15TH NATIONAL CONVENTION ON STATISTICS
03-05 OCTOBER 2022

Organized by the Philippine Statistical System
Spearheaded by the Philippine Statistics Authority

Spatial Autoregressive Epidemic Models for Sparse COVID-19 Deaths in the Philippines

Shirlee R. Ocampo
Department of Mathematics and Statistics
De La Salle University

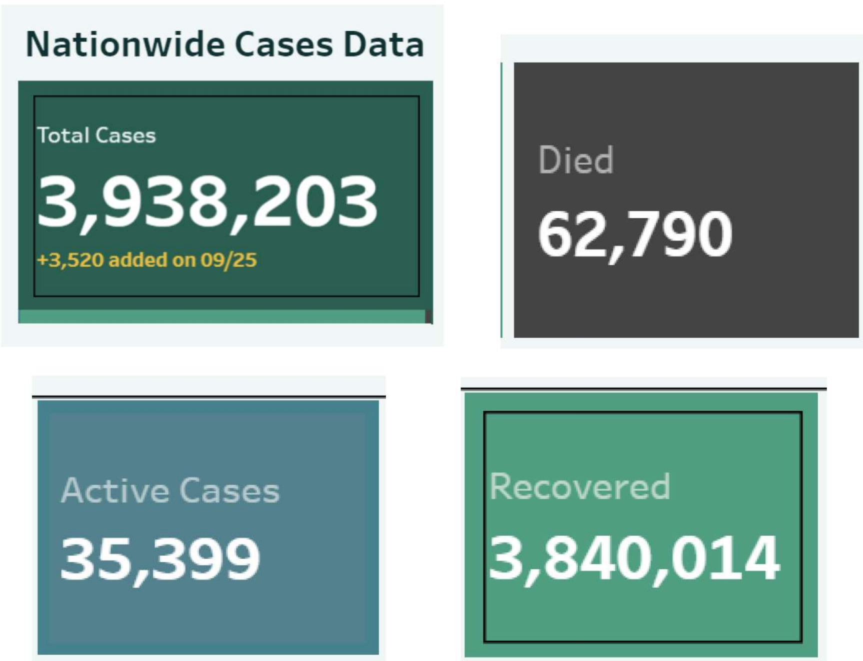
Computational Statistics
Crowne Plaza Manila Galleria
4 October 2022; 1:30 PM

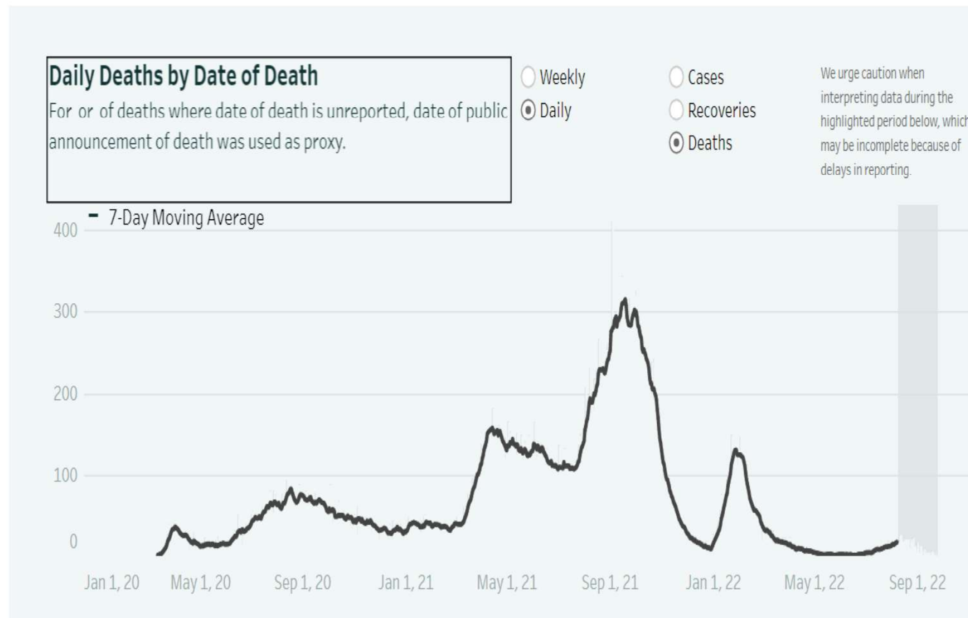
Outline

- Motivation and Background
- Problem and its Significance
- Data and Methodology
  - ➢ Sparsity
  - ➢ Sparse Spatial Autoregressive Model
  - ➢ Spatiotemporal Epidemic Model
  - ➢ Moran's Index
  - ➢ Measures of Predictive Accuracy
- Results and Discussions
- Conclusions and Recommendations

Motivation and Background

- Daily COVID-19 data in the Philippine provinces ➡ MESSY and SPARSE!





https://doh.gov.ph/covid19tracker

Problem and Significance

✓ postulates sparse spatial autoregressive models to link COVID-19 mortality rates with healthcare system, demographic, economic, disease prevalence, vaccination, urbanity, and environmental factors

3 GOOD HEALTH AND WELL-BEING

Healthcare-Related Predictors

- Hospital Beds
- Healthcare Workers
- Licensed COVID-19 Testing Laboratories

Demographic Correlates

- Seniors
- Males

Economic Correlate

Revenue Collection per capita in LGU

Data

daily COVID-19 deaths across provinces and NCR cities (April 1, 2020, to September 15, 2021)

Disease (Comorbidities)
- Tuberculosis
- Cancer
- CVD and Diabetes
- Influenza
- Pneumonia

Urbanity Correlate
Number of Cities

Vaccination Correlate
PPV and Flu Vaccine

Environment Correlates (after PCA)
- Air Quality Index
- Wind Direction
- Relative Humidity

DEPARTMENT OF HEALTH (DOH)

DEPARTMENT OF ENVIRONMENT AND NATURAL RESOURCES (DENR)

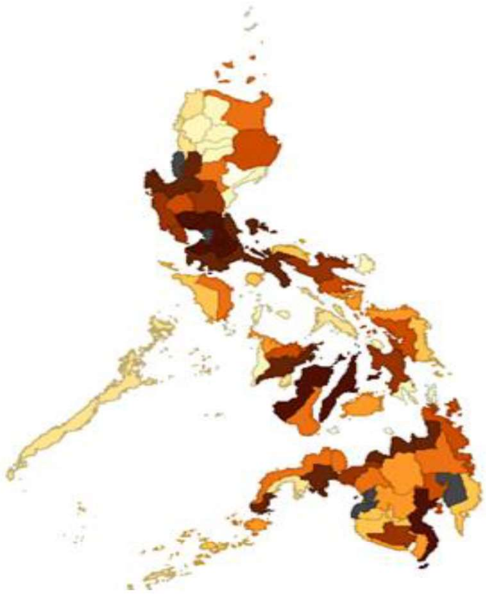
PHILIPPINE STATISTICS AUTHORITY (PSA)

PHILIPPINE ATMOSPHERIC, GEOPHYSICAL AND ASTRONOMICAL SERVICES ADMINISTRATION (PAGASA)

FIELD HEALTH SERVICE INFORMATION SYSTEM (FHSIS)

Sparsity and Moran's I

- presence of too many zero counts spanning over different areas and over time
- sparse data confronted with convergence issues (Aldirawi and Yang, 2022)
- ℓ^p-norm sparsity measures (Hurley and Rickard, 2009)
- ℓ^0-norm sparsity measure of COVID-19 deaths:
  - 14,493 (37,741 zeroes)
- Moran's I for spatial autocorrelation:
  - positive
  - shows significant spatial dependencies ($p < 0.0001$)



Sparse Spatial Autoregressive Model (SAR)

$$Y = X\beta + \alpha D(Y - X\beta) + \varepsilon$$

α – spatial autoregressive parameter, $0 \leq \alpha < 1$

D – spatial weight matrix

- **Neighborhood spatial matrix D_N** (sparsity of 101,419,773 and 2,626,970,983 zeroes)

$$w_{ij} = \begin{cases} 1 & \text{if spatial units } i \text{ and } j \text{ share the same boundaries} \\ 0 & \text{if spatial units } i \text{ and } j \text{ do not share the same boundaries} \end{cases}$$

- **Regional spatial matrix D_R** (sparsity of 174,998,824 with 2,553,391,932 zeroes)

$$w_{ij} = \begin{cases} 1 & \text{if spatial units i and j belong to the same region} \\ 0 & \text{if spatial units i and j belong to different regions} \end{cases}$$

Spatiotemporal Epidemic Model (STEM)

- Landagan and Barrios (2006)

$$Y_{it} = X_{it}\beta_i + w_{it}\gamma_i + \varepsilon_{it}\ ,\ \ i = 1, 2, \dots N, t = 1, 2, \dots T$$

  - w_{it}- set of variables in the neighborhood system of location i at time t

- Estimation: hybrid of Cochranne-Orcutt and backfitting algorithm

(1) Use Cochranne-Orcutt procedure: $Y_{it} = X_{it}\beta_i + \rho\varepsilon_{i(t-1)} + a_{it}$

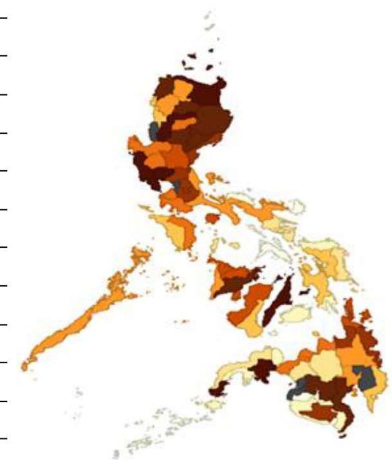
(2) Estimate the spatial component γ_i: $a_{it} = \gamma_i w_{it} + e_{it}$

(3) Compute a new dependent variable: $Y_{it}^1 = Y_{it} - \widehat{\gamma}_i w_{it}$. Iterate.

Results: SAR using DN

Significant Correlates	Parameter Estimate	Standard Eror	t-value	p-value
Intercept	-1.0631	0.0734	14.47	<.0001
Flu	0.0001	0.0013	32.23	<.0001
HW	-0.0000	0.0000	-13.09	<.0001
ImmuSen	-0.0000	0.0000	-3.40	0.0007
LCTL	0.0800	0.0000	84.59	<.0001
Males	-0.0000	0.0009	-4.58	<.0001
NC	0.0124	0.0000	9.39	<.0001
Pneumonia	-0.0000	0.0000	-35.60	<.0001
RCP	-0.0000	0.0000	-40.24	<.0001
CancerBC	0.0000	0.0000	38.08	<.0001
CardioDiab	0.0000	0.0000	30.56	<.0001
Seniors	-0.0000	0.0000	-22.76	<.0001
TB	0.0001	0.0000	39.73	<.0001
WindSpeed	-0.0725	0.0007	-27.25	<.0001
AQI	0.0741	0.0008	113.95	<.0001
RH	0.0048	0.0027	5.78	<.0001
Hosbeds	0.0007	0.0000	86.72	<.0001

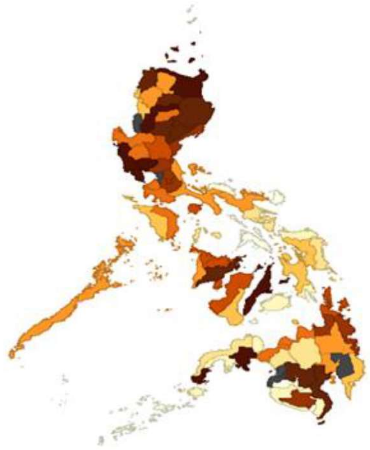
Increased adjusted R square: 60.8%



Results: SAR using Dr

Significant Correlate	Parameter Estimate	Standard Error	t-value	p-value
Intercept	0.5253	0.0486	-10.80	<.0001
Flu	-0.0000	0.0000	-22.41	<.0001
HW	0.0000	0.0000	50.14	<.0001
ImmuSen	-0.0000	0.0000	-4.48	<.0001
LCTL	0.0392	0.0007	55.10	<.0001
Males	0.0000	0.0000	22.51	<.0001
NC	0.0155	0.0009	18.15	<.0001
Pneumonia	-0.0000	0.0000	-38.10	<.0001
RCP	0.0000	0.0000	69.04	<.0001
CancerBC	0.0000	0.0000	35.96	<.0001
CardioDiab	-0.0000	0.0000	-40.48	<.0001
Seniors	-0.0000	0.0000	-149.66	<.0001
TB	0.0001	0.0000	43.37	<.0001
WindSpeed	0.0060	0.0026	2.30	0.0216
AQI	0.0290	0.0004	73.59	<.0001
RH	-0.0041	0.0005	-8.11	<.0001
Hosbeds	0.0006	0.0000	99.36	<.0001

Increased adjusted R square: 87.84%



Results: STEM Spatial Parameter Estimates

Neighborhood Variable	Parameter Estimate after 15 iterations	Parameter Estimate after 50 iterations
Intercept	$3.84(10^{-5})$	$-2.13(10^{-6})$
AQI	$1.22(10^{-7})$	$-8.05(10^{-16})$
Wind speed	$2.72(10^{-9})$	$-4.61(10^{-16})$
Relative Humidity	$-5.49(10^{-7})$	$2.81(10^{-16})$
Hospital beds	$1.98(10^{-9})$	$-3.97(10^{-18})$

- time dimension estimate $\widehat{\rho}$ converged to 0.3535
- parameter estimates converged

Results: STEM Parameter Estimates

Correlates	Parameter Estimate	Standard Error	t-value	p-value
Intercept	0.7627	0.0419	17.40	<0.0001 **
Seniors*	**-0.0000**	**0.0000**	**-1.68**	**0.0926***
Males*	**-0.0000**	**0.0000**	**-1.87**	**0.0611***
HW**	**-0.0001**	**0.0000**	**-10.60**	**<0.0001****
LCTL**	**0.0263**	**0.0044**	**6.02**	**<0.0001****
NC**	**0.0175**	**0.0068**	**2.60**	**0.00953****
Pneumonia	-0.0000	0.0000	-1.60	0.1107
RCP**	**0.0000**	**0.0000**	**2.93**	**0.0034****
ImmuSen	0.0000	0.0000	0.00	0.1985
Flu	-0.0000	0.0000	-1.46	0.1445
CardioDiab**	**0.0000**	**0.0000**	**2.03**	**0.0423****
CancerBC**	**0.0000**	**0.0000**	**2.03**	**0.0420 ****
TB**	**0.0001**	**0.0000**	**9.17**	**<0.0001****

Demographic:
Seniors
Males
Healthcare:
Health workers
LCTL
Urbanity:
Number of cities
Economic:
Revenue Collection
Comorbidities:
CVD and Diabetes
Cancer
Tuberculosis

Results: Predictive Abilities of Models

MODEL	MSE	MAD
Sparse Spatial Autoregressive Model with $\mathbf{D_N}$, $\alpha = 0.05$	3.33	0.95
Sparse Spatial Autoregressive Model with $\mathbf{D_R}$, $\alpha = 0.05$	3.39	0.97
Spatio-Temporal Model	0.08	0.24

Conclusions and Recommendations

- Daily COVID-19 mortality rates are associated with:
  - proportion of health workers, number of licensed COVID-19 testing laboratories, number of cities in an LGU, revenue of the LGU, prevalence rates of cancer, tuberculosis, and cardiovascular disease and diabetes at 5% level of significance
  - proportion of seniors and males at 10% level of significance
- The models emphasize the importance of resources available in the local government that can boost the capabilities of the health care system.
- Pre-existing health conditions (co-morbidities) of the communities also determine mortality rates of COVID-19 in the Philippines.
- Potential research topics: inclusion of a sparse temporal weighing matrix in the SAR model and possible sparsity regularization terms in the spatiotemporal model

References

- ALDIRAWI, H. & YANG, J. 2022. Modeling Sparse Data Using MLE with Applications to Microbiome Data. *J Stat Theory Pract* 16 (1), #13. https://doi.org/10.1007/s42519-021-00230-y
- BASTERO, R. AND BARRIOS, E. 2011. Robust Estimation of a Spatiotemporal Model with Structural Change', Communications in Statistics - Simulation and Computation, 40(3): 448 — 468.
- Department of Health-Republic of the Philippines. 2021. Updates On Novel Coronavirus Disease (COVID-19). https://doh.gov.ph/covid19tracker..
- Hurley, N., & Rickard, S. (2009). Comparing measures of sparsity. *IEEE Transactions on Information Theory*, *55*(10), 4723–4741.
- KAPITSINIS, N. 2020. The underlying factors of the COVID-19 spatially uneven spread. Initial evidence from regions in nine EU countries. Regional Science Policy & Practice, 12(6):1027–1045. https://doi.org/10.1111/rsp3.12340.
- LANDAGAN, O. & BARRIOS, E. 2007. An Estimation Procedure for Spatio-Temporal Model. Science Direct Statistics and Probability, 77: 401-406.
- PACE, R. & BARRY, R. 1997. Sparse Spatial Autoregressions. Statistics and Probability Letters, 33: 291- 297.

Thank you!

Acknowledgment to Dr. Erniel Barrios, DLSU, UPSS, DOH, PSA, PAG-ASA, DENR, and FHSIS.

http://www.psa.gov.ph/ncs

http://openstat.psa.gov.ph

https://twitter.com/PSAgovph

https://www.facebook.com/PSAgovph