



XLVIII Coloquio Argentino de Estadística

VI JORNADA DE EDUCACIÓN ESTADÍSTICA "MARTHA DE ALIAGA"

27 al 30 oct 2020

Poster:

Propuesta de modelo de predicción econométrico espacial para el consumo de agua potable en Chile

Daniel Calderón, Danilo Leal, Orietta Nicolis



Esta obra está bajo una
Licencia Creative Commons
Atribución-NoComercial 4.0
Internacional



FACULTAD
DE CIENCIAS
ECONÓMICAS



Universidad
Nacional
de Córdoba





Universidad Andrés Bello®

PROPUESTA DE MODELO DE PREDICCIÓN ECONOMETRICO ESPACIAL PARA EL CONSUMO DE AGUA POTABLE EN CHILE

DANIEL CALDERÓN, DANILO LEAL, ORIETTA NICOLIS

Facultad de Ingeniería, Universidad Andrés Bello, Viña del Mar, Chile

Centro de Investigación para la Gestión Integrada del Riesgo de Desastres (CIGIDEN), ANID/FONDAP/15110017, Santiago, Chile

Resumen

El consumo de agua potable es un derecho y servicio básico vital para la humanidad debido a que cada persona en el planeta Tierra requiere un consumo promedio entre 20 a 50 litros de agua limpia al día para beber, cocinar o simplemente mantenerse limpios. Este recurso natural es una preocupación a nivel mundial, debido a que se ha ido agotando paulatinamente producto de una serie de factores que están ocurriendo durante el transcurso de los años.

En el presente documento se desarrolla un modelo estadístico para la predicción del consumo del agua potable en Chile considerando las distintas ubicaciones geográficas de todas las comunas del país desde el 2012 hasta el 2019 y su correlación espacial. La primera parte de este trabajo inicia con un análisis estadístico con el objetivo de seleccionar las covariables relacionadas directamente con la variable a predecir (como, por ejemplo: población, precios tarifarios, nivel socio económico, tasa de pobreza, entre otros.). Sucesivamente se realiza un análisis estadístico descriptivo y espacial de todas las variables comprobando la existencia de una dependencia espacial global a través del índice de Moran. Al final se propone la implementación de modelos espacio temporales bayesianos, aplicados a través de la Aproximación de Laplace Anidada Integrada (INLA). Dichos modelos permiten capturar los efectos aleatorios espaciales estructurados (similitudes entre comunas), no estructurados (residuos aleatorios) y los efectos fijos de cada una de las covariables seleccionadas.

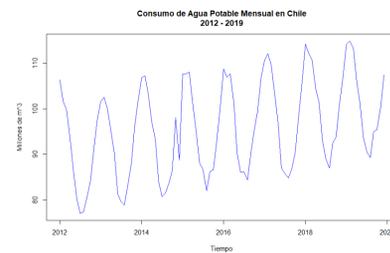
Los resultados obtenidos demuestran que el consumo de agua potable en Chile es más alto en la zona central del país dependiendo esencialmente de la variable índice de nivel socioeconómico, donde el modelo presentó una mayor precisión al momento de considerar la correlación espacial de las variables en comparación a un modelo de regresión multivariado tradicional.

Metodología

- Realizar un estudio bibliográfico para comprender cual son los factores que están relacionas con el consumo de agua potable en Chile y escoger cuales son las variables más importantes a considerar en el planteamiento de los modelos predictivos.
- Desarrollar un análisis estadístico descriptivo detallado de todas las variables que participan en el consumo de agua, con la finalidad de comprender su relación durante el transcurso de los años 2012 a 2019.
- Efectuar un análisis estadístico espacial de información georreferenciada con la finalidad de entender si el consumo de agua potable comunal es dependiente y si este presenta un cierto patrón espacial, utilizando la prueba I de Moran en análisis exploratorios para la autocorrelación espacial.
- Implementar tres modelos de predicción econométrico de la variable consumo de agua, un modelo tradicional y dos modelos, espacio temporales bayesianos para realizar una comparación de los resultados obtenidos basándose en los mejores criterios de evaluación de calidad obtenidos en comparación con los datos reales de la variable consumo de agua potable en Chile.
- Finalmente escoger el mejor modelo predictivo, para utilizarlo en las futuras predicciones del consumo de agua potable en Chile en el Largo plazo.

Análisis Estadístico Descriptivo

Comportamiento del consumo de agua potable mensual de todas las comunas desde el año 2012 al 2019.



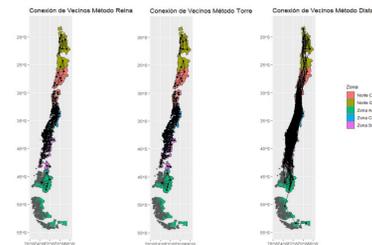
Correlación de las Variables Medias Comunales año 2012 - 2019

PPA	-0.28	-0.26	-0.55	-0.12	-0.56	-0.42	0.41	-0.66	-0.7	1
Tjmed	0.17	0.18	0.55	0.19	0.57	0.17	-0.17	0.74	1	-0.7
Temed	0.25	0.24	0.34	-0.22	0.38	0.14	-0.13	1	0.74	-0.66
PZ	-0.42	-0.39	-0.01	-0.03	-0.03	-0.79	1	-0.13	-0.17	0.41
IDSE	0.63	0.53	0.03	0.04	0.04	1	-0.79	0.14	0.17	-0.42
PSC	0.01	0.02	0.95	0.56	1	0.04	-0.03	0.38	0.57	-0.56
PNP	-0.16	-0.16	0.62	1	0.56	0.04	-0.03	-0.22	-0.19	-0.12
PP	-0.01	-0.02	1	0.62	0.95	0.03	-0.01	0.34	0.55	-0.55
PO	0.93	1	-0.02	-0.16	0.02	0.53	-0.39	0.24	0.18	-0.28
AP	1	0.93	-0.01	-0.16	0.01	0.63	-0.42	0.25	0.17	-0.28

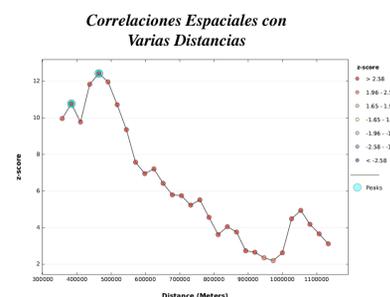
Matriz de correlaciones de todas las variables analizadas en el estudio, se puede observar el grado de dependencia de las variables explicativas con el consumo de agua potable.

Análisis Espacial

Creación de listado de vecindad de las 284 comunas con los métodos Torre, Reina y Distancia para la creación de matriz de pesos espaciales



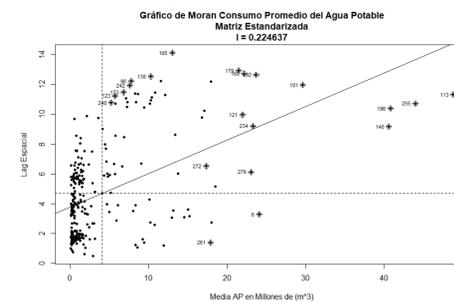
Calculo de distancia máxima a considerar para el calculo de autocorrelación espacial con método de criterio distancia



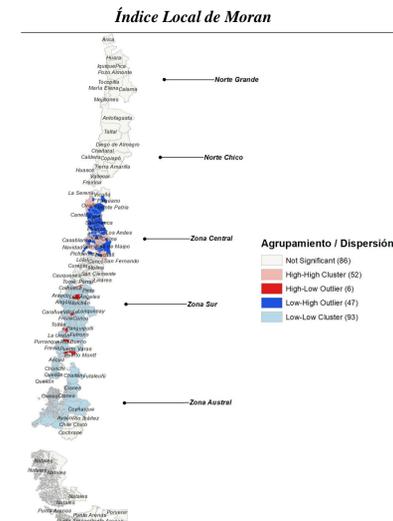
Resultado Índice de Moran

Resultados Obtenidos del Índice de Moran con método de distancia inversa, se puede observar un grado de correlación espacial positivo donde existe la tendencia de agrupamiento de valores en el espacio de todas las variables del estudio.

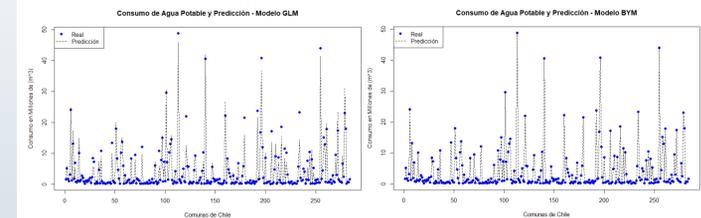
Resultados Índice de Moran			
Variables	Índice	Puntuación Z	Valor P
AP	0.225	16.330	0.000
PO	0.157	11.456	0.000
PP	0.671	47.284	0.000
PNP	0.569	40.356	0.000
PSC	0.688	48.732	0.000
IDSE	0.433	30.534	0.000
PZ	0.427	30.132	0.000
Temed	0.637	44.868	0.000
Tjmed	0.530	37.399	0.000
PPA	0.780	54.743	0.000



Elaboración de Mapa Lisa de Moran, se puede observar el comportamiento de agrupamiento o dispersión de datos de cada una de las comunas consideradas en el estudio, donde la Región Metropolitana presento la mayoría de comunas con agrupamiento de valores en el espacio.

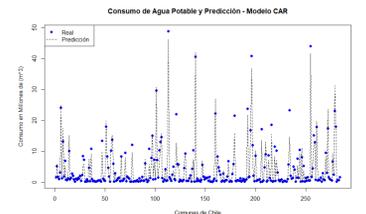


Resultados de Modelos de Predicción



Modelo Lineal Generalizado (Nelder & Wedderburn, 1972)

Modelo Espacio Temporal BYM (Besag, York, & Mollié, 1991)



Modelo Espacio Temporal CAR (Besag, 1974)

Finalmente, después realizar el análisis de conciliación por diferentes métodos entre el valor real del consumo de agua potable (AP) y el valor estimado por los tres modelos implementados en esta investigación se concluye que el modelo BYM (Besag, York, & Mollié, 1991) presenta la mejor estimación a nivel local y global con varianza de error cercana a cero de la variable estimada, el consumo de agua potable en Chile a nivel comunal.

Conclusión

Los resultados obtenidos en este estudio pueden tener implicaciones importantes en las decisiones de algún organismo del estado que lo ayude a identificar sobreconsumos de agua potable de alto riesgo, y poder tomar decisiones con el objetivo de prevenir y controlar el consumo de agua potable en el futuro en las diferentes comunas de Chile.

Entre algunas de las limitaciones en este estudio es que pueden existir, falta de información del consumo de agua potable en algunas comunas no identificadas.

El índice de nivel socioeconómico y la población fueron las covariables más significativas en los ajustes de los tres modelos implementados en esta investigación para el pronostico del consumo de agua potable.

Con esta investigación, se logró demostrar que, utilizando el efecto espacial y específico de cada una de las comunas, se presentan mejores estimaciones que un ajuste de un modelo tradicional.