling, Modeling in GeoSpatial Information Science , 134–139.

Yuniartika, Y., Kusnandar, D., Mara, M.N., 2013. Penentuan generalized cross validation (gcv) sebagai kriteria dalam pemilihan model regresi b-spline terbaik. Bultin Ilmiah Matematika, Statistika, Terapannya 2, 121–126.

Zhang, X., Rao, H., Wu, Y., Huang, Y., Dai, H., 2020. Comparison of the spatiotemporal characteristics of the covid-19 and sars outbreaks in mainland china. Geography and Sustainability doi:10.1101/2020.03.23.20034058.

Zhou, C., Su, F., Pei, T., Zhang, A., Du, Y., Luo, B., Song, C., 2020. Covid-19: challenges to gis with big data. Geography and Sustainability 1, 77–87.

Conflicts of Interest Statement

This statement is to certify that all Authors have seen and approved the manuscript being submitted. We warrant that the article is the Authors original work. We warrant that the article has not received prior publication and is not under consideration for publication elsewhere. On behalf of all Co-Authors, the corresponding Author shall bear full responsibility for the submission. This research has not been submitted for publication nor has it been published in whole or in part elsewhere. We attest to the fact that all Authors listed on the title page have contributed significantly to the work, have read the manuscript, attest to the validity and legitimacy of the data and its interpretation, and agree to its submission to the Journal Spatial and Spatio-temporal Epidemiology.

Best Regards,

Dr. Sifriyani, S.Pd., M.Si.

https://www.editorialmanager.com/sste/default1.aspx

Spatial and Spatio-temporal Epidemiology

Editorial Manager

HOME • LOGOUT • HELP • REGISTER • UPDATE MY INFORMATION • JOURNAL OVERVIEW
MAIN MENU • CONTACT US • SUBMIT A MANUSCRIPT • INSTRUCTIONS FOR AUTHORS • POLICIES

Role: Author Username: sifriyani

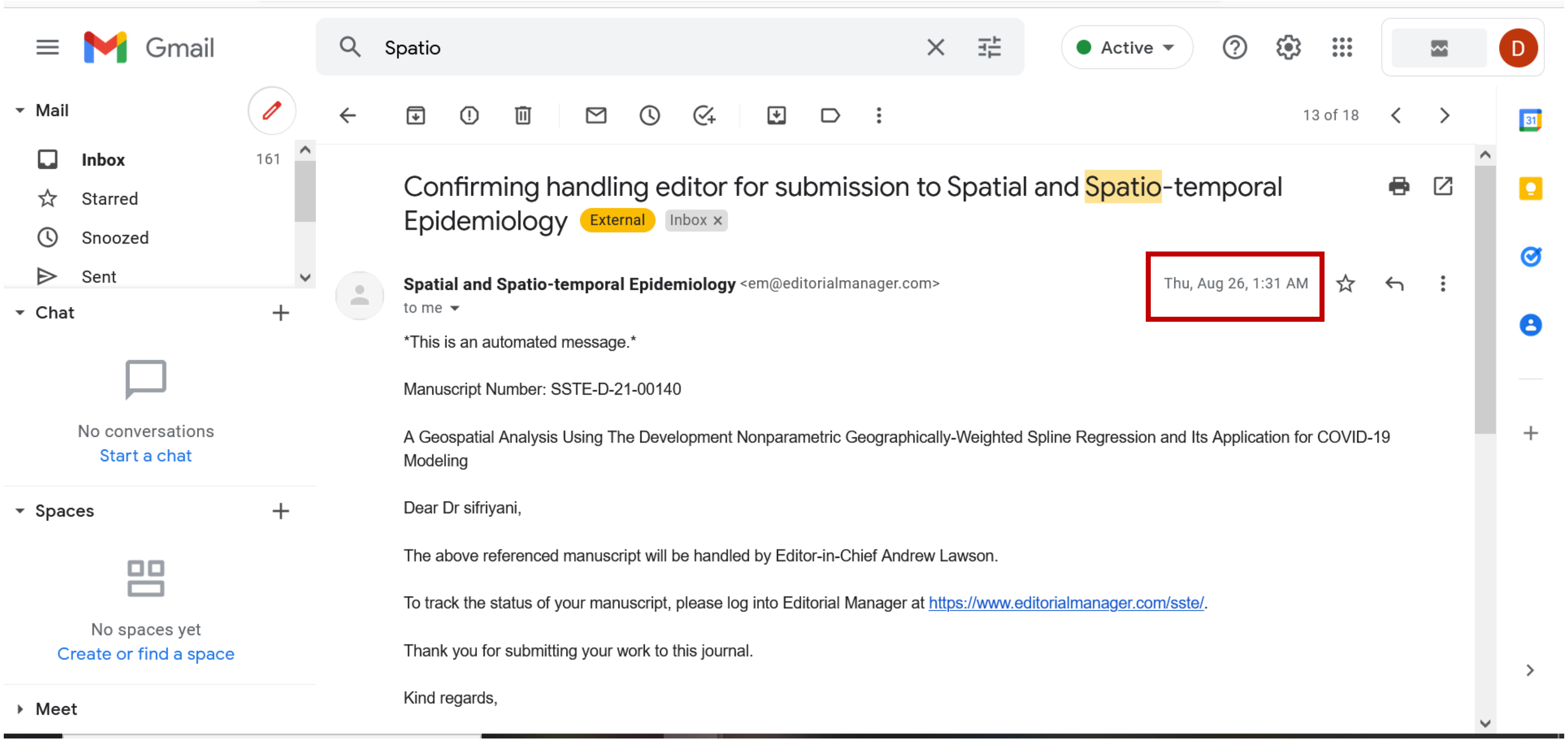
Submissions Being Processed for Author sifriyani sifriyani, M.Si

Page: 1 of 1 (1 total submissions) Display 10 results per page.

Action	Manuscript Number	Title	Initial Date Submitted	Status Date	Current Status
Action Links	SSTE-D-21-00140	A Geospatial Analysis Using The Development Nonparametric Geographically-Weighted Spline Regression and Its Application for COVID-19 Modeling	Aug 12, 2021	Oct 10, 2021	Under Review

Page: 1 of 1 (1 total submissions) Display 10 results per page.

<< Author Main Menu



Confirming handling editor for submission to Spatial and Spatio-temporal Epidemiology External Inbox ×

Spatial and Spatio-temporal Epidemiology <em@editorialmanager.com>
to me

Thu, Aug 26, 1:31 AM

This is an automated message.

Manuscript Number: SSTE-D-21-00140

A Geospatial Analysis Using The Development Nonparametric Geographically-Weighted Spline Regression and Its Application for COVID-19 Modeling

Dear Dr sifriyani,

The above referenced manuscript will be handled by Editor-in-Chief Andrew Lawson.

To track the status of your manuscript, please log into Editorial Manager at https://www.editorialmanager.com/sste/.

Thank you for submitting your work to this journal.

Kind regards,

https://www.editorialmanager.com/japg/default1.aspx

APPLIED GEOGRAPHY

HOME • LOGOUT • HELP • REGISTER • UPDATE MY INFORMATION • JOURNAL OVERVIEW
MAIN MENU • CONTACT US • SUBMIT A MANUSCRIPT • INSTRUCTIONS FOR AUTHORS • POLICIES

Editorial Manager

Role: Author Username: Sifriyani

Submissions Being Processed for Author Sifriyani Sifriyani, Ph.D

Page: 1 of 1 (1 total submissions) Display 10 results per page.

Action	Manuscript Number	Title	Initial Date Submitted	Status Date	Current Status
View Submission View Reference Checking Results Send E-mail	JAPG-D-21-00974	Tuberculosis Cases and Population Density Affect the Increase in COVID-19 Outbreak: A Spatiotemporal Analysis of Geographically and Temporally Weighted Regression	Sep 25, 2021	Oct 04, 2021	under review

Page: 1 of 1 (1 total submissions) Display 10 results per page.

<< Author Main Menu

HighLights

Tuberculosis Cases and Population Density Affect the Increase in COVID-19 Outbreak: A Spatiotemporal Analysis of Geographically and Temporally Weighted Regression

Sifriyani, Dedi Rosadi, Mariani Rasjid

1. Application of geographic theory and methodology in resolution of the increase in positive cases of COVID-19 from the perspective of Spatial and Time in Kalimantan Province.
2. Inovation of Geographically Temporally Weighted Regression model with the development of spatial and time interaction distance functions.
3. This study develops a new model in spatial statistics, offers new methodologies in the fields of Geographic Models and Geographic Information Systems (GIS).
4. This study uses national COVID-19 data from 56 districts/cities until August 2021.
5. The results of the study provide the local governments with decision-making recommendations for overcoming the COVID-19 problems in their respective regions.

Tuberculosis Cases and Population Density Affect the Increase in COVID-19 Outbreak: A Spatiotemporal Analysis of Geographically and Temporally Weighted Regression

Sifriyani[1], Dedi Rosadi[2], Mariani Rasjid[3]

[1]Statistics Study Program, Department of Mathematics, Faculty of Mathematics and Natural Sciences, Mulawarman University, Samarinda, Indonesia
[2]Department of Mathematics, Faculty of Mathematics and Natural Sciences, Gadjah Mada University, Yogyakarta, Indonesia
[3]Pulmonologist, Tadulako General Hospital, Indonesia
[3]Faculty of Medicine, Tadulako University, Kendari, Indonesia

Abstract

COVID-19 cases in Indonesia are increasing and spreading in all provinces, the severity is also occurring in Kalimantan. Application of geographic theory and methodology in resolution of the increase in positive cases of COVID-19 from the perspective of Spatial and Time in Kalimantan Province. The present study uses a spatiotemporal analysis to determine the factors affecting the constantly increase in COVID-19 cases in Kalimantan. The spatiotemporal analysis uses the geographically temporally weighted regression (GTWR) model by developing a spatial and temporal interaction distance function. The GTWR model was applied to data on the number of positive COVID-19 cases at a scale of 56 districts/cities on the Indonesian island of Kalimantan in the period of January 2020 to August 2021. The purposes of the present study was to determine the factors affecting the cumulative increase in COVID-19 cases in Kalimantan and to map the spatial distribution for 56 districts/cities based on the significant predictor variables. Results of the study showed that the GTWR model with the development of a spatial and temporal interaction distance function using the kernel Gaussian fixed bandwidth function is the best model. Factors affecting the positive cases of COVID-19 in Kalimantan for each district/city were varying, based on the significant variables, including the number of doctors, the number of TB cases, the percentage of the elderly population, GRDP, and the number of hospitals. The highest influence that affects COVID-19 cases is the high number of TB cases, population density, and the lack of health services. Furthermore, an area map was produced on the basis of the significant variables affecting the increase in COVID-19 cases. The results of the study provide the local governments with decision-making recommendations for overcoming the COVID-19 problems in their respective regions.

Keywords: Pandemic. Epidemiology COVID-19, Spatio Temporal, Spatial Analysis, Mapping GIS, Applied Geography

1. Introduction

Spatial data are modeled using the spatial regression with geographic weighting or known as the geographically-weighted regression (GWR) model. GWR was first introduced by Fotheringham in 1967 (Fotheringham et al., 2002). The response variable in the GWR model contains predictor variables, each of its regression coefficients being depending on where the data are observed. The GWR uses the point approach, in which each parameter value is estimated at each observation location so that each observation point has a different parameter value

(Fotheringham, et al., 2002). The GWR model is a development of the linear regression analysis into a regression with the addition of geographic weighting for each regression parameter due to spatial variability. Studies using the GWR theory were undertaken by Brunsdon et al. (1999), Crespo et al. (2007), Leung et al. (2000a) and Leung et al. (2000b).

The weakness of the GWR model is that it only uses spatial (location) data at one time, while spatial data are usually affected by time series. Spatial data involving multiple time (temporal) observations are important in spatial analysis to make it possible to estimate parameters more accurately (Fotheringham et al., 2015). Therefore, in order to increase the precision of the parameter estimators in the GWR model, observations are made for each location at a certain time. To overcome the weakness of the GWR model, a geographically temporally weighted regression (GTWR) model was developed (B. Huang et al., 2010) taking into account the elements of location and time. The GTWR model constitutes a development of the GWR model to deal with the non-stationary data, both spatially and temporally, at the same time (Wang P et al., 2006). The advantage of the GTWR model is that the model produced is local for each location and time, making the model more representative (B. Huang et al., 2010). Spatial and temporal information in the GTWR model is an important element to construct a weighting matrix that it is expected to be capable of identifying spatial and temporal variability. The GTWR model has been widely applied to a variety of fields. According Fotheringham et al. (2015), the GTWR model is generally used in relation to the issues of the spread of infectious diseases, water pollution, hydrology and city planning (Fotheringham et al., 2015). The present study used the GTWR model in relation to the spread of COVID-19. COVID-19 has been spreading in various countries around the world, including Indonesia.

Indonesia is an archipelagic country, consisting of various large and small islands. Kalimantan represents the largest island in Indonesia which consists of 5 provinces: East Kalimantan, North Kalimantan, South Kalimantan, Central Kalimantan and West Kalimantan Provinces, with an increasing daily spread of COVID-19. Based on data from the official COVID-19 website of the 5 provinces on the island of Kalimantan showed that the highest cumulative number of positive COVID-19 cases as of 10 August 2021 was East Kalimantan Province with 133,826 cases (www.covid19. kaltimprov.go.id), followed by South Kalimantan Province with 55,257 cases (www.covid19.kalselprov.go.id), Central Kalimantan Province with 38,123 cases (www. covid19.kaltengprov.go.id), North Kalimantan Province with 26,050 cases (www.covid19.kaltaraprov.go.id), and West Kalimantan Province with 17,999 cases (www. .covid19.kalbarprov.go.id). Based on this, it is importantly reasonable to undertake a study to obtain factors causing such an increase from a spatial and temporal point of view. The present study contribute the local governments with regard to overcoming the increase in COVID-19 cases in their respective regions.

Studies that used the spatial and temporal analysis of confirmed cases of COVID-19 include Geographical characteristics of those infected by COVID-19 region of China (Guan et al., 2020), Distribution of the contagion cases and their correlation with the Emigration of the Wuhan population in the initial stage of the epidemic region China (Chen et al., 2020; Z. L. Chen et al., 2020), Spatiotemporal analysis of COVID-19 and its relationship with epidemiological characteristics, control of measures taken and their effects region China (H. Huang et al., 2020), Spatiotemporal analysis of COVID-19 at the national and provincial levels region (Arab-Mazar et al., 2020), Spatiotemporal analysis of COVID-19 at the national and provincial levels region Italy (Giuliani et al., 2020), Reflections on the use of GIS with big data and spatiotemporal analysis of COVID-19 region China (Zhou et al., 2020), Spatiotemporal analysis of COVID-19 at the national and provincial levels region South Korea (Rezaei et al., 2020), Comparison of spatiotemporal evolution between COVID-19 and SARS 2003 region China (Zhang et al., 2020), Spatiotemporal analysis of COVID-19 in Rio Grande do Sul region Brazil (Dagnino et al., 2020), Spatiotemporal analysis and call for the need to do more GIS studies locally region Iran (Ahmadi et al., 2020; M. Ahmadi et al., 2020).

COVID-19 modeling studies using spatiotemporal analysis include Pearson's correlation methods for spatiotemporal analysis region China (Xiong et al., 2020), Levy's flight to explain the spatiotemporal dynamics of the pandemic region China (Gross et al., 2020), Prospective space–time statistics to identify active and emerging COVID-19 groups at the county level region USA (Desjardins et al., 2020), Online questionnaire at the geographical identification of possible symptomatics region Israel (Rossman et al., 2020). Predicting the global spread of COVID-19 based on geographic and climatic data region Caribe Basin (de Ángel Solá et al., 2020), Geographical characteristics and spatiotemporal analysis of infections region USA (Team, 2020), Analysis by provinces of the effectiveness of quarantine on the spread of the pandemic Spain (Orea & Alvarez, 2020), Spatiotemporal analysis of COVID-19 at the national and provincial levels region India (Murugesan et al., 2020), Poisson segmented model for the analysis of changing patterns in different geographic areas region China (Tang et al., 2020), Spatiotemporal analysis of COVID-19 in Estado da Bahia region Brasil (Dagnino et al., 2020), Spatiotemporal analysis and reflections on health geography region Argentina (Buzai, 2020), Spatiotemporal analysis of COVID-19 at the national and provincial levels region Mexico (Santana Juárez, 2020), and Spatiotemporal analysis and reflections on the usefulness of GIS in the pandemic region World (Saha et al., 2020), and Susceptible Infected Recovered (SIR) Model For Estimating Covid-19 Reproduction Number in East Kalimantan and Samarinda (Sifriyani & Rosadi, 2020).

Based on the above description, it is urgent to Application of geographic theory and methodology in resolution of the increase in positive cases of COVID-19 from the perspective of Spatial and Time in Kalimantan Province, undertake a spatiotemporal analysis to model the cumulative data of COVID-19 in the area of Kalimantan Island, Indonesia. The spatiotemporal analysis used the geographically temporally weighted regression (GTWR) model with the development of a model distance function. The present study used data of the accumulative COVID-19 cases in Kalimantan, number of doctors, number of TB cases, percentage of elderly population, population density, Gross Regional Domestic Product as an economic indicator of an area, number of hospitals, number of public health centers, and percentage of poor population. Data were collected for each regency/city in Kalimantan.

The purposes of the present study were, first, to determine the factors affecting the cumulative increase in COVID-19 cases at the regency/city scale in Kalimantan, Indonesia, by using the GTWR model. Second, it mapped the spatial distribution for 56 regencies/cities based on significant predictor variables. One novelty of the present study is the development of the GTWR model with the distance function and the application of the model to COVID-19 data. The second novelty is the use of the latest COVID-19 data (2020 and 2021) of 56 districts/cities on the island of Kalimantan, Indonesia.

2. Material and Methods

2.1 Geographically and Temporally Weighted Regression

The Geographically and Temporally Weighted Regression (GTWR) model represents an effective approach to dealing with the problem of spatial and temporal non-stationarity (Huang et al., 2010). The GTWR model is a development of the GWR model by adding the time (temporal) element. In contrast to the GWR model, GTWR combines temporal and spatial information in a weighted matrix to identify spatial and temporal variability. The GTWR model in Equation (1) is for the independent variable p with the response variable at the location (u_i, v_i, t) for each observation.

$$y_i = \beta_0(u_i, v_i, t_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i, t_i)\, x_{ik} + \varepsilon_i \quad (1)$$

where y_i is the observed value of the response variable for the observation location (u_i, v_i) and time t_i, the parameter $\beta_0(u_i, v_i, t_i)$ is the constant of the intercept value, the parameter $\beta_k(u_i, v_i, t_i)$ is the regression coefficient of the *k*-th independent variable at the observation location (u_i, v_i) and time t_i, the variable x_{ik} is the observed value of the *k*-th explanatory variable at the observation location (u_i, v_i) and time t_i and ε_i is error the *i*-th observations which are assumed to be identical, independent, and $\varepsilon_i \sim N(0, \sigma^2)$.

2.2 GTWR Model Parameter Estimation

The regression coefficient $\hat{\beta}_i(u_i, v_i, t_i)$ at the *i*-th point can be obtained by using the Weighted Least Square. The estimated parameters of the GTWR model are given in Equation (2).

$$\widehat{\boldsymbol{\beta}}(u_i, v_i, t_i) = [\boldsymbol{X}^T \boldsymbol{W}(u_i, v_i, t_i)\boldsymbol{X}]^{-1} \boldsymbol{X}^T \boldsymbol{W}(u_i, v_i, t_i)\boldsymbol{y} \quad (2)$$

where the weight $\boldsymbol{W}(u_i, v_i, t_i) = diag(w_{i1}, w_{i2}, \ldots, w_{in})$ is the weighting matrix at the observation location (u_i, v_i) and time t_i. The diagonal element $w_{ij}(1 \leq j \leq n)$ is the spatiotemporal distance function at the observation point (u_i, v_i, t_i). In the modeling stage, it is assumed that the proximity of the data observation point to the *i* point in the spatiotemporal coordinate system has a greater effect on the parameter estimator $\widehat{\boldsymbol{\beta}}(u_i, v_i, t_i)$ than that of the data located further from the *i* point. The proximity has two elements, spatial proximity and temporal proximity; thus, the definition and measurement of spatiotemporal proximity in the coordinate system constitute major problems in the construction of the GTWR model.

The present study used date located at three dimensions in the spatiotemporal coordinate system and it is known that the observations are close to the *i* point. Therefore, B. Huang et al. (2010) used an ellipsoidal coordinate system to measure the proximity of the regression point to the observation points that surrounds it. An illustration of the spatiotemporal distance is shown in Figure 1.

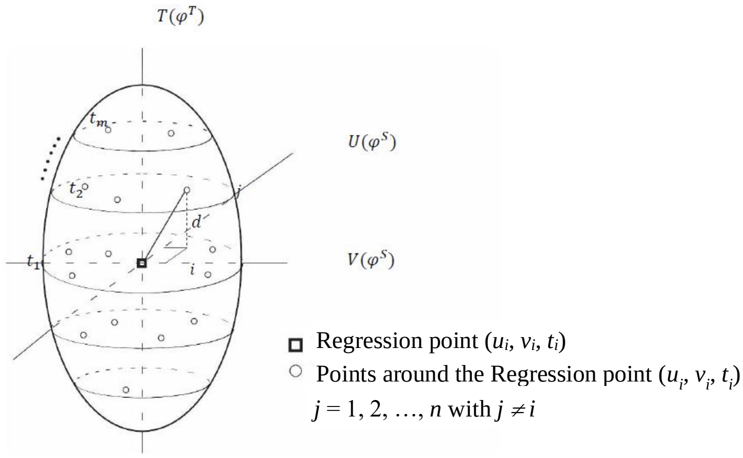


Figure 1. Illustration of spatiotemporal distance (B. Huang et al., 2010)

2.3 Distance Function and Geographical Weight of GTWR Model

The spatiotemporal distance function consists of a combination of the spatial distance function and the temporal distance function, which are given as follows (B. Huang et al., 2010; Wu et al., 2014):

$$\begin{cases} \left(d_{ij}^{S}\right)^2 = \left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2 \\ \left(d_{ij}^{T}\right)^2 = \left(t_i - t_j\right)^2 \\ \left(d_{ij}^{ST}\right)^2 = \varphi^S \left[\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2\right] + \varphi^T \left[\left(t_i - t_j\right)^2\right] \end{cases} \tag{3}$$

with φ^S and φ^T are the affecting factors that balance the different effects used to measure the spatiotemporal distance. Based on the distance function in Equation (3), the geographical weighting function according to Equation (4) is obtained.

$$\begin{aligned} w_{ij} &= \exp\left\{-\left(\frac{\varphi^S \left[\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2\right] + \varphi^T \left[\left(t_i - t_j\right)^2\right]}{h_{ST}^2}\right)\right\} \\ &= \exp\left\{-\left(\frac{\left[\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2\right]}{h_S^2} + \frac{\left[\left(t_i - t_j\right)^2\right]}{h_T^2}\right)\right\} \end{aligned} \tag{4}$$

The value of $h_S^2 = \frac{h_{ST}^2}{\varphi^S}$ and $h_T^2 = \frac{h_{ST}^2}{\varphi^T}$, then Equation (5) is obtained.

$$\begin{aligned} w_{ij} &= \exp\left\{-\left(\frac{\left(d_{ij}^{S}\right)^2}{h_S^2} + \frac{\left(d_{ij}^{T}\right)^2}{h_T^2}\right)\right\} \\ &= \exp\left\{-\left(\frac{\left(d_{ij}^{S}\right)^2}{h_S^2}\right)\right\} \times \exp\left\{-\left(\frac{\left(d_{ij}^{T}\right)^2}{h_T^2}\right)\right\} \\ &= w_{ij}^{S} \times w_{ij}^{T} \end{aligned} \tag{5}$$

where $w_{ij}^{S} = \exp\left\{-\left(\frac{\left(d_{ij}^{S}\right)^2}{h_S^2}\right)\right\}$ and $w_{ij}^{T} = \exp\left\{-\left(\frac{\left(d_{ij}^{T}\right)^2}{h_T^2}\right)\right\}$

h_S is a parameter of the spatial window width, h_T is a parameter of the temporal window width, h_{ST} is a parameter of the spatial-temporal window width.

In most cases, the value of φ^S and φ^T is not equal to zero. Let τ the ratio parameter of $= \frac{\varphi^T}{\varphi^S}$ with $\varphi^S \neq 0$; then, Equation (6) (Liu et al., 2017) is obtained.

$$\frac{\left(d_{ij}^{ST}\right)^2}{\varphi^S} = \left[\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2\right] + \tau\left[\left(t_i - t_j\right)^2\right] \quad (6)$$

Let $\varphi^S = 1$ in order to reduce unknown parameters. In this problem, there is only one unknown parameter, τ. The parameter τ serves to increase or decrease the effect of temporal distance on spatial distance. This parameter is obtained from the minimum cross-validation criteria by initializing the initial τ value as given in Equation (7).

$$(\tau) = \sum_i \left(y_i - \hat{y}_{\neq i}(\tau)\right)^2 \quad (7)$$

The Gaussian kernel function is a weighting function which is used in the GTWR model as given in Equation (8).

$$w_{ij} = exp\left(-\left(\frac{d_{ij}^{ST}}{h_{ST}}\right)^2\right) \quad (8)$$

The weighting matrix is $\boldsymbol{W}_{ij}$ determined by the spatiotemporal distance (d_{ij}^{ST}) and the window width h_{ST}. The window width value can be calculated using the geographic weighted regression model as proposed by Fotheringham et al. (2002). The estimator value of the response variable is determined by Equation (9).

$$\hat{y} = \begin{bmatrix}\hat{y}_1\\ \hat{y}_2\\ \cdots\\ \hat{y}_n\end{bmatrix} = \begin{bmatrix} X_1^T(X^TW(u_1,v_1,t_1)X)^{-1}X^TW(u_1,v_1,t_1)\\ X_2^T(X^TW(u_2,v_2,t_2)X)^{-1}X^TW(u_2,v_2,t_2)\\ \cdots\\ X_n^T(X^TW(u_n,v_n,t_n)X)^{-1}X^TW(u_n,v_n,t_n)\end{bmatrix} y = Sy \quad (9)$$

The selection of the model's goodness of fit can be calculated using the AIC (Akaike information criterion) value. The corrected AIC value (Liu et al., 2017) is used to overcome the spatiotemporal variability as given in Equation (10).

$$AIC = 2\,n\ln(\hat{\sigma}) + n\ln(2\pi) + n\left(\frac{n + tr(S)}{n - 2 - tr(S)}\right) \quad (10)$$

2.4 RESEARCH METHODOLOGY

2.4.1 Data and Data Sources

Data and data sources are described in Table 1.

Table 1. Description of Research Variables and Data Sources

Variable	Symbol	Variable Description	Observation Data Source	Unit	Scale
Response	y	Cumulative positive cases of COVID-19	Official websites www.covid19.kaltimprov.go.id www.coronainfo.kaltaraprov.go.id www.corona.kalselprov.go.id www.corona.kalteng.go.id www.corona.kalbarprov.go.id	People	56 regencies/ cities in the Island of Kalimantan
Predictor	x_1	Number of doctors	Statistics Indonesia of East Kalimantan Province, North Kalimantan Province, South Kalimantan Province, Central Kalimantan Province, West Kalimantan Province, 2020-2021	People	56 regencies/ cities in the Island of Kalimantan
	x_2	Number of TB cases	Public Health Office of East Kalimantan Province, North Kalimantan Province, South Kalimantan Province, Central Kalimantan Province, West Kalimantan Province, 2020-2021	Cases	56 regencies/ cities in the Island of Kalimantan
	x_3	Percentage of elderly population	Statistics Indonesia of East Kalimantan Province, North Kalimantan Province, South Kalimantan Province, Central Kalimantan Province, West Kalimantan Province, 2020-2021	Percentage	56 regencies/ cities in the Island of Kalimantan
	x_4	Population density	Statistics Indonesia of East Kalimantan Province, North Kalimantan Province, South Kalimantan Province, Central Kalimantan Province, West Kalimantan Province, 2020-2021	People/Km^2	56 regencies/ cities in the Island of Kalimantan
	x_5	Gross Regional Domestic Product at Market Price	Statistics Indonesia of East Kalimantan Province, North Kalimantan Province, South Kalimantan Province, Central Kalimantan Province, West Kalimantan Province, 2020-2021	Billion Rupiah	56 regencies/ cities in the Island of Kalimantan
	x_6	Number of hospitals	Public Health Office of East Kalimantan Province, North Kalimantan Province, South Kalimantan Province, Central Kalimantan Province, West Kalimantan Province, 2020-2021	Units	56 regencies/ cities in the Island of Kalimantan
	x_7	Number of villages/ kelurahan with public health centers	Public Health Office of East Kalimantan Province, North Kalimantan Province, South Kalimantan Province, Central Kalimantan Province, West Kalimantan Province, 2020-2021	Units	56 regencies/ cities in the Island of Kalimantan

	x_8	Percentage of poor population	Statistics Indonesia of East Kalimantan Province, North Kalimantan Province, South Kalimantan Province, Central Kalimantan Province, West Kalimantan Province, 2020-2021	%	56 regencies/ cities in the Island of Kalimantan

2.4.2 Stages of Analysis

Based on the objectives of the study, the following is the stages of data analysis using the GTWR model to determine factors affecting the cumulative increase in COVID-19 cases and to map the spatial distribution based on significant predictor variables.

1. To explore the distribution of response variables and predictor variables for the period of 2020-2021 using a spatial distribution mapping.
2. To describe cumulative data of COVID-19 cases and the predictor variables.
3. To perform a multicollinearity test by taking the value of VIF (variance inflation factor) into account.
4. To explore temporal variability using a boxplot of response variables for each year.
5. To perform an analysis using the GTWR method as follows:
   a. To calculate the optimum spatial bandwidth (h_S) using cross-validation based on the GWR optimization approach with the formula as given by Equation (11).

$$CV(h_S) = \sum_i \left(y_i - \hat{y}_{\neq i}(h_S)\right)^2 \quad (11)$$

   b. To calculate the optimum spatiotemporal ratio parameter (τ) using cross-validation based on the GTWR optimization approach with the formula as given by Equation (7).
   c. To calculate parameters φ^S and φ^T using the cross-validation approach with the formula given in point b. Both parameters are based on the spatiotemporal distance function with the formula as given by Equation (12)

$$(d_{ij}^{ST})^2 = \varphi^S[(u_i - u_j)^2 + (v_i - v_j)^2] + \varphi^T(t_i - t_j)^2 \quad (12)$$

   d. To determine the weighting matrix (W) using the spatiotemporal distance measure for each observation location based on the Gaussian kernel function with the formula given by Equation (8).
6. To estimate parameters in the GTWR model at each location using the weighted least square (WLS) according to Equation (2).
7. To perform a parameter significance test for the GTWR model.
8. To map the variable significance for each region.

3 RESULTS AND DISCUSSION

This section begins with providing information on descriptive statistics, especially the presentation of data using a spatial distribution map, followed by information on the measure of concentration and the measure of data distribution. The results of statistical inference research begin with regression analysis and a test of spatial effects, GTWR modeling, GTWR model estimation, GTWR model significance test and spatial mapping based on GTWR model results. The analysis to determine factors affecting the increase in confirmed COVID-19 cases is based on the regency/city scale on the island of Kalimantan, Indonesia.

3.1 Spatial Distribution Mapping

The observation data in Table 1 are subjected to a descriptive statistical analysis and statistical inference. Observation data are categorized based on variables and are described in Figures 2 to 10.

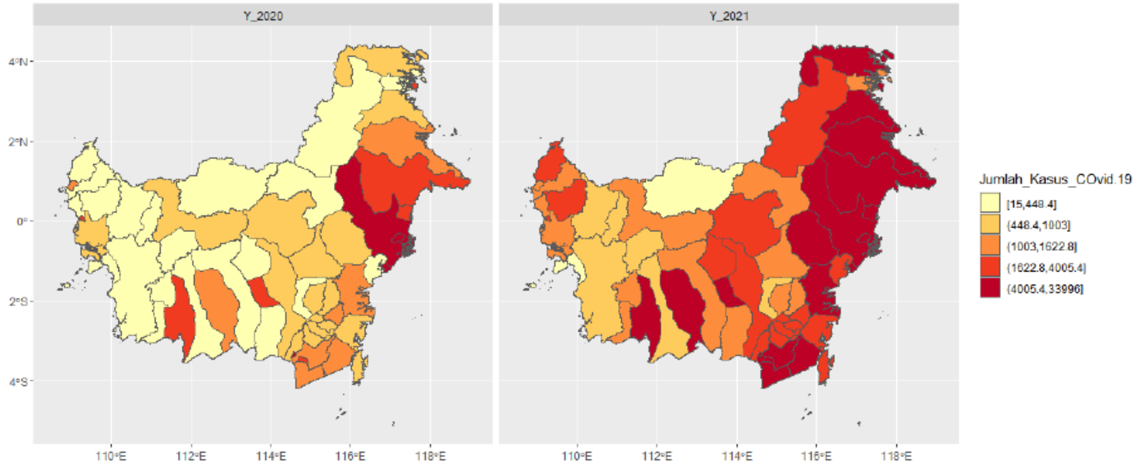


Figure 2. Map of the number of confirmed positive COVID-19 cases 2020-2021

Figure 2 shows that the number of confirmed positive COVID-19 cases in 2020 spread evenly across the regencies/cities in Kalimantan, as indicated by the similar distribution colors. However, in comparison, in 2020 the number of confirmed cases was relative scant, while the number of positive COVID-19 cases increased 2021. This is clearly shown by the color changes for regencies/cities in East Kalimantan Province.

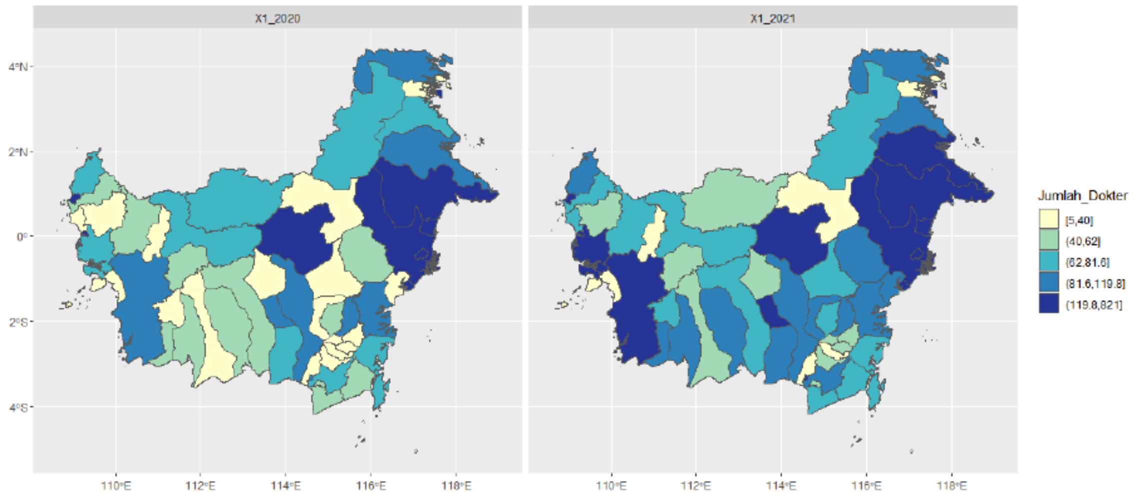


Figure 3. Map of the number of doctors 2020-2021

Figure 3 shows the distribution of the number of doctors by the scale of regency/city in Kalimantan 2020-2021. Most regencies/cities in Kalimantan had a little and evenly distributed number of doctors. In 2021 the number of doctors increased quite significantly in Samarinda and Balikpapan. This is shown by the dark color contrast, indicating quite a lot of the number of doctors in those areas. This is commensurate with the increasing number of COVID-19 cases.

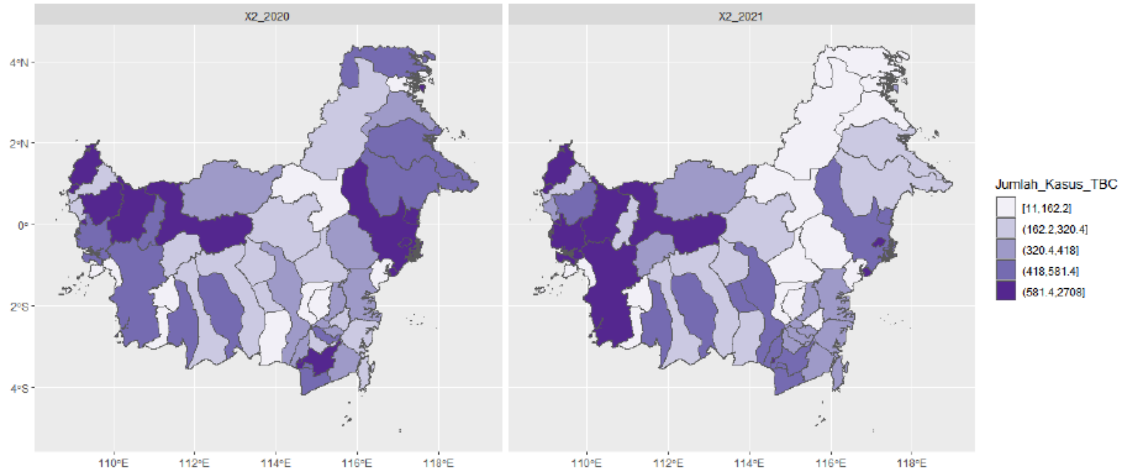


Figure 4. Map of the number of TB cases 2020-2021

Figure 4 shows the distribution of the number of TBC cases in 2020- 2021. In general, there is no increase in the number of TBC cases in each regency/city in Kalimantan. This is shown by the almost similar and evenly distributed color pattern in each area. However, the City of Banjarmasin has a higher TB cases than that of other regencies/cities, as indicated by its darker color.

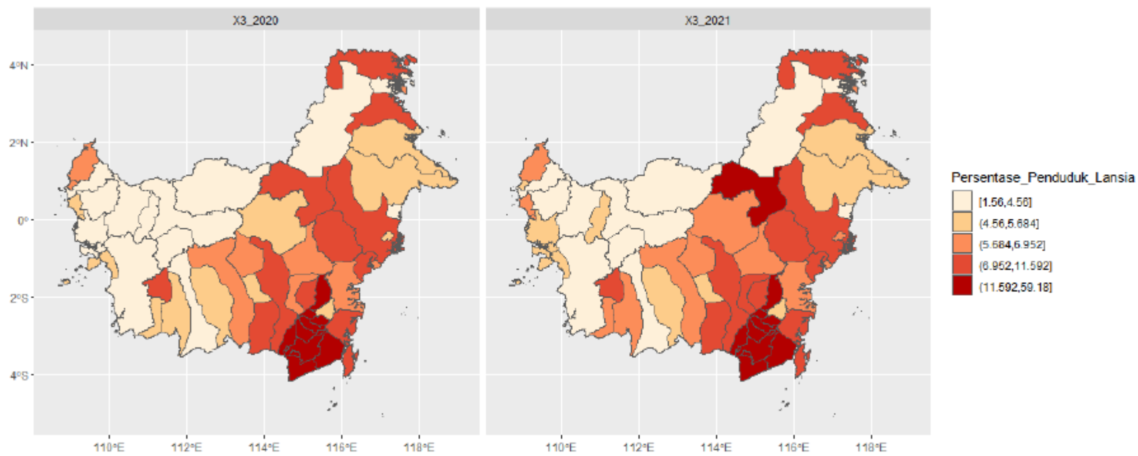


Figure 5. Map of the Percentage of the Elderly Population in 2020-2021

Figure 5 shows the distribution of the percentage of the elderly population in 2020-2021. In general, there is no increase in the percentage of the elderly population in every regency/city in Kalimantan. This is shown by the almost similar and evenly distributed color pattern in each area. However, some areas, such as Kutai Kertanegara and Kutai Timur, have a higher percentage of the elderly population than that of other regencies/cities, as indicated by the darker color.

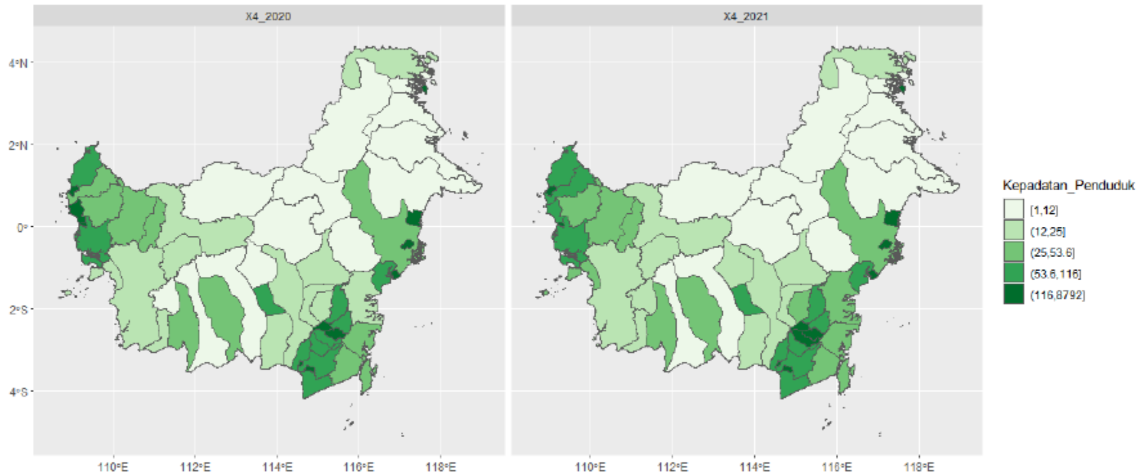


Figure 6. Map of population density in 2020-2021

Figure 6 shows the distribution of population density in 2020-2021. In general, there is no increase in population density in every regency/city in Kalimantan. This is shown by the almost similar and evenly distributed color pattern in each area. However, the City of Samarinda has a higher population density than that of other regencies/cities, as indicated by the darker color. Furthermore, the City of Samarinda has a decrease in population density from 2020 to 2021, as indicated by the color difference which is getting brighter.

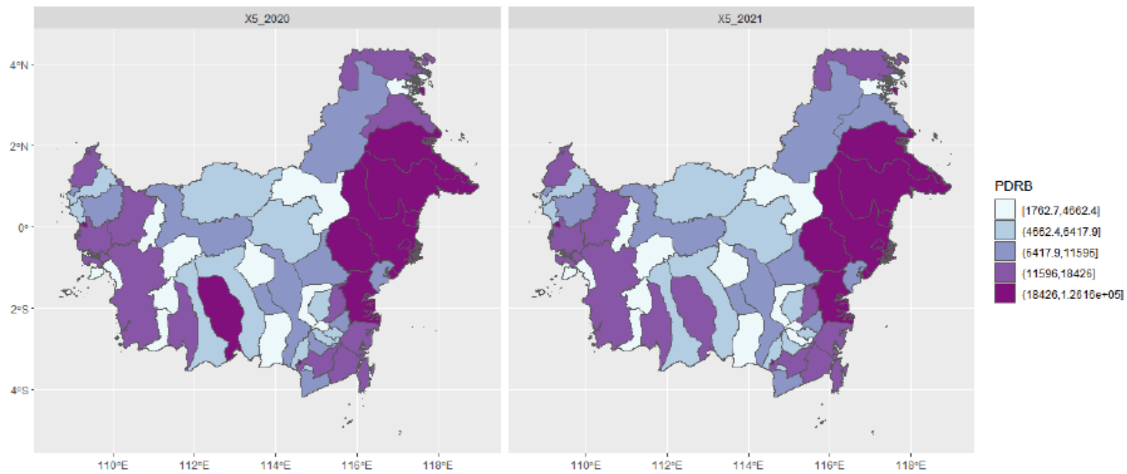


Figure 7. Map of Gross Regional Domestic Product (GRDP) 2020-2021

Figure 7 shows the distribution of GDP in 2020-2021. In general, there is no increase in GRDP in each regency/city in Kalimantan. This is shown by the almost similar and evenly distributed color pattern in each area. However, most areas have higher GRDP values, as indicated by the dark color.

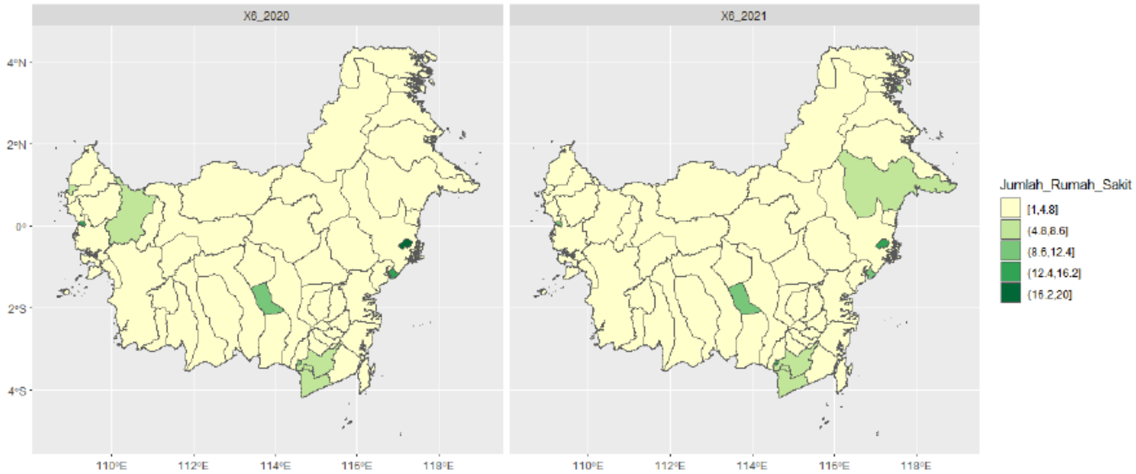


Figure 8. Map of the number of hospitals in 2020-2021

Figure 8 shows the distribution of the number of hospital in 2020-2021. In general, there is no increase in the number of hospitals in each regency/city in Kalimantan. This is shown by the almost similar and evenly distributed color pattern in each area. However, the City of Banjarmasin has more hospitals than other regencies/cities, as indicated by the darker color.

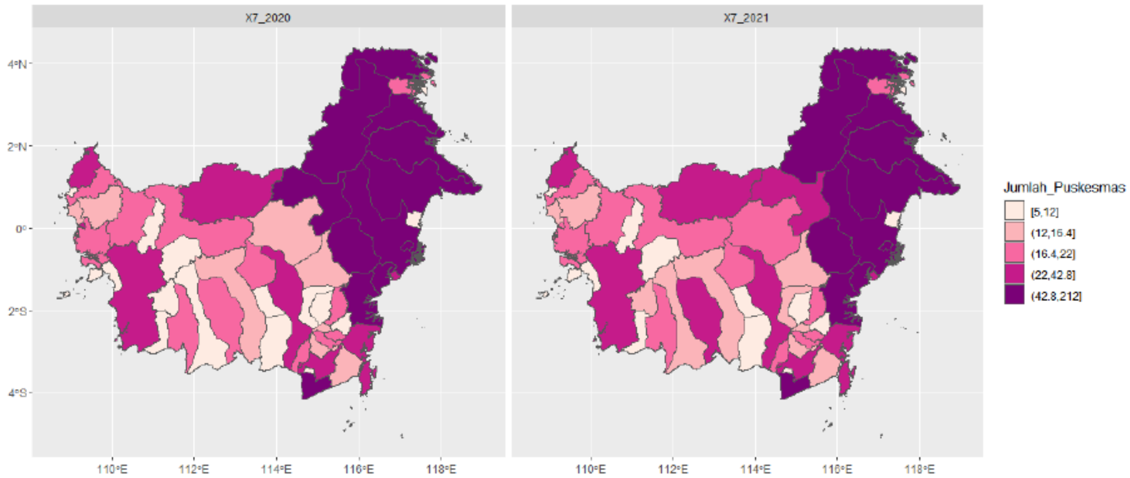


Figure 9. Map of the number of public health centers in 2020-2021

Figure 9 shows the distribution of the number of public health centers in 2020-2021. In general, there is no increase in the number of public health centers in each regency/city in Kalimantan. This is shown by the almost similar and evenly distributed color pattern in each area. However, some areas, such as the provinces of North Kalimantan and East Kalimantan, have a large number of public health centers, as indicated by the darker color than that of other provinces.

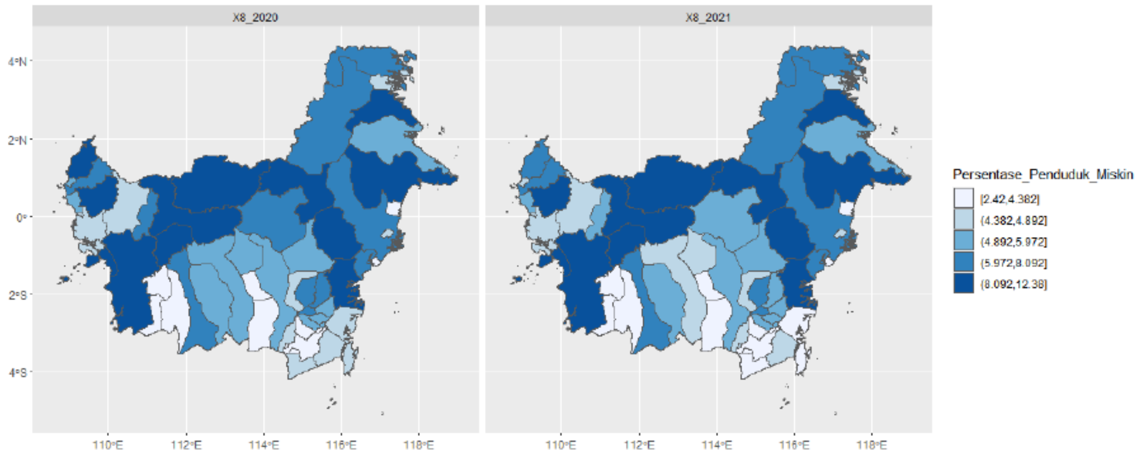


Figure 10. Map of the percentage of poor population in 2020-2021

Figure 10 shows the distribution of the percentage of poor population in 2020-2021. In general, there is no increase in the percentage of poor population in each regency/city in Kalimantan. This is shown by the almost similar and evenly distributed color pattern in each area. However, some areas such as Kutai Kertanegara, East Kutai, West Kutai, Paser, and Berau, have a higher percentage of poor population than that of other regencies/cities, as indicated by the darker color.

3.2 Description of COVID-19 Cumulative Data and Predictor Variables

The description of the COVID-19 cumulative data and predictor variables for the observation data in Table 1 is shown in Table 2.

Table 2 Summary of Variable Statistics

Descriptive statistics	**Confirmed positive cases of COVID-19**	**Number of doctors (x_1)**	**Number of TB cases (x_2)**	**Percentage of elderly population (x_3)**	**Population density (x_4)**	**GRDP (x_5)**	**Number of hospitals (x_6)**	**Number of public health centers (x_7)**	**Percentage of poor population (x_8)**
Minimum	15	5	11	2	1	1,763	1	5	2
Maximum	33,996	821	2,708	59	8,792	126,160	20	212	12
Range	33,981	816	2,697	58	8,791	124,397	19	207	10
Sum	332,489	12,084	49,283	1,149	45,054	1,818,623	382	3,861	675
Median	1,183	73	355	6	31	8,512	2	20	5
Mean	2,969	108	440	10	402	16,238	3	34	6
SE.Mean	464	12	39	1	133	2,122	0	4	0
Variance	24,119,005	16,810	166,245	118	1,969,598	504,525,521	11	1,609	5
Std.dev	4,911	130	408	11	1,403	22,462	3	40	2

Correlation between the variable y and each variable $x_1, x_2, x_3, x_4, x_5, x_6, x_7,$ and x_8 is given in Table 3.

Table 3. Correlation of independent variable to the number of positive COVID-19 cases.

Variable	**Correlation**	***p-value***
x_1	0.684	0.000*
x_2	0.255	0.006*
x_3	0.048	0.612
x_4	0.232	0.013*

x_5	0.628	0.000*
x_6	0.501	0.000*
x_7	0.353	0.000*
x_8	-0.144	0.129

Note: (*) = significant at 5% significance level.

The value of the correlation of the explanatory variable to the response variable shows that the variable x_1 has a high and positive correlation to the variable y. In addition, the variable x_1 has a significant correlation to the variable y. It can be concluded that the higher the number of positive COVID-19 cases, the higher the number of doctors will be.

Table 4. Multicollinearity Test

Predictor Variable	**VIF**
x_1	2.521
x_2	2.481
x_3	1.294
x_5	1.455
x_6	3.557

The results of the multicollinearity test in Table 4 show that all variables have a VIF value of < 5; thus, all independent variables have no multicollinearity.

Breusch-Pagan	**p-Value**
0.90079	4.642e-07

Table 5. Spatial variability test value

The results of the spatial variability test using the Breusch-Pagan test are shown in Table 5. It shows a p-value of 4.642e-07 < 0.05; thus, there is spatial variability in the multiple linear regression model.

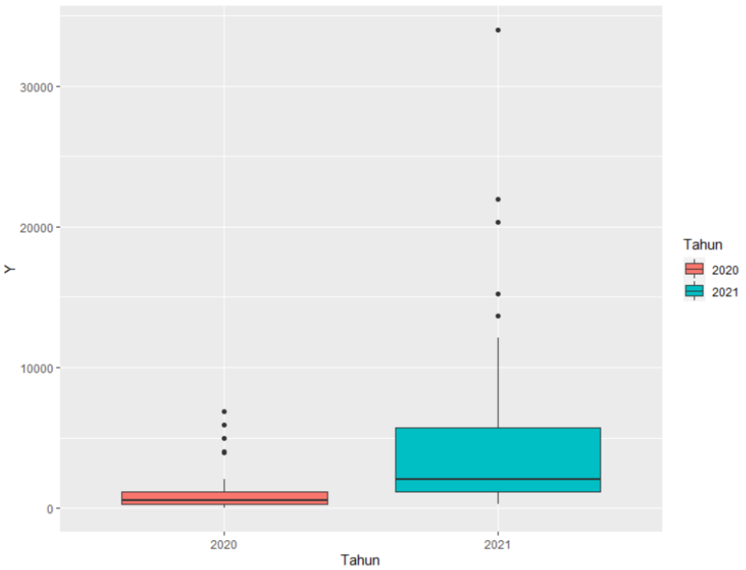


Figure 11. Boxplot of temporal variability for each year

Figure 11 shows the visualization results of the number of positive COVID-19 cases in 2020 to 2021 by using Boxplot. Figure 11 shows that in 2021 the variability in the number of positive COVID-19 cases is larger than that of 2020. This difference in variability indicates a variability between years or so-called temporal variability. The results of the analysis of the

Breusch-Pagan and Boxplot tests lead to a conclusion that the GTWR modeling can be performed in the study of 56 regencies/cities in Kalimantan.

3.3 Geographically Temporally Weighted Regression (GTWR) Modeling for 56 regencies/cities in Kalimantan

3.3.1 Estimation of GTWR Model

The analysis of the GTWR model estimation uses Equation (2) at the *i*-th location where the location $i = 1,2,\cdots,56$ is the initials for 56 regencies/cities in Kalimantan and the the-*t* time is 1 for 2020 and 2 for 2021. The estimation results of the GTWR model are given in Equation (13).

$$\hat{y}_{it} = \hat{\beta}_0(u_i,v_i,t_i) + \hat{\beta}_1(u_i,v_i,t_i)x_{it1} + \hat{\beta}_2(u_i,v_i,t_i)x_{it2} + \hat{\beta}_3(u_i,v_i,t_i)x_{it3} + \hat{\beta}_5(u_i,v_i,t_i)x_{it5} + \hat{\beta}_6(u_i,v_i,t_i)x_{it6} \quad ,i = 1,2,\dots,56\,;\ t = 1,2 \qquad (13)$$

3.3.2 Distance Function Values

The first step to estimate the parameters of the GTWR model is to determine the spatiotemporal distance matrix with the Euclidean function among points of observation based on geographic position (longitude and latitude) at the *t*-th time. The Euclidean distance matrix for the 1st, 2nd, and 56th observation locations with other observation locations at each observation time is shown in Table 6.

Table 6. Spatiotemporal distance matrix with Euclidean function

Location (i,j)		1	2	3	4	5	6	⋯	56
	Waktu (t)	2020	2020	2020	2020	2020	2020	⋯	2021
1	2020	0	3.873	2.229	3.837	4.899	1.135	⋯	17.014
2	2020	3.873	0	1.646	0.081	1.410	2.786	⋯	19.358
3	2020	2.229	1.646	0	1.609	2.793	1.148	⋯	18.337
4	2020	3.837	0.081	1.609	0	1.489	2.743	⋯	19.380
5	2020	4.899	1.410	2.793	1.489	0	3.920	⋯	19.318
6	2020	1.135	2.786	1.148	2.743	3.920	0	⋯	17.820
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋱	⋮
56	2021	17.014	19.358	18.337	19.380	19.318	17.820	⋯	0

3.3.3. Geographical Weighting Function Value

The spatiotemporal weighting function used in the present study is a Gaussian Kernel function with a fixed bandwidth. The Gaussian Kernel function is one that generally gives better GTWR modeling results. This is due to the fact the Gaussian kernel function is a function that resembles the normal distribution function. The spatiotemporal weighting matrix with a Gaussian kernel function at the 1st, 2nd, to 56th observation locations with other observation locations at each observation time is shown in Table 7.

Table 7. Spatiotemporal weighting matrix with Gaussian kernel function

Location (i,j)		1	2	3	4	5	6	⋯	56
	Waktu (t)	2020	2020	2020	2020	2020	2020	⋯	2021
1	2020	1	0.191	0.578	0.197	0.071	0.868	⋯	1.38E-14
2	2020	0.191	1	0.742	0.999	0.803	0.425	⋯	1.14E-18
3	2020	0.578	0.742	1	0.752	0.423	0.865	⋯	7.92E-17
4	2020	0.197	0.999	0.752	1	0.783	0.436	⋯	1.03E-18

5	2020	0.071	0.803	0.423	0.783	1	0.184	···	1.35E-18
6	2020	0.868	0.425	0.865	0.436	0.184	1	···	6.20E-16
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋱	⋮
56	2021	1.38E-14	1.14E-18	7.92E-17	1.03E-18	1.35E-18	6.20E-16	...	1

3.3.4 GTWR Model Parameter Estimation

Table 8. Summary of the estimated values of the GTWR model parameters

Estimasi Parameter	Minimum	Q_1	Median	Q_3	Maximum
$\hat{\beta}_0$	-1612.200	-886.460	-282.050	-64.537	1206.736
$\hat{\beta}_1$	-3.750	-0.609	0.033	5.815	23.556
$\hat{\beta}_2$	-4.870	-0.634	-0.197	0.560	2.702
$\hat{\beta}_3$	-20.633	-4.809	5.696	29.529	110.782
$\hat{\beta}_5$	0.030	0.033	0.086	0.170	0.210
$\hat{\beta}_6$	-308.440	170.570	220.270	849.960	1024.984

Table 8 shows the summary results of GTWR modeling using the Gaussian kernel function with a fixed bandwidth on the spatial and temporal weighting function. The variable number of doctors (x_1) has a coefficient value ranging from –3.750 to 23.5555. The variable number of TB cases (x_2) has a coefficient value ranging from –4,869 to 2,702. The variable percentage of elderly population () has a coefficient value ranging from –20,633 to 110,781. The variable GRDP (x_5) has a coefficient value ranging from 0.0303 to 0.2104. The variable number of hospitals (x_6) has a coefficient value ranging from –308.44 to 1024,983. The coefficient values for each of these variables are spread across all regencies/cities in Kalimantan.

The results of parameter estimation provide GTWR model estimators which states the correlation of the independent variables number of doctors (x_1), number of TB cases (x_2), percentage of elderly population (x_3), GRDP (x_5), and number of hospitals (x_6) to the percentage of positive COVID-19 cases in Kalimantan Provinces. Four GTWR models are given for 4 regency/city locations in Equation (14), equation (15), Equation (16) and Equation (17).

Samarinda City, East Kalimantan Province 2020

$$\hat{y}_{it} = -206.539 - 0.898X_{it1} + 0.248X_{it2} - 2.802X_{it3} + 0.034X_{it5} + 264.725X_{it6} \tag{14}$$

Samarinda City, East Kalimantan Province 2021

$$\hat{y}_{it} = -515.123 + 12.700X_{it1} + 1.194X_{it2} + 4.734X_{it3} + 0.149X_{it5} + 432.961X_{it6} \tag{15}$$

Kapuas Hulu Regency, West Kalimantan Province 2020

$$\hat{y}_{it} = -405.751 - 0.096X_{it1} - 0.149X_{it2} + 19.508X_{it3} + 0.04X_{it5} + 188.398X_{it6} \tag{16}$$

Kapuas Hulu Regency, West Kalimantan Province 2021

$$\hat{y}_{it} = -1382.853 + 5.676X_{it1} - 4.156X_{it2} + 74.872X_{it3} + 0.185X_{it5} + 903.160X_{it6} \tag{17}$$

3.3.5 Measure of Model's Goodness of Fit

The measure of the goodness used to compare the OLS model and GTWR model is the coefficient of determination (R^2), adjusted R^2, Akaike information criterion (AIC) and the root mean square error (RMSE). The results of the comparison of the value of the goodness-of-fit measure are shown in Table 9.

Table 9. Comparison of models in terms of the number of positive COVID-19 cases

Kriteria	OLS	GTWR
R^2	0.6134	0.95713
Adjusted R^2	0.5952	0.92855
AIC	2128.229	1900.76
RMSE	3039.91	1302.99

The above comparison of models shows that the GTWR model is better than the OLS model. This is indicated by the higher values of R^2 and adjusted R^2, as well as the smaller values of AIC and RMSE criteria.

3.3.6 Simultaneous Significance Test of GTWR Model Parameters

The first hypothesis testing conducted is simultaneous tests of the model in order to test the goodness of fit of the GTWR model. The hypothesis testing for the goodness of fit of the GWPR model is as follows:

H_0: $\hat{\beta}_k(u_i, v_i, t_i) = \hat{\beta}_k \quad , k = 1,2, \dots ,5 \,; i = 1,2, \dots ,56 \,; t = 1,2$

(There is no significant difference between multiple linear regression models and GTWR models.)

H_1: There is at least one $\hat{\beta}_k(u_i, v_i, t_i) \neq \hat{\beta}_k \; k = 1,2, \dots ,5 \,; i = 1,2, \dots ,56 \,; t = 1,2$

(There is no significant difference between multiple linear regression models and GTWR models.)

Table 10. Values of simultaneous hypotheses testing of the model's goodness of fit

F-*Statistics*	F tabel	*p-value*	Keputusan Uji
14.440	1.537	0.000	Tolak Ho

Table 10 shows that F- *Statistics* = 14,440 > F-table = 1,537 or *p*-value = 0.000 < α = 0.05. Thus, H_0 is rejected and there is a significant difference between the multiple linear regression model and the GTWR model.

3.3.7 Partial Significance Test of GTWR Model Parameters

Partial parameter tests aim to determine the partial effects of the independent variables on the dependent variable. The hypothesis of partial tests of regression model parameters for the parameter $\hat{\beta}_k(u_i, v_i, t_i)$ is as follows:

H_0: $\hat{\beta}_k(u_i, v_i, t_i) = 0 \quad , k = 1,2, \dots ,5 \,; i = 1,2, \dots ,56 \,; t = 1,2$

(The independent variable X_{kt} has no effect on the number of positive COVID-19 cases in Kalimantan Provinces.)

H_1: $\hat{\beta}_k(u_i, v_i, t_i) \neq 0 \quad , k = 1,2, \dots ,5 \,; i = 1,2, \dots ,56 \,; t = 1,2$

(The independent variable X_{kt} has an effect on the number of positive COVID-19 cases in Kalimantan Provinces.)

The test statistic of the partial parameter testing is the statistic of the *t*-test. The criteria for rejecting H_0 at the significance level of α = 0.05 is to reject H_0 if the *p*-value < 0.05. The results of the partial test of parameters are shown in Table 11.

Table 11. The test statistical value of partial hypothesis testing of the GTWR model parameters

Location	**Year**	**Parameter**	**Estimator Value**	**Standard Error**	**T-Value**	**p-value**
Samarinda	2020	β_0	-206.539	421.063	-0.491	0.625
		β_1	-0.898	4.825	-0.186	0.853
		β_2	0.248	0.859	0.289	0.773
		β_3	-2.802	21.995	-0.127	0.899
		β_5	0.034	0.011	3.016	0.003*
		β_6	264.725	129.780	2.040	0.044*
	2021	β_0	-515.123	422.710	-1.219	0.226
		β_1	12.700	2.200	5.773	0.000*
		β_2	1.194	1.109	1.076	0.284
		β_3	4.734	24.235	0.195	0.845
		β_5	0.149	0.011	13.838	0.000*
		β_6	432.961	142.856	3.031	0.003*
Kapuas Hulu	2020	β_0	-405.751	388.938	-1.043	0.299
		β_1	-0.096	3.698	-0.026	0.979
		β_2	-0.149	0.770	-0.194	0.846
		β_3	19.508	20.966	0.930	0.354
		β_5	0.040	0.010	4.088	0.000*
		β_6	188.398	111.859	1.684	0.095
	2021	β_0	-1382.853	400.964	-3.449	0.001*
		β_1	5.676	2.290	2.478	0.015*
		β_2	-4.156	0.738	-5.631	0.000*
		β_3	74.872	22.760	3.290	0.001*
		β_5	0.185	0.010	17.661	0.000*
		β_6	903.160	117.218	7.705	0.000*

Note: (*) Significant at the 5% significance level

The table above shows that the factors affecting the number of positive COVID-19 cases in Berau Regency are the human development index, life expectancy, gross regional domestic income, and population growth rate, and so on for all observation locations in Kalimantan Provinces. This is shown by the *p*-value of those variables that is lower than 0.05.

3.3.8 Mapping based on the significance of GTWR model parameters

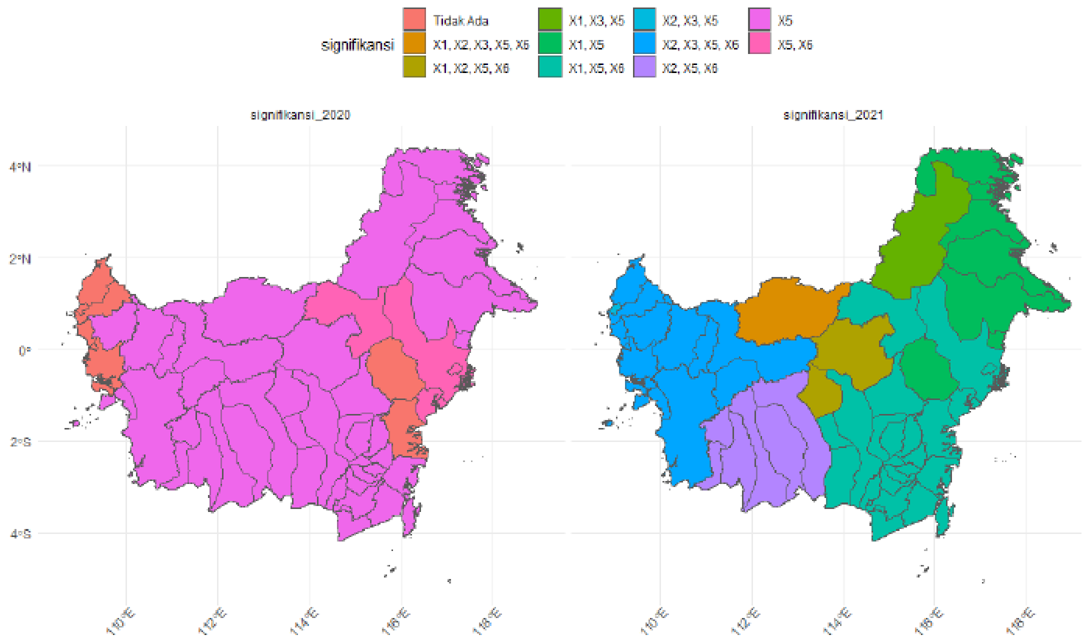


Figure 12. Significance of variables at 5% significance level

Figure 12 shows the result of a GTWR model analysis, indicating variables significantly affecting the number of positive COVID-19 cases in Kalimantan in 2020-2021. In 2020, GRDP (x_5) and number of hospitals (x_6) had a significant effect on the number of positive COVID-19 cases in the majority of regency/City. Meanwhile, in 2021, x_2, x_3, x_5 and x_6 were those variables with a significant effect on the number of positive COVID-19 cases in West Kalimantan Province. In North Kalimantan Province, the variables x_1, x_3 and x_5 had a significant effect on the number of positive COVID-19 cases. Furthermore, in South Kalimantan Province, the variables with a significant effect on the number of positive COVID-19 cases were x_1, x_5 and x_6.

Conclusions

The present study developed a geographically temporally weighted regression (GTWR) model by constructing a distance function with spatial and temporal interactions. The GTWR model uses a Gaussian kernel function with a fixed bandwidth on its spatial and temporal weighting functions. The GTWR model has the greatest goodness of fit as shown by the coefficient of determination $R^2 = 0.957$, adjusted $R^2 = 0.928$, Akaike information criterion (AIC) = 1900.76 and root mean square error (RMSE) = 1302.99. Based on the Spatio-temporal analysis using the GTWR model, the factors that influence the increase in positive cases of COVID-19 are different for each district/city in Kalimantan. Overall, the factors that affect COVID-19 are the number of doctors, the number of hospitals, the number of villages that have puskesmas, and the number of tuberculosis cases. The population sector includes the percentage of the elderly population, population density, and the percentage of the poor. The highest effect based on the GTWR model is tuberculosis cases, health services, and elderly population percentage. So that local governments need to pay attention to patients with tuberculosis, health services, and population density who are the easiest to contract the COVID-19 virus. The mapping of the spread of COVID-19 based on the model's significant variables is grouped into 11 groups so that each region can find out the factors that can be considered to prevent an increase in positive cases of COVID-19.

Conflict of Interest

The authors have no conflict of interest related to this research.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors gratefully acknowledge the funding of KEMENDIKBUD RISTEK Indonesia in 2021 [597/UN17.L1/PG/2021]

References

Ahmadi, A., Fadaei, Y., Shirani, M., & Rahmani, F. (2020). Modeling and forecasting trend of COVID-19 epidemic in Iran until may 13, 2020. *Medical Journal of the Islamic Republic of Iran*, *34*(1). https://doi.org/10.34171/mjiri.34.27

Ahmadi, M., Sharifi, A., Dorosti, S., Jafarzadeh Ghoushchi, S., & Ghanbari, N. (2020). Investigation of effective climatology parameters on COVID-19 outbreak in Iran. *Science of the Total Environment*, *729*(April), 138705. https://doi.org/10.1016/j.scitotenv.2020.138705

Arab-Mazar, Z., Sah, R., Rabaan, A. A., Dhama, K., & Rodriguez-Morales, A. J. (2020). Mapping the incidence of the COVID-19 hotspot in Iran – Implications for Travellers. *Travel Medicine and Infectious Disease*, *34*(March). https://doi.org/10.1016/j.tmaid.2020.101630

BPS Provinsi Kalimantan Barat. 2020. Kalimantan Barat Dalam Angka 2020. Badan Pusat Statistik

BPS Provinsi Kalimantan Tengah. 2020. Kalimantan Tengah Dalam Angka 2020. Badan Pusat Statistik

BPS Provinsi Kalimantan Timur. 2020. Kalimantan Timur Dalam Angka 2020. Badan Pusat Statistik

BPS Provinsi Kalimantan Selatan. 2020. Kalimantan Selatan Dalam Angka 2020. Badan Pusat Statistik

BPS Provinsi Kalimantan Utara. 2020. Kalimantan Utara Dalam Angka 2020. Badan Pusat Statistik

BPS Provinsi Kalimantan Barat. 2021. Kalimantan Barat Dalam Angka 2021. Badan Pusat Statistik

BPS Provinsi Kalimantan Tengah. 2021. Kalimantan Tengah Dalam Angka 2021. Badan Pusat Statistik

BPS Provinsi Kalimantan Timur. 2021. Kalimantan Timur Dalam Angka 2021. Badan Pusat Statistik

BPS Provinsi Kalimantan Selatan. 2021. Kalimantan Selatan Dalam Angka 2021. Badan Pusat Statistik

BPS Provinsi Kalimantan Utara. 2021. Kalimantan Utara Dalam Angka 2021. Badan Pusat Statistik

Brunsdon, C., Fotheringham, A. S., & Charlton, M. (1999). Some notes on parametric significance tests for geographically weighted regression. *Journal of Regional Science*, *39*(3), 497–524. https://doi.org/10.1111/0022-4146.00146

Buzai, G. D. (2020). De Wuhan a Luján. Evolución espacial del COVID-19. *Posición*, *3*(2683–8915), 1–21. http://ri.unlu.edu.ar/xmlui/handle/rediunlu/683

Chen, S., Li, Q., Gao, S., Kang, Y., & Shi, X. (2020). Mitigating COVID-19 outbreak via high testing capacity and strong transmission-intervention in the United States. *MedRxiv*, 2020.04.03.20052720. https://doi.org/10.1101/2020.04.03.20052720

Chen, Z. L., Zhang, Q., Lu, Y., Guo, Z. M., Zhang, X., Zhang, W. J., Guo, C., Liao, C. H., Li, Q. L., Han, X. H., & Lu, J. H. (2020). Distribution of the COVID-19 epidemic and correlation with population emigration from Wuhan, China. *Chinese Medical Journal*, *133*(9), 1044–1050. https://doi.org/10.1097/CM9.0000000000000782

Crespo, R., Fotheringham, S., & Charlton, M. (2007). Application of geographically weighted regression to a 19-year set of house price data in London to calibrate local hedonic price models. *Proceedings of the 9th …*. https://mural.maynoothuniversity.ie/5816/1/MC_application.pdf

Dagnino, R., Weber, E. J., & Panitz, L. M. (2020). *Monitoramento do Coronavírus (Covid-19) nos municípios do Rio Grande do Sul, Brasil*. 1–8. https://osf.io/download/5e7ed0488140b4001c426fa8/

de Ángel Solá, D. E., Wang, L., Vázquez, M., & Méndez-Lázaro, P. A. (2020). Weathering the pandemic: How the Caribbean Basin can use viral and environmental patterns to predict, prepare, and respond to COVID-19. *Journal of Medical Virology*, *92*(9), 1460–1468. https://doi.org/10.1002/jmv.25864

Desjardins, M. R., Hohl, A., & Delmelle, E. M. (2020). Rapid surveillance of COVID-19 in the United States using a prospective space-time scan statistic: Detecting and evaluating emerging clusters. *Applied Geography*, *January*. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7139246/pdf/main.pdf

Dinas Kesehatan Provinsi Kalimantan Barat. 2021. Profil Kesehatan Provinsi Kalimantan Barat Tahun 2021

Dinas Kesehatan Provinsi Kalimantan Barat. 2020. Profil Kesehatan Provinsi Kalimantan Barat Tahun 2020

Dinas Kesehatan Provinsi Kalimantan Tengah. 2020. Profil Kesehatan Provinsi Kalimantan Tengah Tahun 2020

Dinas Kesehatan Provinsi Kalimantan Timur. 2020. Profil Kesehatan Provinsi Kalimantan Timur Tahun 2020

Dinas Kesehatan Provinsi Kalimantan Selatan. 2020. Profil Kesehatan Provinsi Kalimantan Selatan Tahun 2020

Dinas Kesehatan Provinsi Kalimantan Utara. 2020. Profil Kesehatan Provinsi Kalimantan Utara Tahun 2020

Dinas Kesehatan Provinsi Kalimantan Tengah. 2021. Profil Kesehatan Provinsi Kalimantan Tengah Tahun 2021

Dinas Kesehatan Provinsi Kalimantan Timur. 2021. Profil Kesehatan Provinsi Kalimantan Timur Tahun 2021

Dinas Kesehatan Provinsi Kalimantan Selatan. 2021. Profil Kesehatan Provinsi Kalimantan Selatan Tahun 2021

Dinas Kesehatan Provinsi Kalimantan Utara. 2021. Profil Kesehatan Provinsi Kalimantan Utara Tahun 2021

Fotheringham, A.S, Brundson, C and Charlton, M., Geographically Weighted Regression : The Analysis of Spatially Varying Relationships, John Wiley & Sons Ltd, England. 2002

Fotheringham, A. S., Crespo, R., & Yao, J. (2015). Geographical and Temporal Weighted Regression (GTWR). *Geographical Analysis*, *47*(4), 431–452. https://doi.org/10.1111/gean.12071

Giuliani, D., Dickson, M. M., Espa, G., & Santi, F. (2020). *Modelling and Predicting the spatio-temporal spread of coronavirus disease 2019 (COVID-19) in Italy*. *March*. https://doi.org/10.2139/ssrn.3559569

Gross, B., Zheng, Z., Liu, S., Chen, X., Sela, A., Li, J., Li, D., & Havlin, S. (2020). Spatio-temporal propagation of COVID-19 pandemics. *Medrxiv.Org*, 1–7. https://www.medrxiv.org/content/10.1101/2020.03.23.20041517v3.full.pdf

Guan, W., Ni, Z., Hu, Y., Liang, W., Ou, C., He, J., Liu, L., Shan, H., Lei, C., Hui, D. S. C., Du, B., Li, L., Zeng, G., Yuen, K.-Y., Chen, R., Tang, C., Wang, T., Chen, P., Xiang, J., … Zhong, N. (2020). Clinical Characteristics of Coronavirus Disease 2019 in China. *New England Journal of Medicine*, *382*(18), 1708–1720. https://doi.org/10.1056/nejmoa2002032

Huang, B., Wu, B., & Barry, M. (2010). Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. *International Journal of Geographical Information Science*, *24*(3), 383–401. https://doi.org/10.1080/13658810802672469

Huang, H., Wang, Y., Wang, Z., Liang, Z., Qu, S., Ma, S., Mao, G., & Liu, X. (2020). *rin ep Pr t n ot er r ed Pr t n ot pe er ed*. http://doi.org/10.2139/ssrn.3550007

Leung, Y., Mei, C. L., & Zhang, W. X. (2000a). Statistical tests for spatial nonstationarity based on the geographically weighted regression model. *Environment and Planning A*, *32*(1), 9–32. https://doi.org/10.1068%2Fa3162

Leung, Y., Mei, C. L., & Zhang, W. X. (2000b). Testing for spatial autocorrelation among the residuals of the geographically weighted regression. *Environment and Planning A*, *32*(5), 871–890. https://doi.org/10.1068%2Fa32117

Liu, J., Zhao, Y., Yang, Y., Xu, S., Zhang, F., Zhang, X., Shi, L., & Qiu, A. (2017). A mixed geographically and temporallyweighted regression: Exploring spatial-temporal variations from global and local perspectives. *Entropy*, *19*(2), 1–20. https://doi.org/10.3390/e19020053

Murugesan, B., Karuppannan, S., Mengistie, A. T., Ranganathan, M., & Gopalakrishnan, G. (2020). Distribution and Trend Analysis of COVID-19 in India: Geospatial Approach. *Journal of Geographical Studies*, *4*(1), 1–9. https://doi.org/10.21523/gcj5.20040101

Orea, L., & Alvarez, I. C. (2020). How effective has the Spanish lockdown been to battle COVID-19? A spatial analysis of the coronavirus propagation across provinces. *Documento de Trabajo - 2020/03*, 1–27. https://documentos.fedea.net/pubs/dt/2020/dt2020-03.pdf

Rezaei, M., Nouri, A. A., Park, G. S., & Kim, D. H. (2020). Application of geographic information system in monitoring and detecting the COVID-19 outbreak. *Iranian Journal of Public Health*, *49*, 114–116. https://doi.org/10.18502/ijph.v49is1.3679

Rossman, H., Keshet, A., Shilo, S., Gavrieli, A., Bauman, T., Cohen, O., Belicer, R., Geiger, B., Dor, Y., & Segal, E. (2020). A framework for identifying regional outbreak and spread of COVID-19 from one-minute population-wide surveys. *Nature Medicine*, *26*(5), 632. https://doi.org/10.1038/s41591-020-0853-0

Saha, A., Gupta, K., Patil, M., & Urvashi. (2020). Monitoring and epidemiological trends of coronavirus disease (COVID-19) around the world. *Matrix Science Medica*, *4*(4), 121. https://www.researchgate.net/profile/Arnab-Saha-14/publication/346151257_Monitoring_and_epidemiological_trends_of_coronavirus_disease_COVID-19_around_the_world/links/5ffd380292851c13fe06b512/Monitoring-and-epidemiological-trends-of-coronavirus-disease-COV

Santana Juárez, M. V. (2020). COVID-19 en México: Comportamiento espacio temporal y condicionantes socioespaciales, febrero y marzo de 2020. *Posicion*, *3*, 1–27. https://716132a6-9cf5-45de-baee-6a15e46210f7.filesusr.com/ugd/df634b_96bb0dd9fa6b4621b96a2d722105f2bd.pdf

Sifriyani, S., & Rosadi, D. (2020). Susceptible Infected Recovered (SIR) Model for Estimating Covid-19 Reproduction Number in East Kalimantan and Samarinda. *Media Statistika*, *13*(2), 170–181. https://doi.org/10.14710/medstat.13.2.170-181

Tang, W., Liao, H., Marley, G., Wang, Z., Cheng, W., Wu, D., & Yu, R. (2020). The changing patterns of coronavirus disease 2019 (COVID-19) in China: A tempogeographic analysis of the severe acute respiratory syndrome coronavirus 2 epidemic. *Clinical Infectious Diseases*, *71*(15), 818–824. https://doi.org/10.1093/cid/ciaa423

Team, C. C.-19 R. (2020). Geographic Differences in COVID-19 Cases, Deaths, and Incidence — United States, February 12–April 7, 2020. *US Department of Health and Human Services/Centers for Disease Control and Prevention*, *69*(April 17, 2020), 15. https://covid-19.conacyt.mx/jspui/handle/1000/2490

Wang, P. 2006. *Exploring spatial effects on housing price: the case study of the city of Calgary*. Master dissertation. University of Calgary, Canada.

Wu, B., Li, R., & Huang, B. (2014). A geographically and temporally weighted autoregressive model with application to housing prices. *International Journal of Geographical Information Science*, *28*, 1186–1204. https://doi.org/10.1080/13658816.2013.878463

Xiong, Y., Wang, Y., Chen, F., & Zhu, M. (2020). Spatial Statistics and Influencing Factors of the Epidemic of Novel Coronavirus Pneumonia 2019 in Hubei Province, China. *Research Square*, 1–25. https://doi.org/10.21203/rs.3.rs-16858/v1

Zhang, X., Rao, H., Wu, Y., Huang, Y., & Dai, H. (2020). Comparison of spatiotemporal characteristics of the COVID-19 and SARS outbreaks in mainland China. *BMC Infectious Diseases*, *20*(1), 1–7. https://doi.org/10.1186/s12879-020-05537-y

Zhou, C., Su, F., Pei, T., Zhang, A., Du, Y., Luo, B., Cao, Z., Wang, J., Yuan, W., Zhu, Y., Song, C., Chen, J., Xu, J., Li, F., Ma, T., Jiang, L., Yan, F., Yi, J., Hu, Y., … Xiao, H. (2020). COVID-19: Challenges to GIS with Big Data. *Geography and Sustainability*, *1*(1), 77–87. https://doi.org/10.1016/j.geosus.2020.03.005

www.corona.kalbarprov.go.id

www.corona.kalteng.go.id

www.covid19.kaltimprov.go.id

www.corona.kalselprov.go.id

www.coronainfo.kaltaraprov.go.id

Tuberculosis Cases and Population Density Affect the Increase in COVID-19 Outbreak: A Spatiotemporal Analysis of Geographically and Temporally Weighted Regression

Sifriyani[1], Dedi Rosadi[2], Mariani Rasjid[3]

[1]*Statistics Study Program, Department of Mathematics, Faculty of Mathematics and Natural Sciences, Mulawarman University, Samarinda, Indonesia*
[2]*Department of Mathematics, Faculty of Mathematics and Natural Sciences, Gadjah Mada University, Yogyakarta, Indonesia*
[3]*Pulmonologist, Tadulako General Hospital, Indonesia*
[3]*Faculty of Medicine, Tadulako University, Kendari, Indonesia*

Abstract

COVID-19 cases in Indonesia are increasing and spreading in all provinces, the severity is also occurring in Kalimantan. Application of geographic theory and methodology in resolution of the increase in positive cases of COVID-19 from the perspective of Spatial and Time in Kalimantan Province. The present study uses a spatiotemporal analysis to determine the factors affecting the constantly increase in COVID-19 cases in Kalimantan. The spatiotemporal analysis uses the geographically temporally weighted regression (GTWR) model by developing a spatial and temporal interaction distance function. The GTWR model was applied to data on the number of positive COVID-19 cases at a scale of 56 districts/cities on the Indonesian island of Kalimantan in the period of January 2020 to August 2021. The purposes of the present study was to determine the factors affecting the cumulative increase in COVID-19 cases in Kalimantan and to map the spatial distribution for 56 districts/cities based on the significant predictor variables. Results of the study showed that the GTWR model with the development of a spatial and temporal interaction distance function using the kernel Gaussian fixed bandwidth function is the best model. Factors affecting the positive cases of COVID-19 in Kalimantan for each district/city were varying, based on the significant variables, including the number of doctors, the number of TB cases, the percentage of the elderly population, GRDP, and the number of hospitals. The highest influence that affects COVID-19 cases is the high number of TB cases, population density, and the lack of health services. Furthermore, an area map was produced on the basis of the significant variables affecting the increase in COVID-19 cases. The results of the study provide the local governments with decision-making recommendations for overcoming the COVID-19 problems in their respective regions.

Keywords: Pandemic. Epidemiology COVID-19, Spatio Temporal, Spatial Analysis, Mapping GIS, Applied Geography

Biografi Authors

Sifriyani https://scholar.google.co.id/citations?user=79VWksEAAAAJ&hl=en

Lecturer at Statistics Study Program, Mulawarman University, Samarinda, Indonesia. Now she is an associate professor.
Graduated from Master's Education in the Master's Degree in Statistics at the Sepuluh November Institute of Technology, Surabaya, Indonesia and Doctoral Education in the Doctoral Program in Statistics at Gadjah Mada University. Short-study as visitor for the Spatial Statistics Research at the Faculty of Science and Technology at Lancaster University, United Kingdom, studying with Professor Peter M. Atkinson. Doing some researches since 2017 until now particularly focused on developing researches in the field of Spatial Statistics and has produced several reputable articles about spatial statistics.
She earned a Research Grant for the Topic of Spatial Statistics Development from the Ministry of Education and Culture in 2019-2020 and 2021-2022. In 2020 as a member of the Task Force Team to accelerate the handling of COVID-19 at Mulawarman University, Samarinda, Indonesia.
ID Orchid 0000-0002-4616-775X

Dedi Rosadi. https://acadstaff.ugm.ac.id/dedirosadi
https://scholar.google.co.id/citations?user=3WlY-YUAAAAJ&hl=en

I am a full time professor in the Department of Mathematics, Faculty of Mathematics and Sciences, Universitas Gadjah Mada (UGM). I graduated from Statistics Study Program in Gadjah Mada University in February 1996 and started to work as the lecturer in UGM. In August 1997, I continue my study to obtain the master of science degree in Stochastic Modeling (Applied Statistics) in the University of Twente, The Netherland, graduated in June 1999. From September2001- September 2004 I took my doctoral study (in Econometrics) in the Institute of Econometrics and Operation Research (EOS 119), the Vienna University of Technology (TU Wien), Austria. After finished my doctoral study, i come back to Yogyakarta and continue my work in the Universitas Gadjah Mada until now. In September 2013, I receive the full professor title.
https://orchid.org/0000-0003-2689-253X

Mariani Rasjid.

https://scholar.google.co.id/citations?user=6-mDEVEAAAAJ&hl=id

Lecturer at the Faculty of Medicine, Tadulako University, Palu, Central Sulawesi, Indonesia. Now she is also a doctor at Tadulako General Hospital, Palu, Central Sulawesi, Indonesia. During the COVID-19 pandemic, the author served as the Head of the Task Force Team to accelerate the handling of COVID-19 at the Tadulako General Hospital, Indonesia, the head of the COVID-19 Task Force Team at Muhammadiyah, Central Sulawesi, Indonesia. The author has created a Pocket Book of Getting to Know COVID-19.

Graduated from the Medical Faculty of Hasanuddin University in 2009 and successfully completed the Pulmonology Specialist Education and Respiratory Medicine at Medical Faculty of Airlangga University in 2017. Joined the Professional Organizations called Indonesian Medical Association (IDI) for Palu City Branch and the Indonesian Society of Respiralogy for Sulawesi Branch.
https://orcid.org/0000-0002-1724-3565