

Assessment of the importance of including meteorological variables in models for predicting hospitalization due to respiratory diseases

Gregori Moreira, Alexandre Cacheffo, Andrés Pereira, Antônio Gomes, Dirk Thielen, Ediclê Duarte, Eduardo Landulfo, Ezequiel Z. Ledesma, Fábio Lopes, Henrique Aveiro, Jonatan Silva, Juan Pallota, Juan Rascado, Laura Pelicer, Luiz Souza, Maria Torres, Maria Costa, Marlon Puig, Mary Puche, Náíade Rosa, Paloma González, Samara Carbone, Thiago Diaz





PROJETO
CAIPORA



Introducción

■ Introducción

Modelo Univariado

Es una herramienta que analiza solo una variable para hacer predicciones. Se centra en cómo se comporta esta variable sola, sin considerar la influencia de otras variables.

Ej: ARIMA, SARIMA, Suavizado exponencial, TBATS, RNN, LSTM, etc.

Sencillez

- Son sencillos y fáciles de implementar e interpretar, lo que los hace accesibles incluso para principiantes en el análisis de datos.

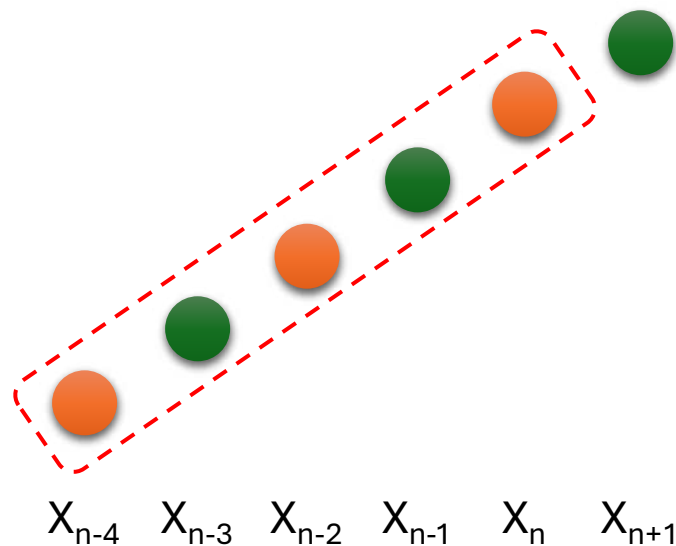
Eficiencia computacional

- Dado que solo se considera una variable, se reduce la carga computacional, lo que puede ser crucial para conjuntos de datos grandes.

Facilidad de comunicación

- Los resultados de un modelo univariado son generalmente más fáciles de comunicar y comprender, lo que facilita la toma de decisiones.

■ Introducción



■ Introducción

Modelo Multivariado

Es una herramienta que utiliza múltiples variables para predecir o analizar resultados, permitiendo analizar la interacción e impacto de varias variables simultáneamente.

Ej: Random Forest, XGBoost, LSTM, RNN...

Análisis de interacción

- Permite estudiar cómo diferentes variables influyen en un resultado, teniendo en cuenta sus interacciones.

Modelado de relaciones complejas

- Capaz de modelar relaciones no lineales y complejas entre variables, como es el caso en muchos problemas del mundo real.

Pronóstico más preciso

- Al considerar múltiples variables, los resultados de las previsiones pueden ser más precisos, reduciendo errores e incertidumbres.

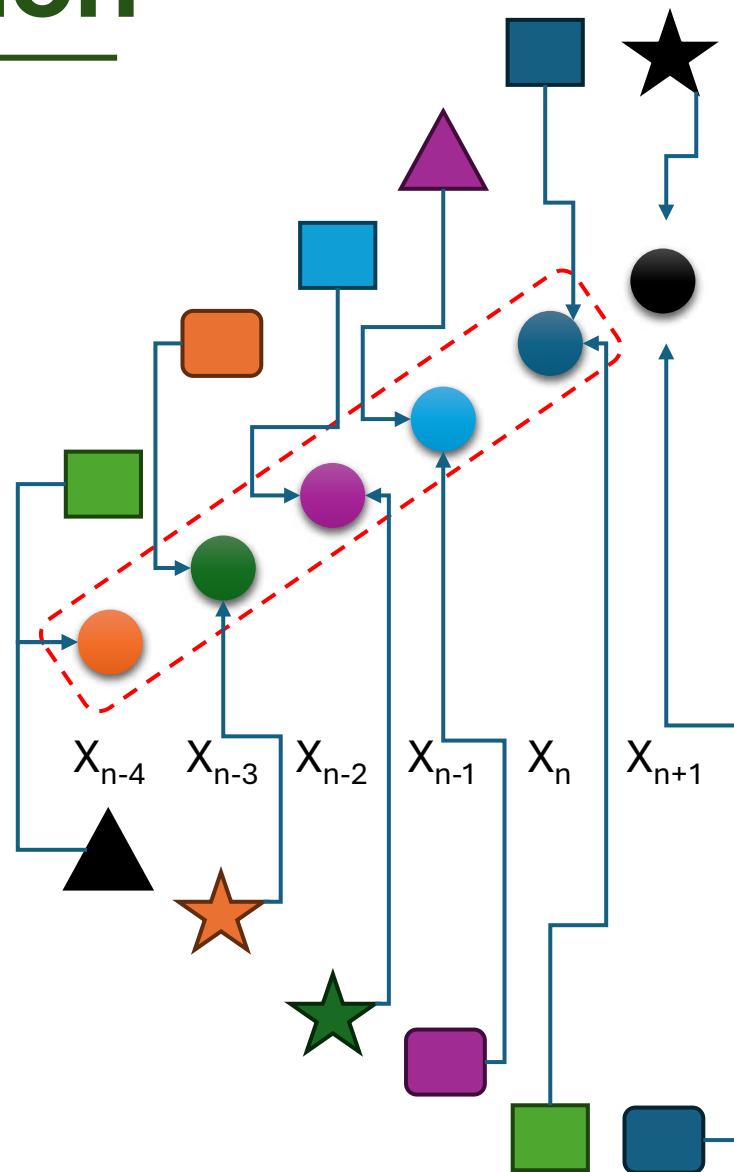
Detección de patrones ocultos

- Pueden revelar patrones y relaciones que no serían evidentes en análisis univariados.



PROJETO
CAIPORA

Introducción



■ Introducción

Cuáles lo mejor?

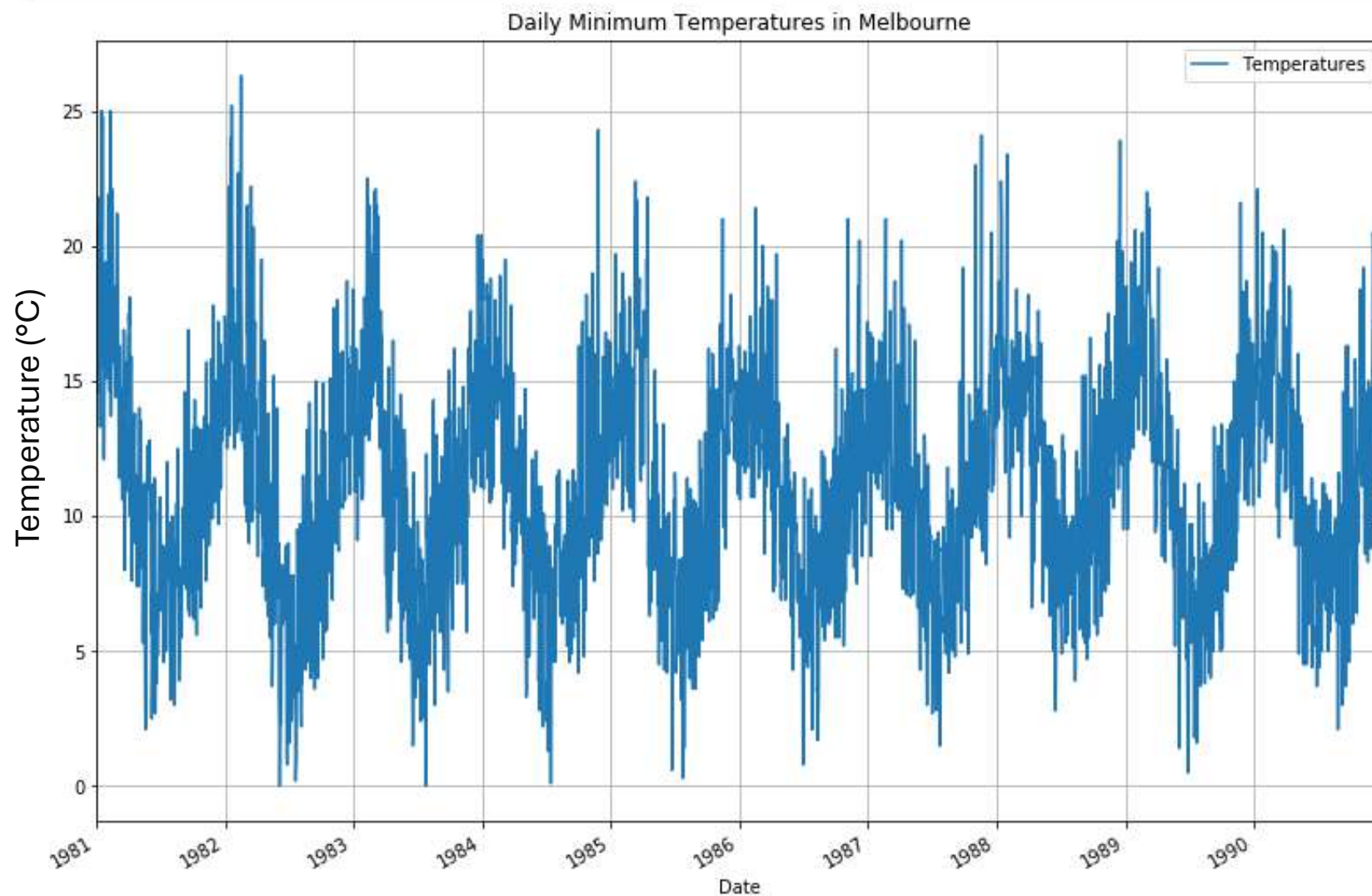




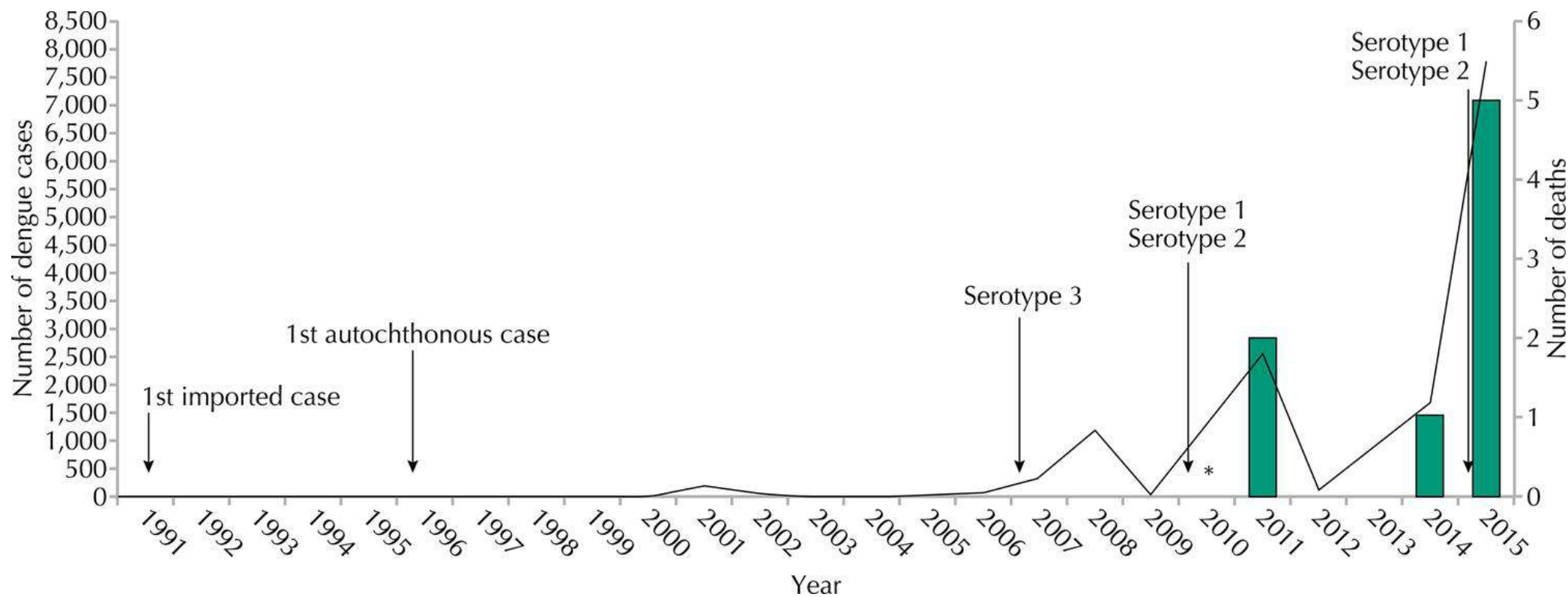
PROYECTO
CAIPORA



Introducción



Introducción





PROJETO
CAIPORA



Estudios de Caso

■ Objetivo General

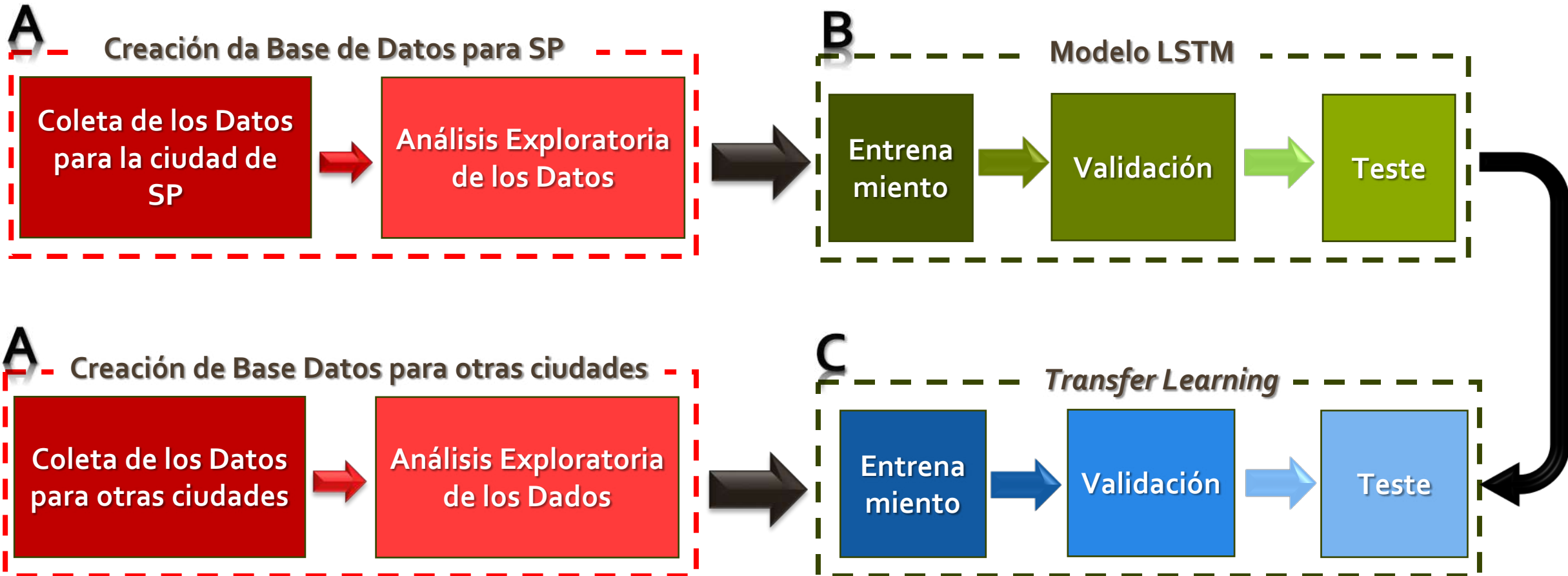
- Desarrollar, probar y validar un **modelo predictivo** que, a partir de series temporales de **variables meteorológicas, concentración de contaminantes e indicadores sociales**, sea capaz de predecir, para las regiones metropolitanas brasileñas, **casos de mortalidad y hospitalizaciones por enfermedades cardiovasculares, cerebrovasculares y enfermedades respiratorias**, así como hospitalizaciones de pacientes, ambas registradas en los distintos sistemas de Bases de Datos del Departamento de Tecnología de la Información del SUS (DATASUS), tomando como referencia la CIE-10

■ Estudios de Caso

• Estaciones de monitoreo en Brasil

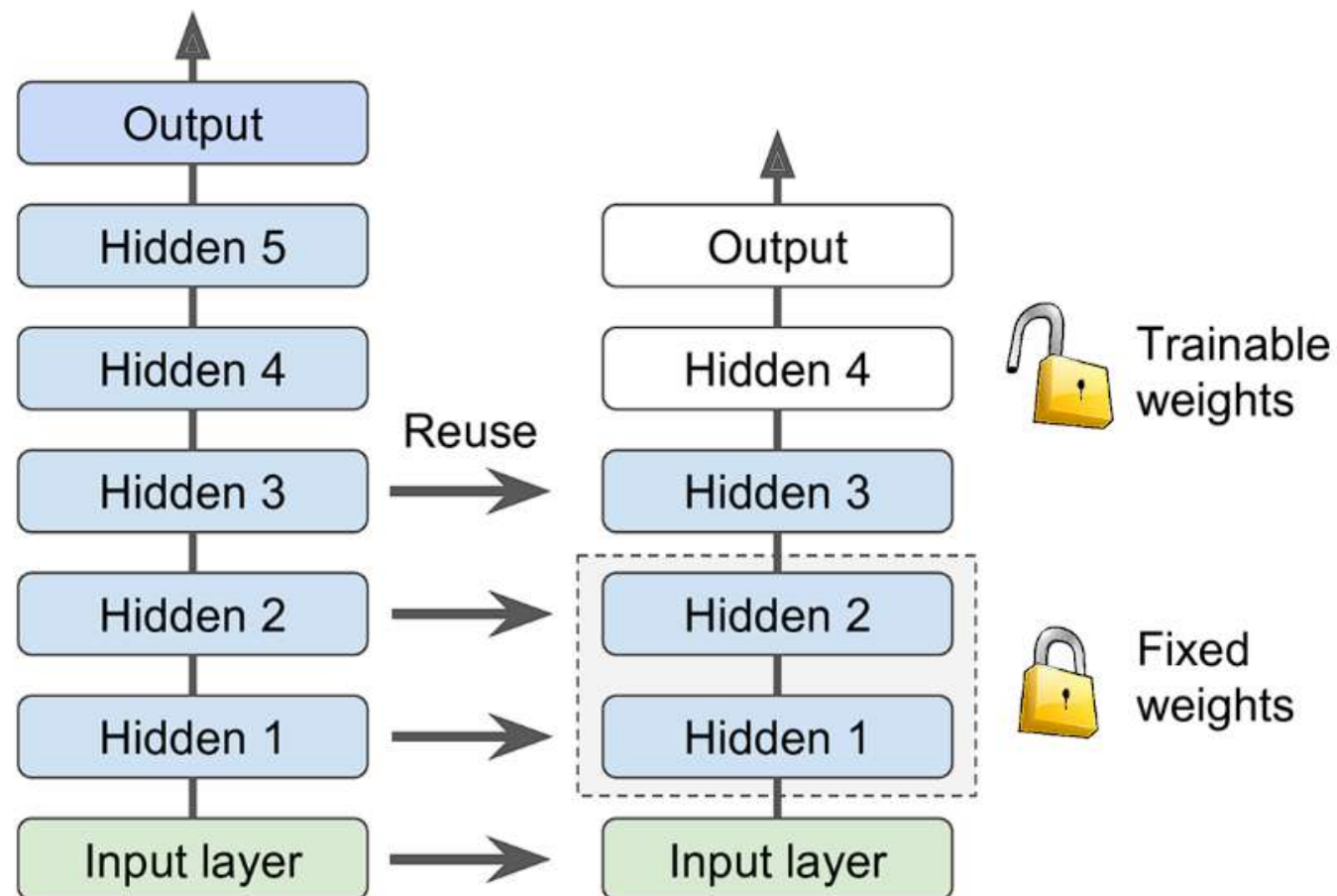


Metodología



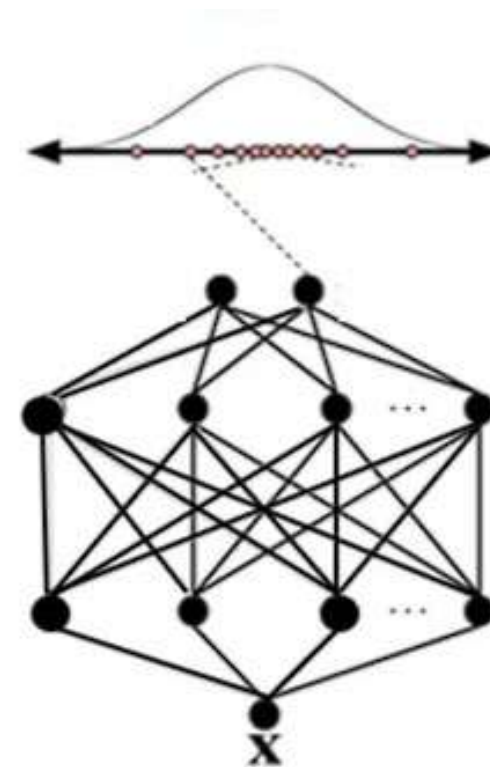
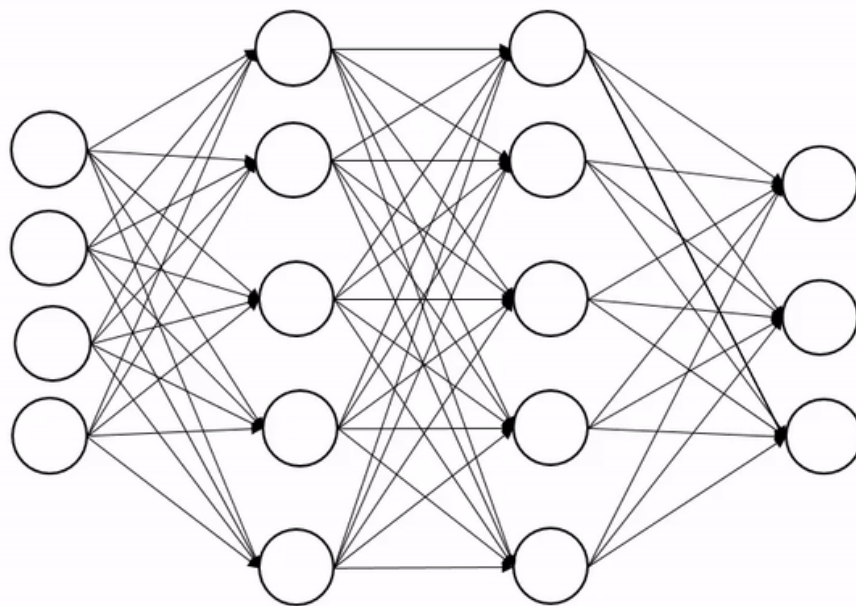
Estudios de Caso

Transfer Learning



■ Estudios de Caso

Monte Carlo - Dropout



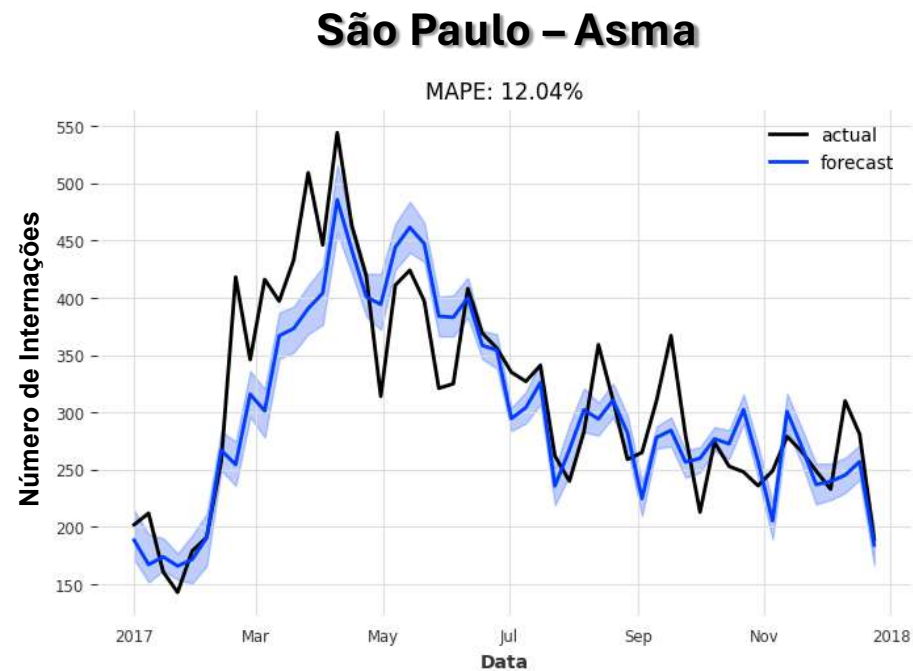


**PROJETO
CAIPORA**



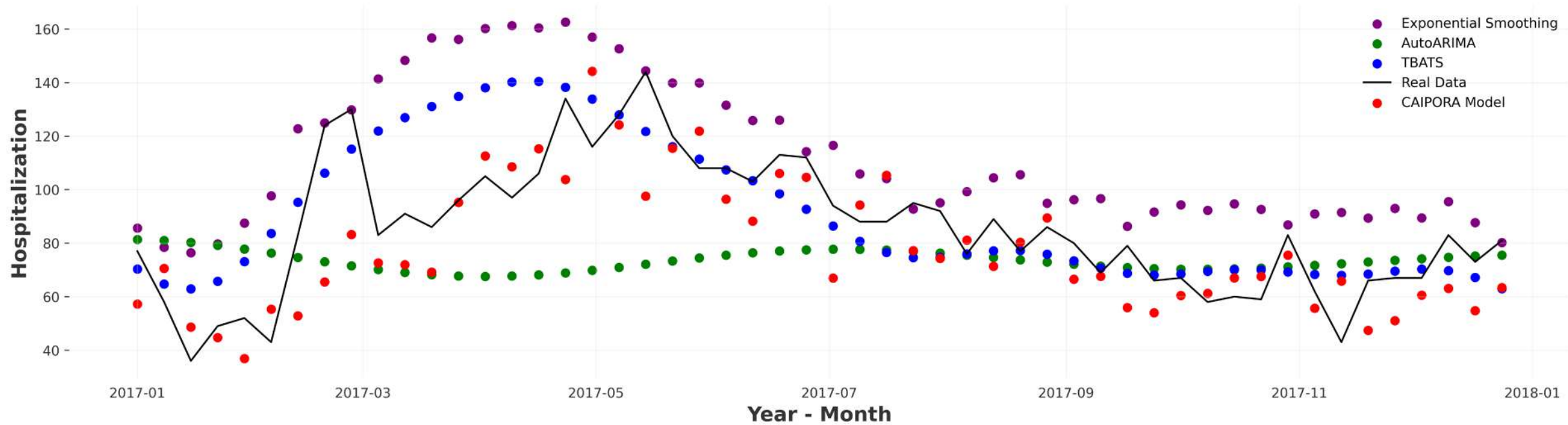
Resultados

Comparaciones

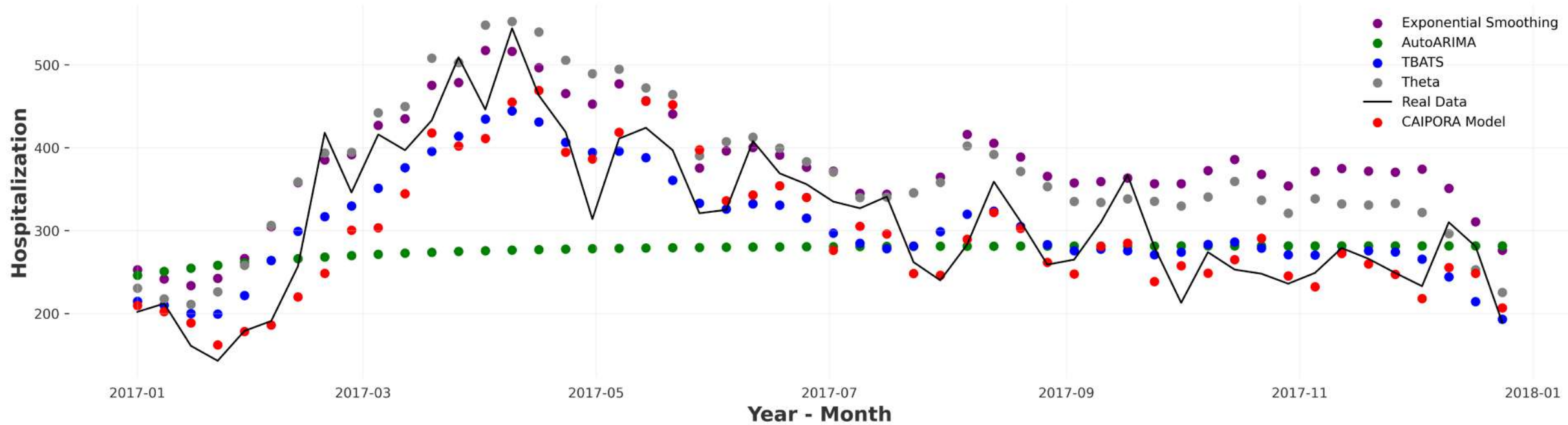


Modelo	MAE	R ²
CAIPORA (LSTM)	33	0,73
TBATS	37	0,73
Exponential Smoothing	69	0,20
Theta	71	0,36
AutoARIMA	73	-0,12

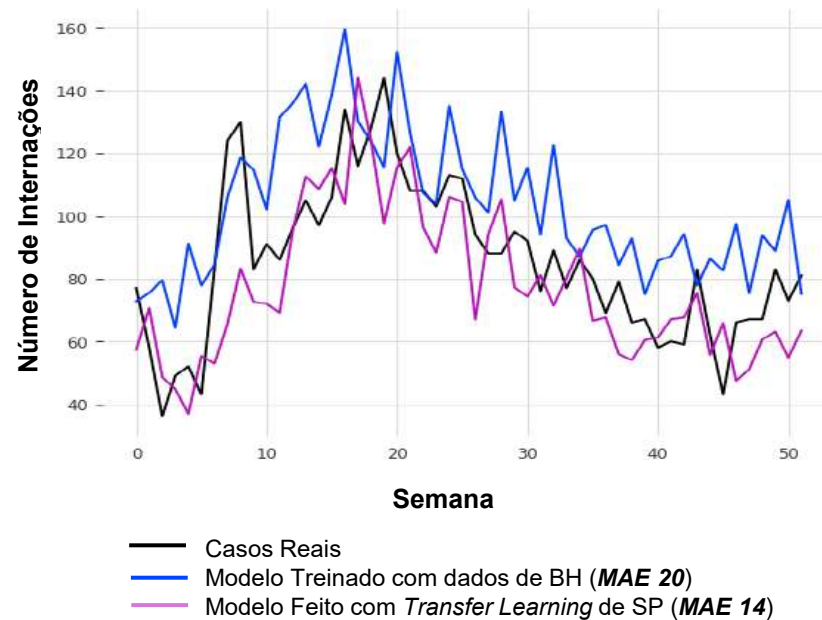
Comparaciones



Comparaciones



Belo Horizonte - 2018



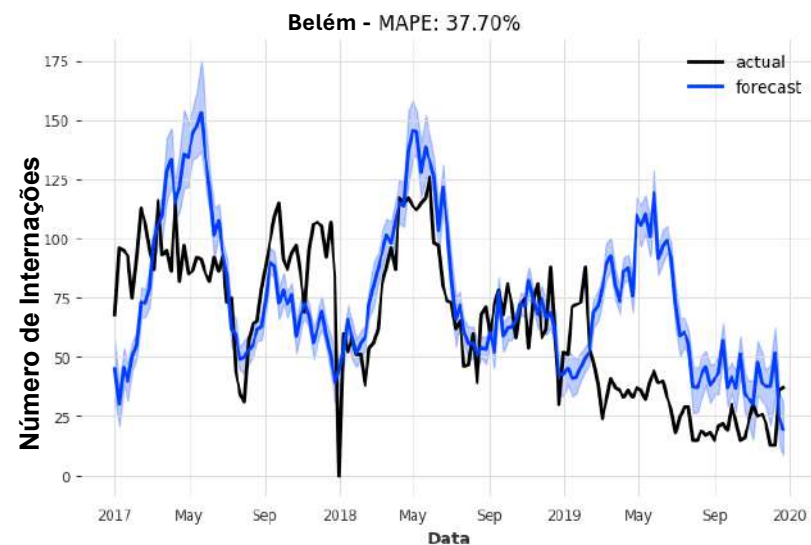
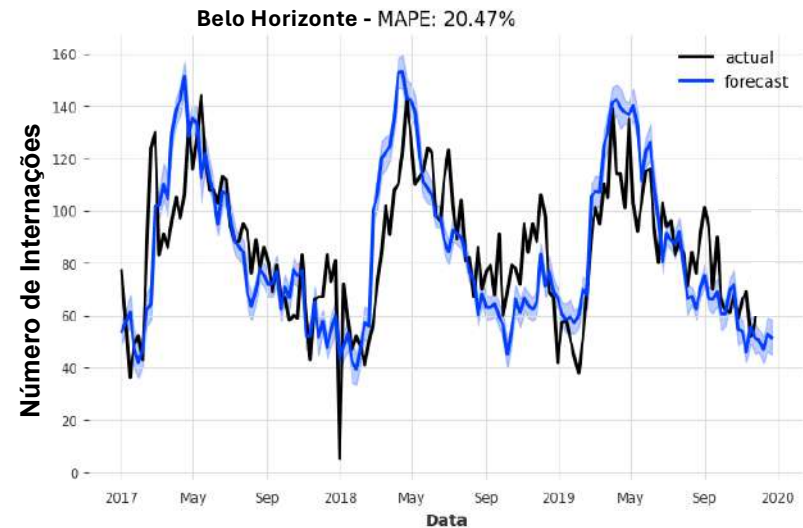
Pros

- Mejora la performance del Modelo
- Posible identificación de subnotificaciones

Contras

- Alto custo computacional
- Limitación de las features del modelo base

Comparaciones



Capital	MAE	MAPE
São Paulo	46	16,10 %
Belo Horizonte	86	20,47 %
Manaus	11	32,15 %
Belém	25	37,70 %



PROJETO
CAIPORA



Conclusión

■ Conclusión

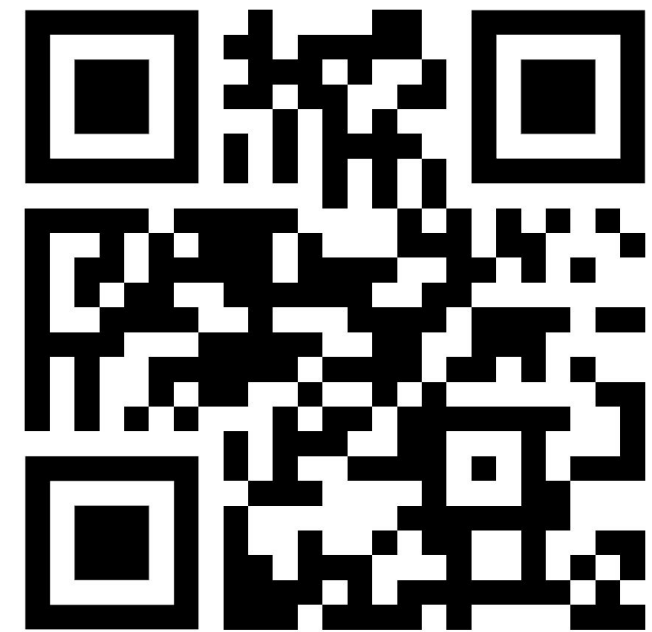
- Los enfoques univariados y multivariados tienen sus ventajas y desventajas;
- Los modelos univariados proporcionan simplificaciones que pueden pasar por alto características importantes;
- Los modelos multivariados permiten una mejor detección de la dinámica real;
- El Transfer Learning combinado con datos satelitales y/o reanálisis son alternativas cuando existe una gran ausencia de datos;



Medellín, Colombia

CASAP X

CONGRESO COLOMBIANO Y
CONFERENCIA INTERNACIONAL
DE CALIDAD DE AIRE, CAMBIO CLIMÁTICO Y SALUD PÚBLICA



¡Gracias!

Más información



<https://casap.science/>



casap@casap.science