12th Seminar for Homogenization and Quality Control in Climatological Databases and the 7th Interpolation Conference, followed by the Danube-Adapt Project meeting.

Reconstruction of Maximum Temperature Time Series Using Machine Learning Models

Eduarda Agnolin

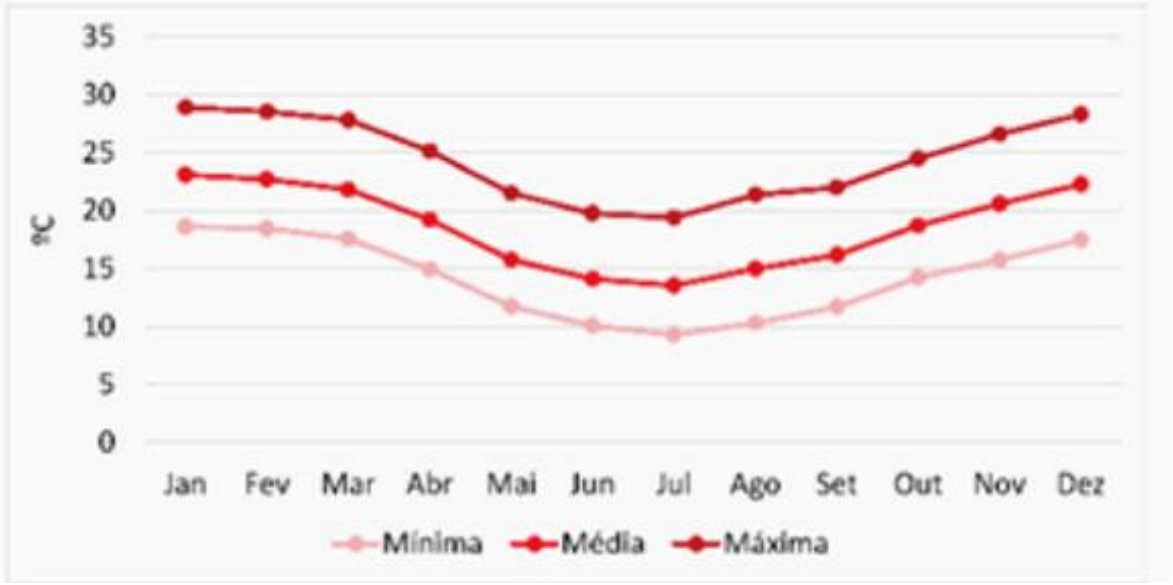
Fiorella Acquaotta

Introduction

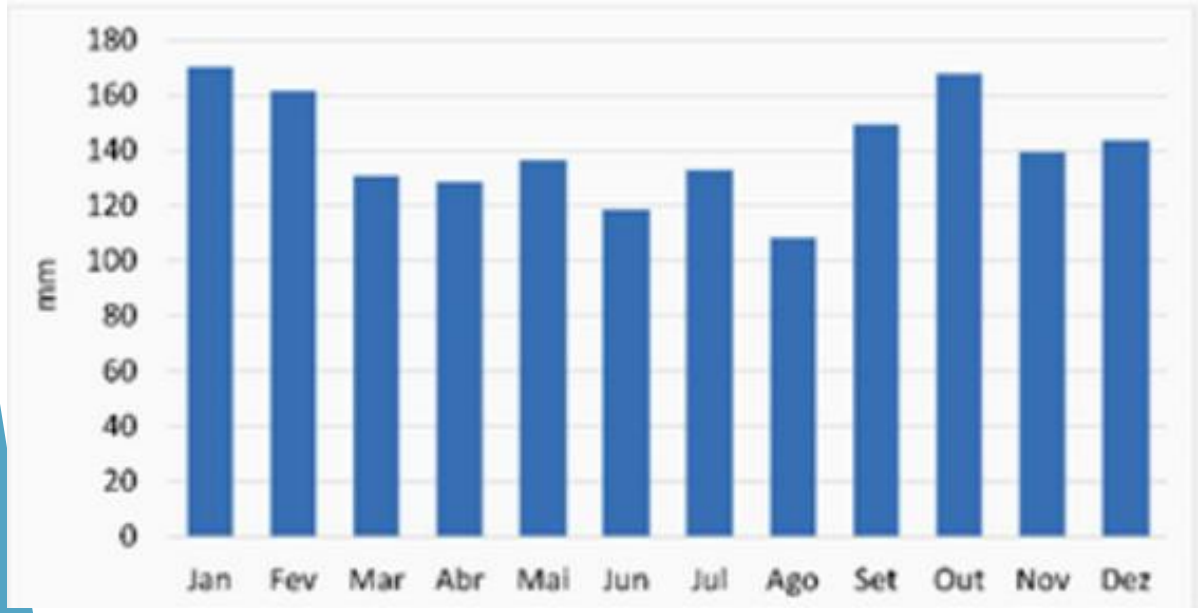
- The presence of gaps in meteorological series is a recurring problem in climate studies, potentially compromising the analysis of extreme events such as heat waves. In this context, this study aimed to evaluate different data reconstruction methodologies for filling gaps in maximum temperature series.

Location of the study area

Monthly temperature in the subtropics (1981-2010)



Monthly precipitation in the subtropical zone(1981 -2010)



Map of Brazil showing Rio Grande do Sul and Santa Catarina.

World map with Brazil

Methodology

Table 01 - Weather stations used

Station	Number of failures in the series	% of failures in the series
Florianópolis	360	5,48 %
Porto Alegre	26	0,40%

Number of failures in the series per year

Years	Porto Alegre	Florianópolis
2008	6	1
2009	1	2
2010	0	0
2011	0	0
2012	0	4
2013	5	1
2014	0	0
2015	0	21
2016	5	0
2017	1	0
2018	0	0
2019	0	0
2020	0	0
2021	4	0
2022	0	0
2023	2	0
2024	0	101
2025	2	230

Methodology

Multiple Linear Regression (MLR): statistical method that models a linear relationship between a dependent variable and multiple independent variables to realize predictions.

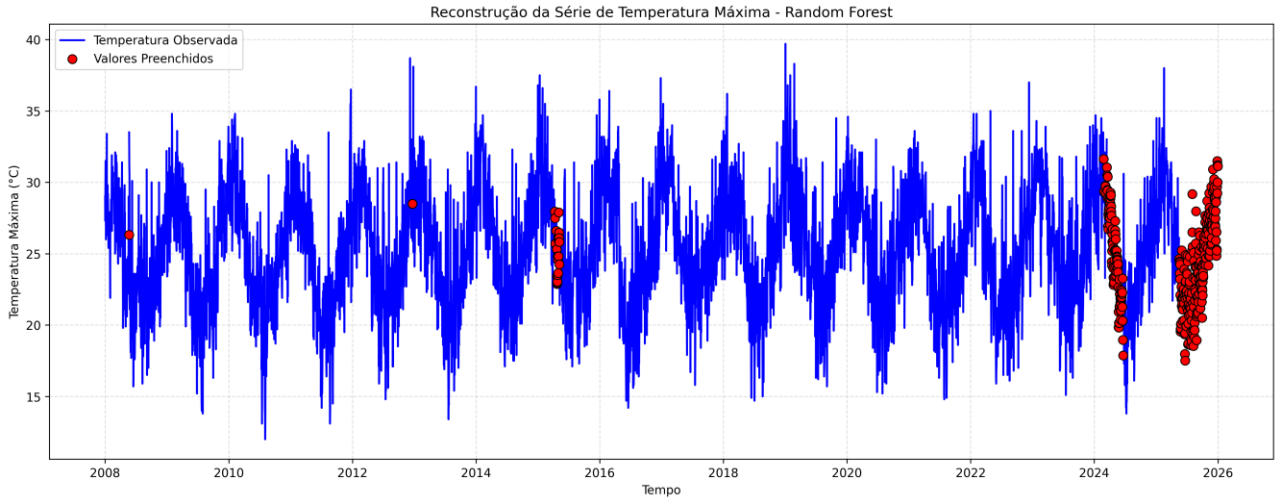
Random Forest (RF): Machine-learned algorithm based on multiple decision factors that combines multiple results to improve accuracy and reduce overfitting.

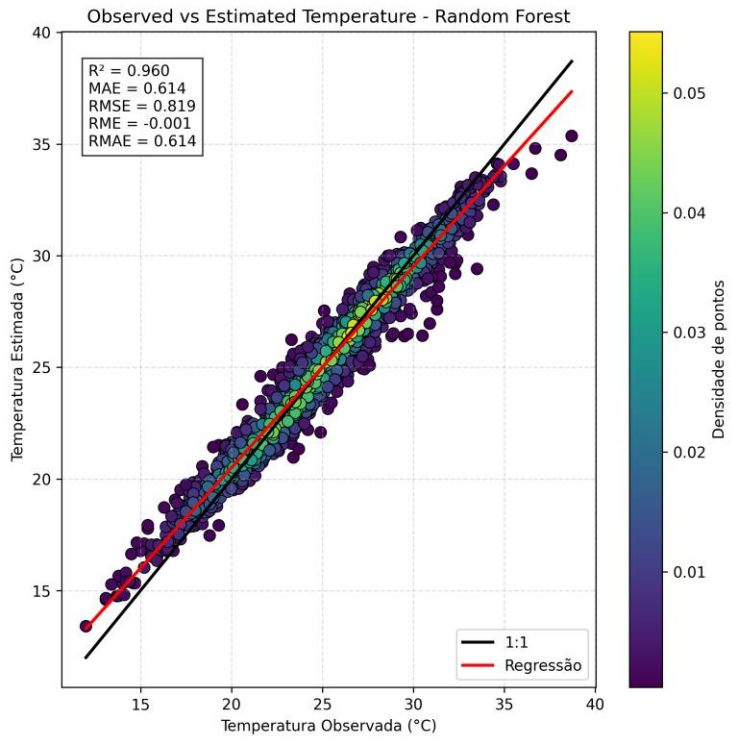
Multilayer Perceptron (MLP): type of artificial neural network composed of interconnected neuron cells capable of learning complex relationships and not linear data.

Support Vector Regression (SVR): Learned machine technique derived from support vector machines that perform regression to find a function that minimizes errors within a tolerated range.

Randon Forest model

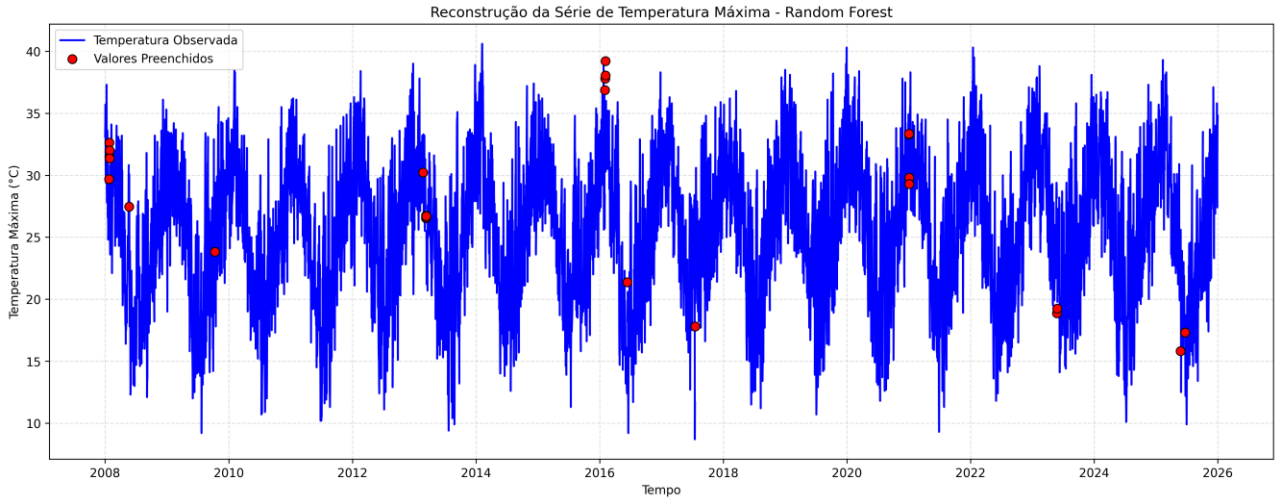
Florianópolis

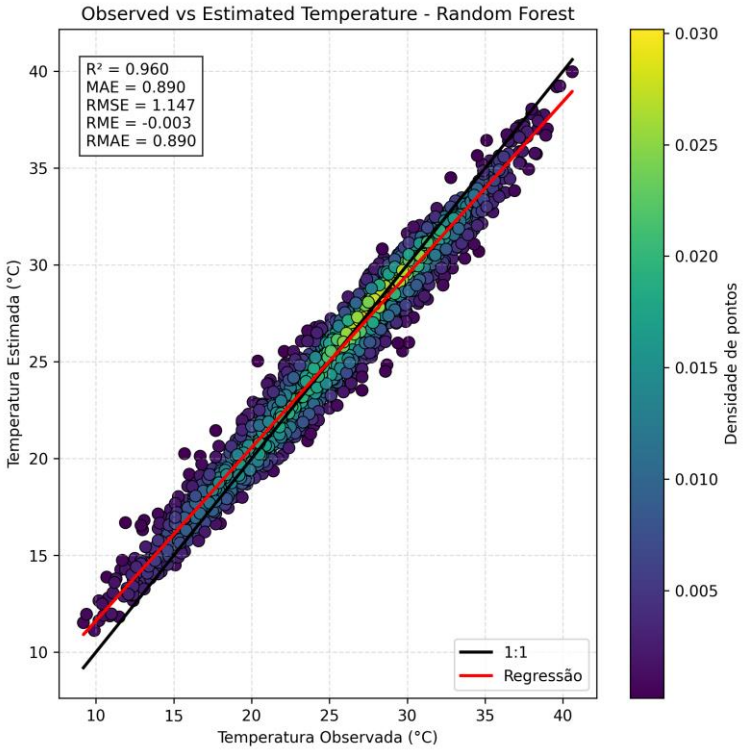




$R^2 = 0.960$
MAE = 0.614
RMSE = 0.819
RME = -0.001
RMAE = 0.614

Porto Alegre

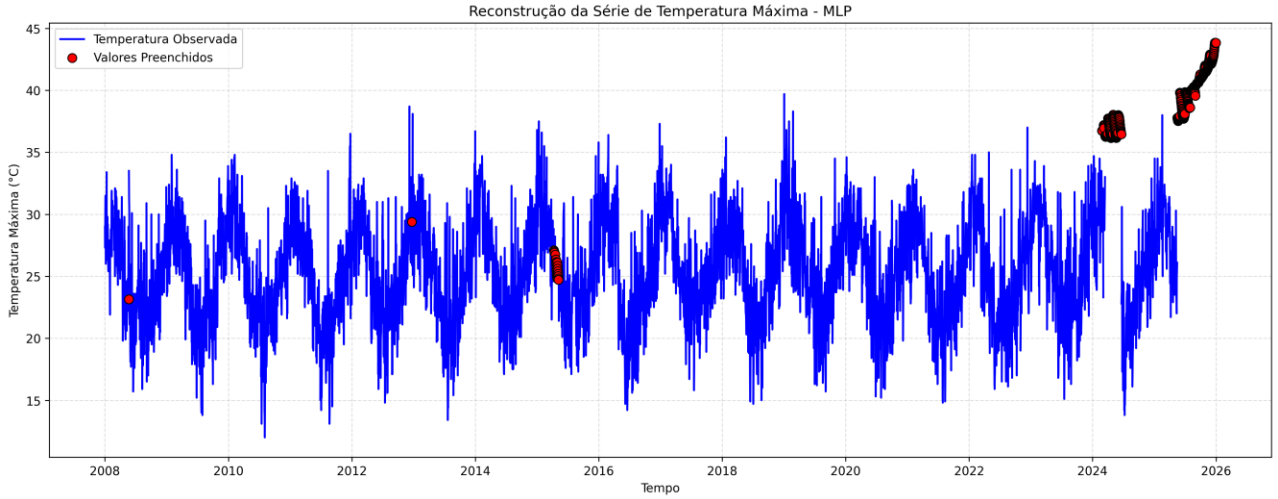


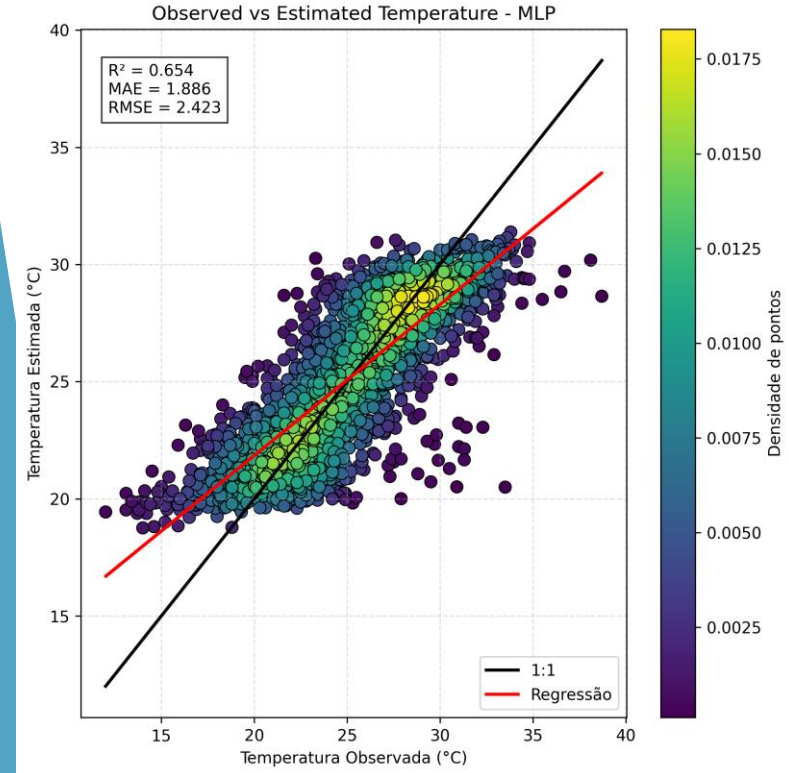


$R^2 = 0.960$
MAE = 0.890
RMSE = 1.147
RME = -0.003
RMAE = 0.890

Multilayer Perceptrons (MLP)

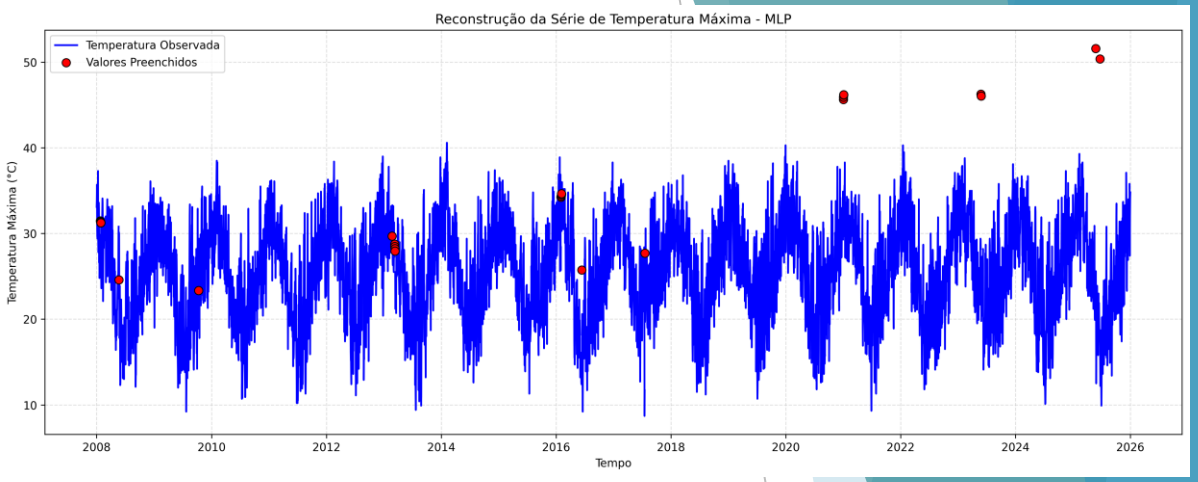
Florianópolis

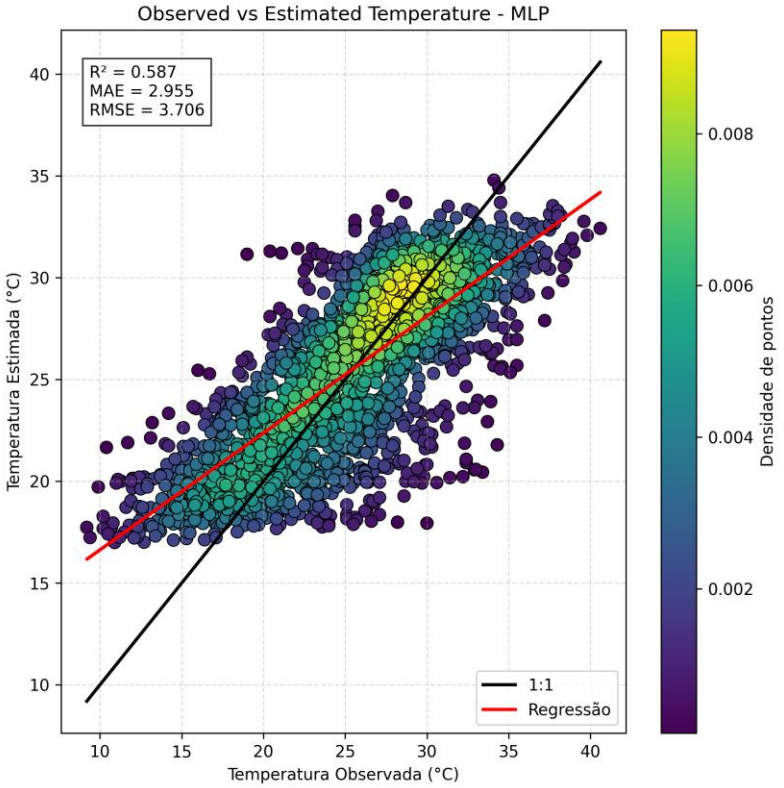




$R^2 = 0.654$
$MAE = 1.886$
$RMSE = 2.423$

Porto Alegre

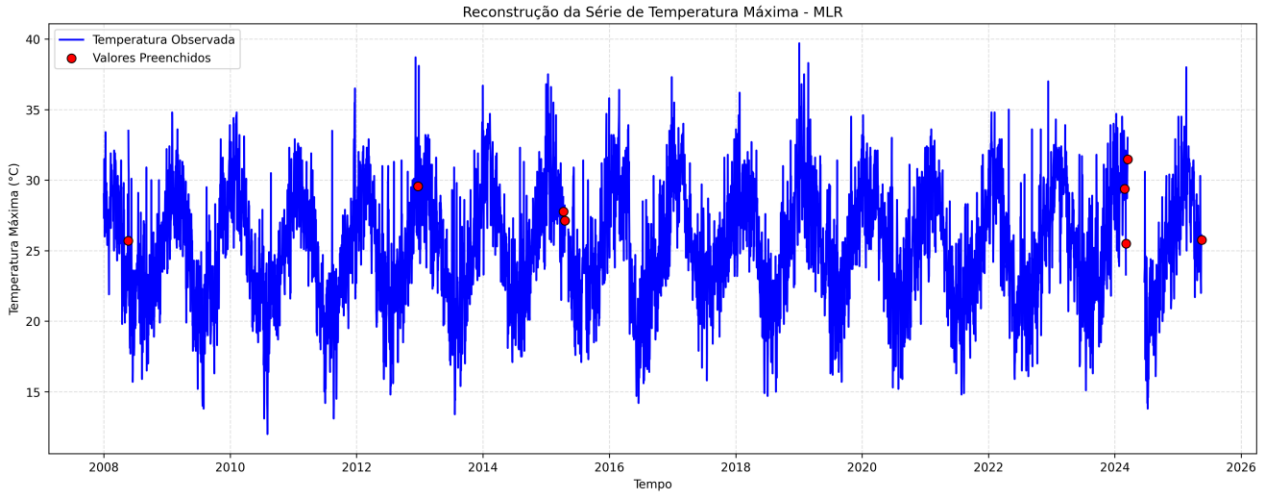


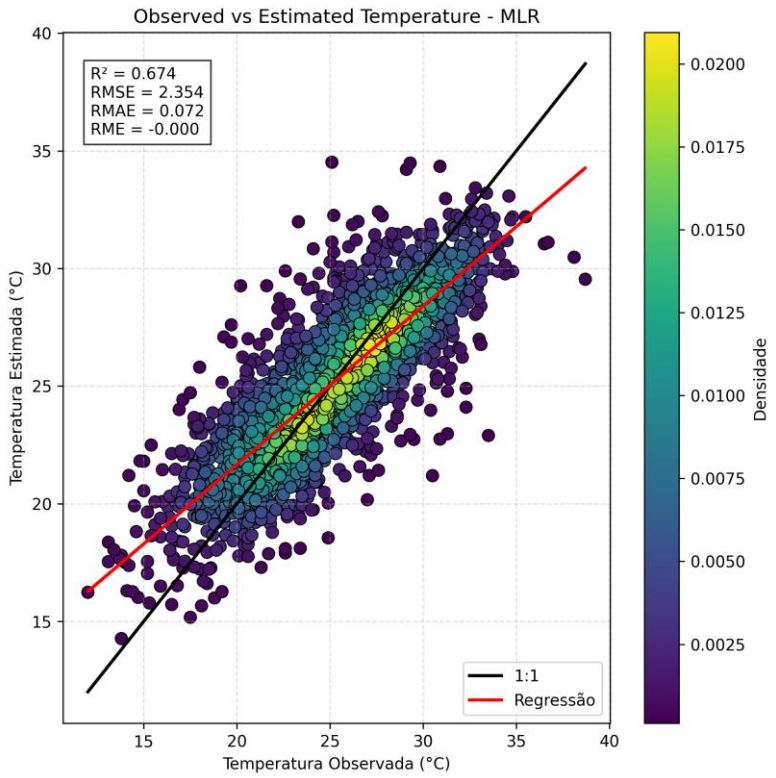


$R^2 = 0.587$
$MAE = 2.955$
$RMSE = 3.706$

MLR (Multiple Linear Regression)

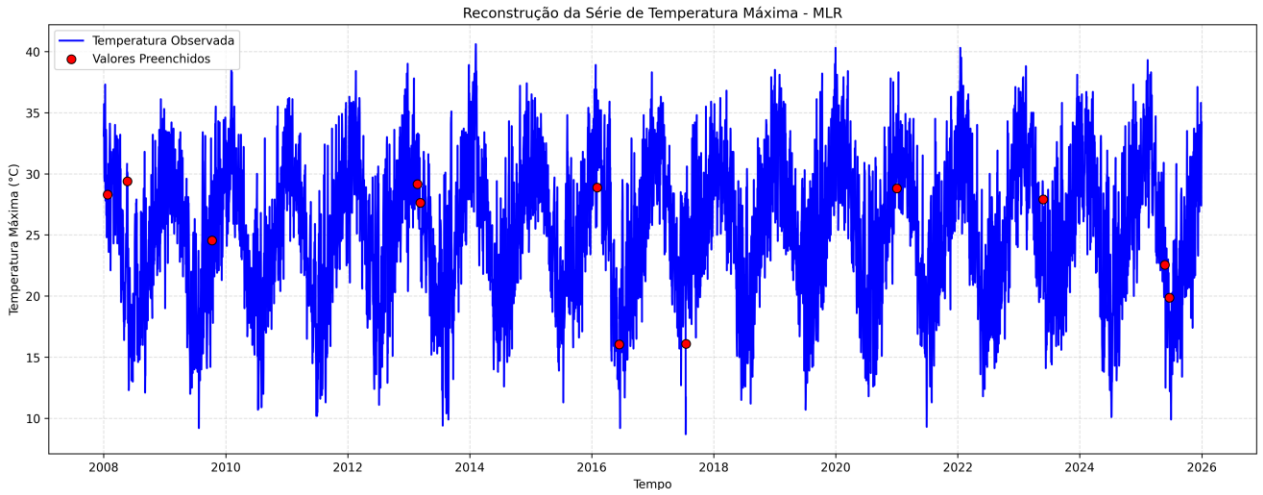
Florianópolis

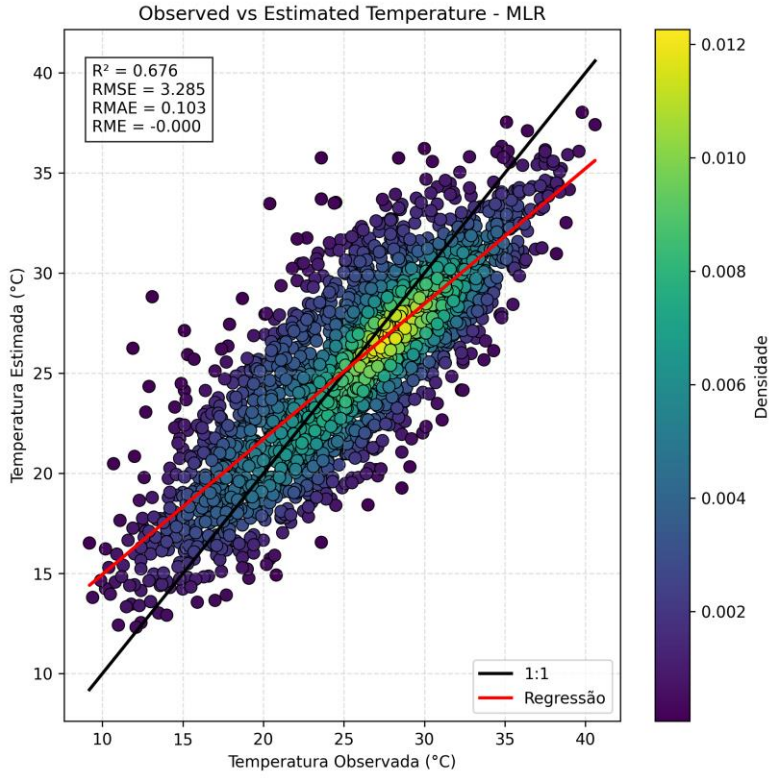




$R^2 = 0.587$
MAE = 2.955
RMSE = 3.706

Porto Alegre

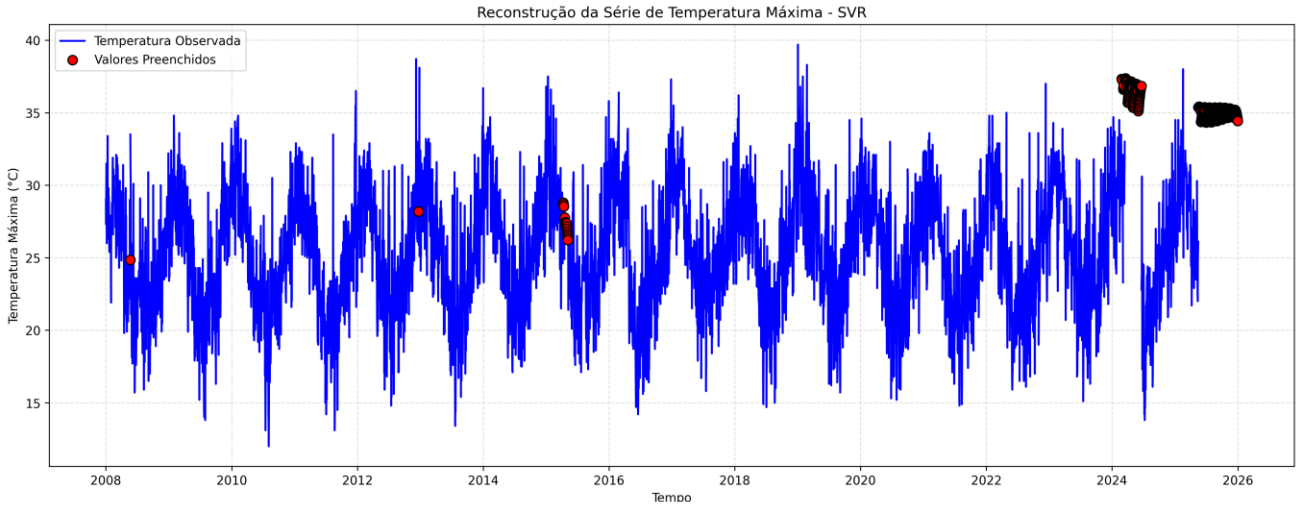


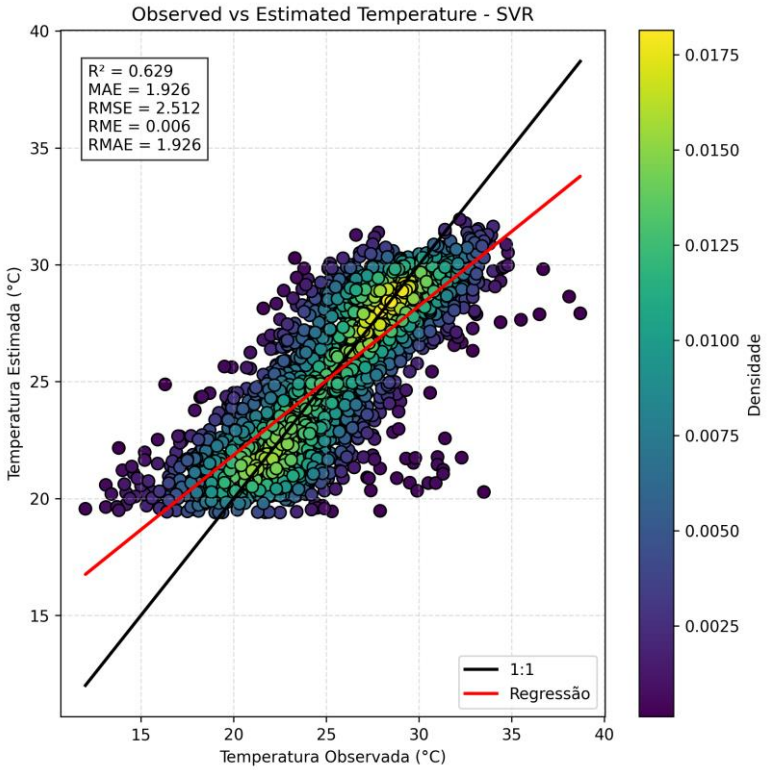


$R^2 = 0.676$
RMSE = 3.285
RMAE = 0.103
RME = -0.000

SVR (Support Vector Regression)

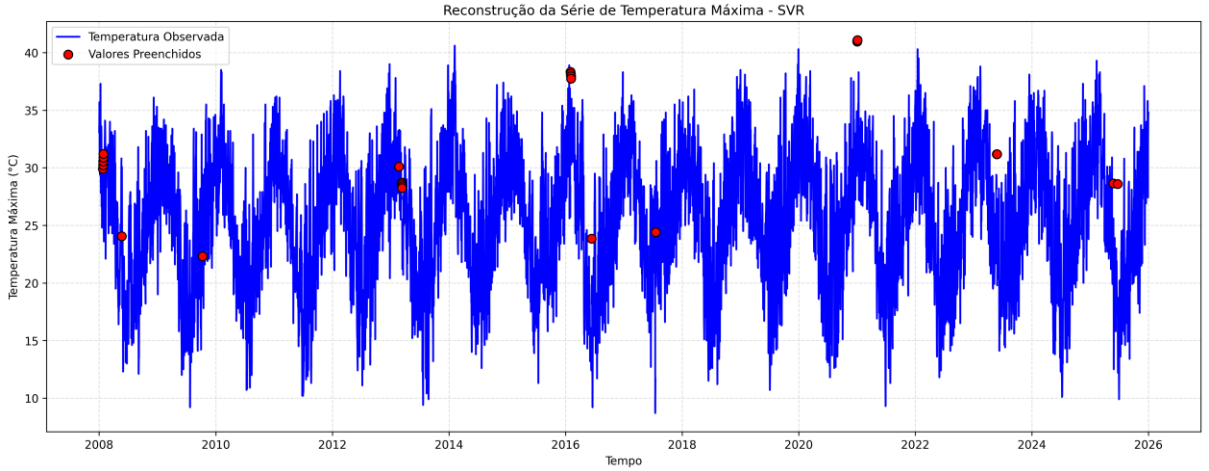
Florianópolis

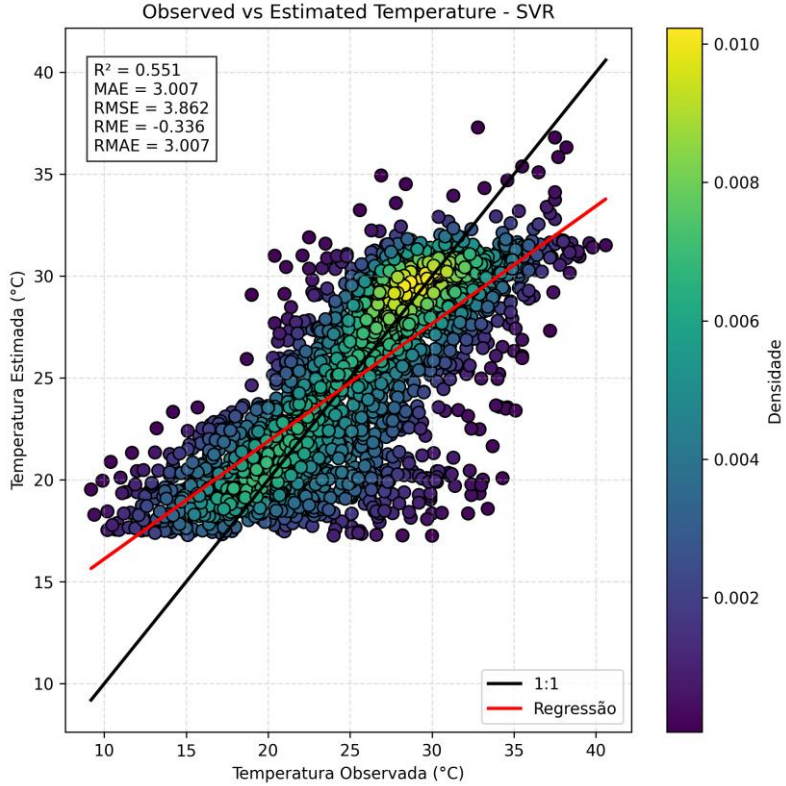




$R^2 = 0.629$
MAE = 1.926
RMSE = 2.512
RME = 0.006
RMAE = 1.926

Porto Alegre





$R^2 = 0.551$
MAE = 3.007
RMSE = 3.862
RME = -0.336
RMAE = 3.007

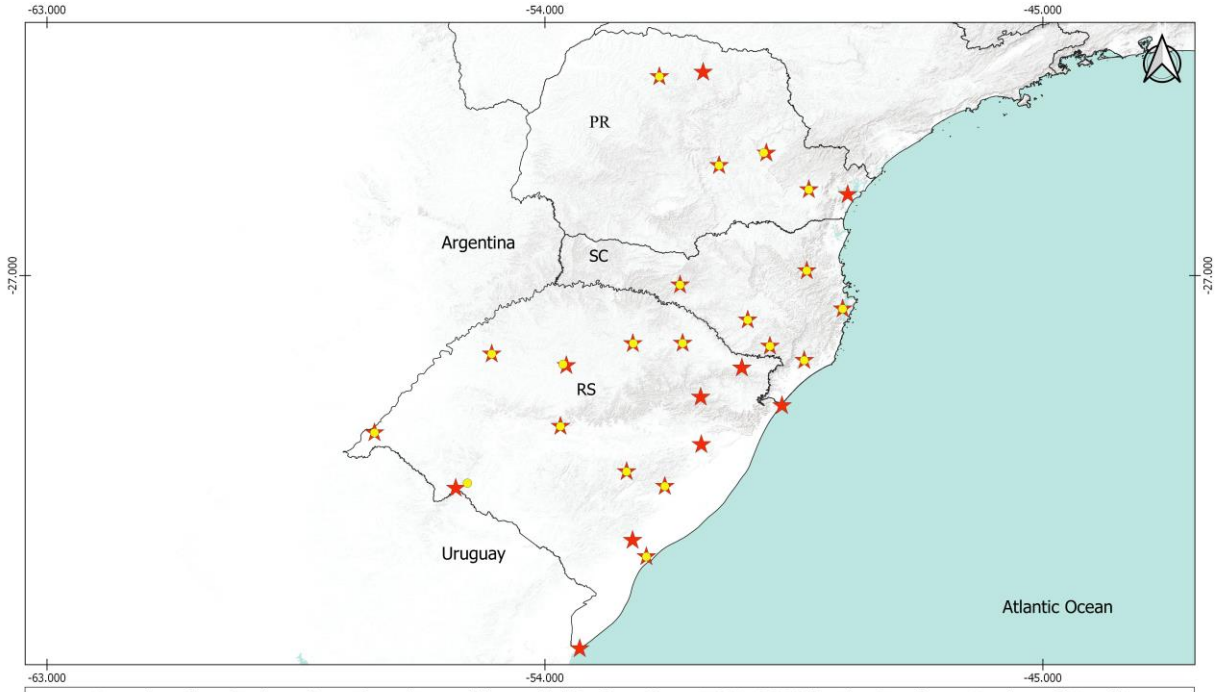
Results:

Table 02 - Model trained on the (2009 - 2014) dataset and applied to the reconstruction of the (2008 - 2025) series.

Model	Estação	R^2	MAE	RMSE	RMAE	RME
MLR	Florianópolis	0.67	1.80	2.35	0.07	-1.07
	Porto Alegre	0.67	2.59	3.28	0.10	-4.62
RF	Florianópolis	0.96	0.61	0.81	0.61	-0.00
	Porto Alegre	0.96	0.88	1.14	0.88	-0.00
MLP	Florianópolis	0.65	1.88	2.42	1.88	-0.03
	Porto Alegre	0.58	2.95	3.70	2.95	-0.12
SVR	Florianópolis	0.62	1.92	2.51	1.92	0.00
	Porto Alegre	0.55	3.00	3.86	3.00	-0.33

Work in Progress

- The research will continue using the thirty stations. I will reconstruct them using Random Forest and then homogenize by Climatol. After that, I will use Climpact to perform a climate analysis of heat waves and cold spells in the subtropics from Nineteen ninety-one - two thousand twenty-five.



Map showing the location of stations with available data from 1991-2020 in the Southern Region of Brazil.

Legend:

Automatic station | Conventional stations | State limit

Preparation: the authors (2026).

Scale:
0 100 200 km

Cartographic and thematic data:

Cartographic Database
Brazilian National Institute of Meteorology (INMET, 2026).
Brazilian Institute of Geography and Statistics (IBGE, 2026).

References

- He, M., Luo, Z., Xie, X., Wang, P., Wang, H., & Zapata-Lancaster, G. (2025). Gap filling crowdsourced air temperature data in cities using data-driven approaches. *Building and Environment*, *271*, 112593.

- da Silva Junior, A. A., Gomes, R. D. S. R., De Musis, C. R., Novais, J. W. Z., Maionchi, D., & de Figueiredo, J. M. (2024). Preenchimento de falhas em séries temporais da temperatura do ar: uma comparação entre modelos de Machine Learning. *Revista Brasileira de Climatologia*, *35*, 362-377.

- GUIJARRO, Jose A. Homogenization of climatic series with Climatol. **Reporte técnico State Meteorological Agency (AEMET), Balearic Islands Office, Spain**, 2018.

Thank you

E-mail: eduardagnolin@gmail.com

UNIVERSIDADE FEDERAL
DE SANTA CATARINA

LabClima
Laboratório de Climatologia Aplicada
Departamento de Geociências - UFSC

CAPES