

Forecasting traffic flow time series with Vine-Transform ARMA Copula models

Sara Selvaggia Guerini

saraselvaggia.guerini@unibg.it

University of Bergamo

Rodolfo Metulini

University of Bergamo

Research Article

Keywords: Vine-Transform ARMA Copulas, Rank-Graduation Approach, Predictive Performance, Time Series Modeling, Urban Traffic Analysis

Posted Date: October 16th, 2025

DOI: <https://doi.org/10.21203/rs.3.rs-7011611/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Forecasting Traffic Flow Time Series with Vine-Transform ARMA Copula Models

Abstract

The prediction of traffic flows in urban areas is gaining importance in order to carry out urban planning for early warning systems and optimized logistics. Hence, there is a growing need for simple and high-performing statistical models. This study leverages the vine-transform autoregressive moving-average copula model to predict traffic data, with particular emphasis on evaluating forecasting performance. To do so, real-life data on origin-destination signals extracted from mobile phone data have been used. Performance evaluation was conducted using the rank-graduation box approach along with a moving window cross-validation strategy, incorporating rank-graduation accuracy for precision and rank-graduation explainability for component analysis. As a benchmark for comparison, the VARX-DHR model was used. Preliminary results reveal that the vine-transform ARMA copula approach performs well in terms of accuracy. Furthermore, we show that the copula component presents greater explainability compared to the autoregressive and moving average components. Residual diagnostics show significantly lower autocorrelation and partial autocorrelation with respect to the original data, and that they are approximately normally distributed. The developed method could provide valuable insights supporting urban planners and analysts in making informed decisions.

Keywords: Vine-Transform ARMA Copulas, Rank-Graduation Approach, Predictive Performance, Time Series Modeling, Urban Traffic Analysis

1 Introduction

The study of traffic in urban areas has become increasingly important in recent years for various aspects of life in smart cities, playing a critical role not only in optimizing transportation systems (Tao et al. 2023) but also in enhancing urban planning (Shahriari et al. 2023), improving emergency response strategies (Yuan et al. 2012), and supporting environmental sustainability initiatives (Zhang et al. 2017). Studying human mobility also helps track disease spread, identify high-risk areas, and design effective COVID-19 containment strategies (Ahdika et al. 2023; Guardabascio et al. 2024). Recently, research on this topic has proliferated, with an increasing number of studies proposing methods to model and forecast traffic (often validated through specific case studies) at the intersection of informatics, machine learning, statistics, and

operations research (Fernández-Ares et al. 2017; Alam et al. 2019; Nagy and Simon 2021). Overall, accurately predicting the movement of people between urban areas is essential for policy makers to make informed and effective decisions (World Health Organization (2022)).

Various types of data are commonly used to proxy urban traffic flows. One of the most traditional sources of traffic data comes from roadside sensors and inductive loop detectors embedded in roads. These devices provide real-time information on traffic volume, speed, and occupancy, which is essential to assess traffic flow (Po et al. 2019). Another increasingly valuable source of data is GPS information collected from vehicles, smartphones, and ride-sharing services. GPS data offer detailed information on vehicle speeds, travel times, and routes, allowing for the modeling of traffic patterns across various times of day and geographic locations (Kan et al. 2019). Giorgini and Sartori (2016) analyze vehicle and pedestrian movements from GPS and video recordings related to urban traffic in Torino, Marseille, and Venice. Platforms like Twitter, Waze, and Google Maps have emerged as valuable sources of crowdsourced traffic data. Social media offers real-time, user-generated reports on traffic conditions, accidents, and road closures (Clemente et al. 2024). This type of data allows for the rapid identification of unforeseen traffic disruptions, providing authorities with the ability to make adjustments in real time. Data from public transportation systems, including buses, trams, and metros, provides valuable insights into traffic demand and congestion. These systems typically generate large amounts of real-time data that can be used to track vehicle locations, travel times, and capacity utilization (Burzacchi et al. 2024).

In this work, we use mobile phone anonymized location data, which has become a powerful tool for forecasting traffic flows. By analyzing the movement patterns of individuals, mobile phone data can provide insights into the flow of people across cities, helping to predict traffic trends and commuter behavior. Mobile phone data are increasingly adopted in research with manifold objectives, from producing dynamic information related to the density of people’s presences (Metulini and Carpita 2021), to analyze movements (Tettamanti and Varga 2014), from monitoring the impact of social and cultural events (Carpita and Simonetto 2014), to the variability in the distribution of presences in the neighborhoods of a large city (Mariotti et al. 2022), the seasonality of the second homes in a tourist area (Curci et al. 2022), or the increase of remote working in sparsely populated areas (Manfredini et al. 2022). The use of mobile phone data is not new to the natural risk literature as well. For example, Balistrocchi et al. (2020) developed spatio-temporal dynamic maps of flood exposure using mobile phone crowding data.

Here we resort on hourly-based origin-destination mobile phone data provided by TIM. The issue of forecasting traffic flow using mobile phone data was considered by proposing a time series modeling strategy based on Vector AutoRegression model with eXplanatories and Dynamic Harmonic Regression (VARX-DHR, hereafter) (Perazzini et al. 2023; Metulini and Carpita 2024). Another class of advanced models such as Facebook Prophet and Neural Prophet (Taylor and Letham 2018; Triebe et al. 2021) have recently gained popularity for their effectiveness in forecasting data with complex seasonality. However, it was demonstrated that its performance in predicting traffic

flows is scarce (Perazzini and Metulini 2024). Traffic flow time series are characterized by complex temporal dynamics that include serial correlation, multiple seasonalities, and non-linearity, meaning that they require robust predictive models capable of capturing such patterns. VARX-DHR is a trivariate model with a double autoregressive structure accounting for, respectively, daily and weekly serial dependence, a dynamic harmonic regression component to account for temporal periodicity, a set of calendar-related dummy covariates, and with cross-correlation among flows modeled via the variance-covariance matrix. The forecast of the model is quite good (despite improvable) but authors found leptokurtic distributions of estimated residuals with heavy tails, which has been used in Carpita et al. (2024) to cluster urban areas in terms of similarities in the tail using a copula approach. Moreover, the model is not parsimonious due to the large amount of covariates, and the non-linearity has not been considered. This leaves ample room for the development of simple models that also effectively consider extreme events (which in our case are very high or very low traffic).

Copula models, which are traditionally effective at capturing extreme events and non-linearities, have been recently applied to the analysis of traffic flow data. Fang et al. (2022) proposed some copula models, such as Gaussian, Farlie-Gumbel-Morgenstern (FGM), Gumbel, Clayton, Frank, Ali-Mikhail-Haq (AMH) copula, to construct bivariate joint distributions of traffic variables. Cervellera et al. (2022) introduced a data-driven method for scenario generation based on copula models, through which the marginals of single input parameters can be chosen freely without altering the joint multivariate dependence structure of the inputs. Lu et al. (2023) introduced the use of a regular-vine copula, combined with a latin hypercube sampling. However, the combination of vine transforms which allow us to map the data to a structure where marginals are independent, and the AutoRegressive Moving Average (ARMA hereafter) component which captures temporal autocorrelation within the vine framework, has never been employed for traffic flow data.

In this study, we apply a recent approach based on copulas, specifically, Vine-Transform ARMA models (from now on, VT-ARMA), proposed by McNeil (2021). Copula-based methods provide a flexible framework for modeling dependencies between variables, which means that they can be suitable for capturing the temporal relationships found in traffic data.

The primary objective of this study is to assess the forecasting performance of the VT-ARMA models on traffic data. In doing so, we compare them against the VARX-DHR, which serves as the benchmark. We assess the forecasting accuracy of the model proposing a blocked k -folds cross-validation strategy for time series, and the Rank-Graduation Accuracy (RGA), in the spirit of Babaei et al. (2025). We also aim to identify the key components of copula models employing their contribution to the model's performance. To these aim, we employ the Rank-Graduation Explainability (RGE). In this study we also conduct a diagnostic analysis of the model residuals to assess the adequacy of the model fit. This involves examining residual plots and applying statistical tests to detect issues such as normality and residual autocorrelation. Periods in which the model performs worse are detected and discussed.

We apply and validate our methodological approach using real-world data through a case study based on a one-year length hourly time series of traffic flows, derived

from mobile phone data for Cellatica, a representative ISTAT "Area di CEnsimento" (ACE) located in the province of Brescia.

The paper is structured as follows. Section 2 describes the data; Section 3 introduces the methods; Section 4 describes the application of the methods to the case study; Section 5 presents the results, while Section 6 concludes the paper.

2 Data

The dataset employed in this analysis is based on origin-destination (OD) flows derived from mobile phone data. It spans one full year of observations, from September 1st, 2020 to August 31st, 2021, and covers the ACE within the province of Brescia.

The data captures traffic flows, denoted as $flow_{ij,t}$, representing the number of SIM cards moving from ACE i to ACE j during the t -th hour of a day. These OD flows are computed by identifying SIM cards first detected by an antenna in ACE i and then - after at least five minutes - detected by an antenna in ACE j , all within the same one-hour interval. SIM card locations are recorded at fixed 5-minute intervals, namely: [00–05), [05–10), ..., [55–60). Only the first position observed within each 5-minute interval is retained for analysis. For example, consider the one-hour time window t corresponding to 8:00–8:59 AM on January 1st, 2021. Suppose a SIM card is first detected in ACE i between 8:00 and 8:04 AM, then moves to ACE j and is detected there between 8:05 and 8:09 AM, and finally reaches ACE z within the same 5-minute window. In this case, the flow is counted only as $flow_{iz,t}$, and not as $flow_{ij,t}$ nor $flow_{jz,t}$. For each time interval t , the dataset includes a nonsymmetric square matrix of size $N \times N$, where $N = 235$ corresponds to the number of ACEs in the province of Brescia. Each row of the matrix represents the ACE of origin, while each column indicates the ACE of destination. The diagonal elements capture internal flows - i.e., those where origin and destination are the same ACE. Thus, for each ACE i and time t , three types of flows are identifiable: inflows to i , outflows from i , and internal flows from i to i . In total, the database comprises $24 \times 364 = 8,736$ hourly matrices, each of dimension 235×235 , because the data for September 2nd have not been not collected. The data only includes SIM cards connected to the TIM network, including foreign SIMs using roaming. Two types of SIM cards are present: human SIMs (about 85% of the total) and machine-to-machine (M2M) SIMs (around 15%), which are used in automated devices. To avoid double-counting, when a user possesses both a human SIM and one or more M2M SIMs, the analysis is restricted to human SIM cards only.

In this analysis, we restrict the attention to a single ACE, that one of Cellatica (Latitude: 45.584, Longitude: 10.180) belonging to the Mandolossa region, a critical zone with flood episodes located in the northwest outskirts of Brescia. We chose Cellatica for two main reasons: the first is that it constitutes a key node in the road network connecting to/from Brescia; the second reason is to facilitate methodological comparison, as previous studies based on the same mobile phone data have focused on the time series to/from Cellatica (Metulini and Carpita 2024).

In this work, we focus on three hourly time series: $inflow_t$, $outflow_t$, and $internal_t$, where:

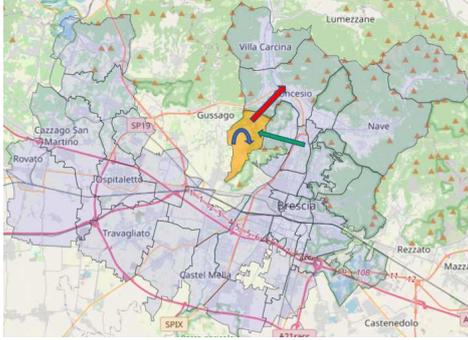
$$inflow_t = \sum_{j=1}^n flow_{ij,t}, \quad i = \text{Cellatica}, \quad (1)$$

$$outflow_t = \sum_{j=1}^n flow_{ji,t}, \quad i = \text{Cellatica}, \quad (2)$$

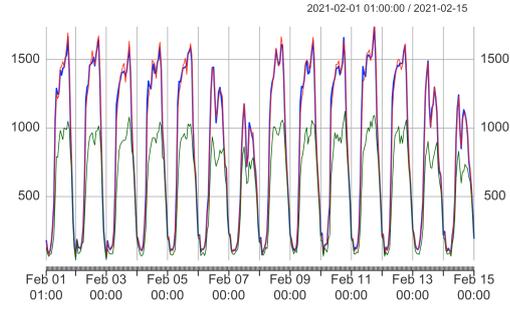
$$internal_t = flow_{ii,t}, \quad i = \text{Cellatica}, \quad (3)$$

and $n = 38$ is the number of neighboring ACEs identified in [Metulini and Carpita \(2024\)](#) as having strong traffic links to Cellatica.

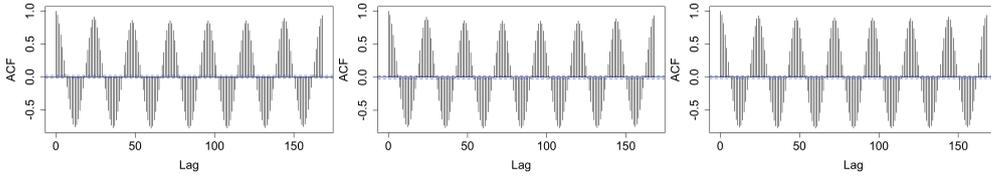
The map of the study area is depicted in Figure 1a. As can be seen in Figure 1b, which shows, as a sample, the time series of the first two weeks of February, the series presents a strong intra-day and intra-weekly pattern.



(a) Map of Cellatica (orange colored), and a stylized representation of internal flows (blue arrow), outflows (red arrow), and inflows (green arrow) to/from the 38 neighbor ACEs (in light blue).



(b) Original data for the first two weeks of February, with Outflows in Blue, Inflows in Red and Internal Flows in green.



(c) Autocorrelation function of Inflow

(d) Autocorrelation function of Outflow

(e) Autocorrelation function of Internal flow

Fig. 1: Map of Cellatica (top left) time series (top right), autocorrelation functions (bottom).

The AutoCorrelation Functions (ACFs) presented in Figure 1 show the pattern of (linear) serial correlation for, from left to right, input, output and internal flow. Both the time series and the ACFs indicates a strong serial dependence that follows a daily path.

3 Methods

Vine-Transform copula (VT-copula) processes are stationary time series processes obtained by using a v-transform and an ARMA process. As shown by McNeil (2021), these models make it easier to capture non-linear serial correlations that traditional autoregressive models cannot identify. Copulas link multivariate density functions to their univariate margins. Hence, they can be applied to study the dependence between the components of a random vector. VT-copulas can be seen as copula models of dependence, used to handle non-Gaussian marginal forms and non-linear and non-Gaussian serial dependence behaviors, specifically used to define a class of tractable non-Gaussian ARMA processes.

Given $Y = \{Y_1, \dots, Y_T\}$ a T-length vector of stationary time series (where Y can be either *inflow*, *outflow* or *internal* flows, and $T=8712$), the v-transform describes the relationship between the quantiles of Y and the quantiles of the distribution of a predictable volatility proxy series, a transformed series such as the absolute value of Y , in which volatility is revealed through serial correlation. They are defined by:

$$\mathcal{V}_{\delta, \kappa, \zeta}(u) = \begin{cases} (1-u) - (1-\delta)\Psi\left(\frac{u}{\delta}\right), & u \leq \delta \\ u - \delta\Psi^{-1}\left(\frac{1-u}{1-\delta}\right), & u > \delta \end{cases} \quad (4)$$

where u is a pseudo-copula data, constructed by calculating the standardized ranks of Y , δ is the parameter for the v-transform of order 1, the Cumulative Density Function (CDF) of Y , and Ψ is a continuous and strictly increasing distribution function on $[0,1]$, which we assume $\Psi(x) = \exp(-\kappa(-(\ln x)^\zeta))$ (McNeil 2021). This is used to divide the v-transform in $u \leq \delta$ and $u > \delta$.

In our case, we will use a specific type of v-transform, which is of order 1, setting $\kappa = 1$ and $\zeta = 1$. This corresponds to a uniform CDF for Ψ and which subsumes the symmetric case $\mathcal{V}_{0.5} = |2u - 1|$ (Bladt and McNeil 2022a):

$$\mathcal{V}_\delta(u) = \begin{cases} (\delta - u)/\delta, & u \leq \delta \\ (u - \delta)/(1 - \delta), & u > \delta \end{cases} \quad (5)$$

V-Transforms can be represented as copulas, allowing formulation and estimation of models that combine arbitrary marginal distributions with copula processes for the dynamics of the volatility proxy. According to Bladt and McNeil (2022b), the best configuration of vines for time series is the drawable vine (d-vine), which describes the strict stationarity of a vector, following some translation-invariance principles on the vine structure.

The first step in copula modeling involves creating a pseudo-copula data by applying the standardized rank transformation to the original data, which we defined as X , in order to estimate the parameter δ of the vine transform, and of the autoregressive component $\{\phi_i\}_{i=1}^p$ and the moving average component $\{\theta_i\}_{i=1}^q$ of a classical ARMA

model (with p and q to be chosen), that make up our copula model. This estimation is performed using the maximum likelihood method, which identifies the fulcrums (i.e., the maximum points of the likelihood functions).

Then one creates a copula model with the previously estimated parameters δ , $\{\phi_i\}_{i=1}^p$, and $\{\theta_i\}_{i=1}^q$.

Next, the procedure consists of applying the margin to the copula model. There are various margins that could be applied, like the normal margin, the Student-t margin, and the Laplace margin. In our case we decided to focus on the Student-t margin, because this type of distribution allows us to capture heavier tails and more extreme values, which characterize the data in our application, as will be shown in Section 4.

4 Empirical procedure

Here we describe how original data undergoes data processing and how the methods are applied to processed data.

First of all, since the observations for September 2nd were missing in the original data, we decided to remove the first two days of the data (i.e. September 1st and September 2nd).

Because of the stationary assumption of the copula model, the first step in our application aims at obtaining a stationary series. We do so with a decomposition of the series through Multiple Seasonal-Trend decomposition using LOESS (MSTL) (Bandara et al. 2025), a recently proposed method to decompose a time series into trend, cycle, and multiple seasonal components. This decomposition model was chosen as it enables the separation of the time series into more than one seasonal component, making it particularly appropriate for our context, where the data display both daily and weekly seasonal patterns. This approach uses LOcally Estimated Scatterplot Smoothing (LOESS), a non-parametric method that locally fits data to capture complex trends, to estimate time series components through a multi-stage decomposition process. First, LOESS estimates the trend component. Then, seasonal decomposition techniques identify and model the seasonal patterns of the data. Finally, residuals (also called “irregulars”) are obtained by subtracting the trend and seasonal estimates from the original series, representing unexplained variation. Since our data shows seasonality patterns that are both daily and weekly, as seen in Figure 1b, our MSTL can be specified through an additive decomposition model of the following form:

$$Y_t = T_t + C_t + S_{d,t} + S_{w,t} + I_t \quad t = 1, \dots, n, \quad (6)$$

where Y_t denotes the original time series at time t , with y representing either *inflow*, *outflow*, or *internal* flows, T_y denotes the long-term trend, C_t denotes the business cycle, $S_{d,t}$ denotes the daily seasonality, $S_{w,t}$ denotes the weekly seasonality, and I_t the irregulars, or residual term. The decomposition has been done through the `mstl` function in the R package `forecast` (Hyndman et al. 2008).

The irregular components are proposed as the dependent variable in the VT-copula model. To support this modeling decision, we assess their stationarity properties in Table 1. To test the stationarity of the irregular component I_t , we applied both the Augmented Dickey–Fuller (ADF) and the Kwiatkowski–Phillips–Schmidt–Shin

(KPSS) tests, using a lag order of $k = 24$ due to the strong daily seasonality of the time series.¹ Both tests support the hypothesis that I_t is stationary, as presented in Table 1. Looking at the ADF statistic and the corresponding p-value, we can reject the null hypothesis that the series has a unit root, implying that it is not non-stationary. Moreover, based on the KPSS statistic and its p-value, we fail to reject the null hypothesis that the series is stationary.

Table 1: Results of ADF and KPSS test applied to Outflow, Inflow and Internal Flow, carried out with lag order of $k = 24$.

		Outflow	Inflow	Internal Flow
ADF	Statistic	-16.633	-16.732	-18.576
	p-value	< 0.01*	< 0.01	< 0.01
KPSS	Statistic	0.007	0.007	0.007
	p-value	> 0.1	> 0.1	> 0.1

* It's worth noting that the results of the p-values shown in the table are approximated as a consequence of the construction of the r functions employed.

We then consider the VT-copula models with a v -transform of order 1 and all possible combinations of AR and MA orders in the set $\{0, 1, 2\}$. As already said, these are calculated on the irregulars I_t of the MSTL decomposition because by only considering the irregulars, we can guarantee the stationarity of the data.

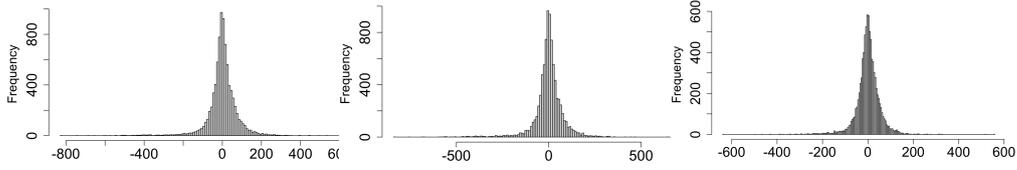
Using this approach, we first fit the vine transform copula with AR and MA components to our pseudo-copula data u . Then, we fit the Student-t marginal distribution to the irregulars found through MSTL. Lastly, we fit the fitted copula and the margins to irregulars. To support the choice of the Student-t margins, we can look to the distribution of the data in Figure 2, which is a normal distribution with heavy tails.

To implement the copula modeling, we apply the semiparametric estimation method by Genest et al. (1995) using the R package `tscopula` (McNeil and Bladt 2024).

To select the best model among all combinations of AR and MA orders, we conducted an analysis based on the Akaike Information Criterion (AIC). However, since the differences in AIC values were not substantial, all models were retained for the subsequent evaluation of forecasting performance.

To assess the forecasting performance of the models, a blocked k -folds cross-validation for time series with a moving window, as depicted in Figure 3, is adopted. This cross-validation strategy is an adaptation of proposed blocked k -folds cross-validation strategies for time series (Snijders 1988; Hyndman and Athanasopoulos 2021), and presents many similarities with that proposed in Metulini and Carpita

¹We also performed the tests with $k = 168$ (corresponding to a full week); the results were similar and are available upon request.



(a) Distribution of the irregulars for Inflows (b) Distribution of the irregulars for Outflows (c) Distribution of the irregulars for Internal flows

Fig. 2: Distributions of the irregulars.

(2024) and Perazzini et al. (2023). According to our strategy, one day of observations



Fig. 3: Illustration of the blocked k -fold cross-validation approach adopted. The yellow time intervals represent the training periods, each with a sample size of 2904, while the green intervals correspond to the validation periods, each of length 24.

is chosen as the validation set and the previous four months of data are taken as the training set. As a result, each validation set has sample size $n_v = 24$ intervals of one hour, and each training set has sample size $n_t = 2902$ (24 observations \times 121 days - 2 hours employed for the autoregressive lag term). In order to evaluate the performance consistently over all the days of the year, we replicate the analysis on different sets of training and validation samples, for a total of 242 different folds to be used for validation². We forecast the flows in the time intervals of the validation set using the VT-ARMA model estimated on the training set.

²Note that, being data available for 363 days, 121 days (from September, 3rd to December 31st) cannot be validated, as they are adopted for training.

To evaluate the accuracy of the predictions coming from the cross-validation strategy, and the explainability of each model’s component, recently proposed Rank-graduation box (RGB) measures (Babaei et al. 2025) are used. RGB measures are statistical tools used to assess and compare the fit of models involving both quantitative as well as ordered or ranked data. These measures evaluate the alignment between observed and predicted rankings. By letting Y be our original data, Y^* and Y^{**} be any two statistical distributions, \hat{Y} be the predicted data and $\hat{Y}^{(-X_k)}$ be the predicted data removing the covariate k , and by letting:

$$RG\cdot = \frac{1}{2} + \frac{1}{2} \frac{\text{cov}(Y^*, \text{rank}(Y^{**}))}{\text{cov}(Y^*, \text{rank}(Y^*))} \quad (7)$$

the general form for a rank-graduation box measure, the RGA is obtained by replacing Y^* with Y and Y^{**} with \hat{Y} , and the RGE by replacing Y^* with \hat{Y} and Y^{**} with $\hat{Y}^{(-X_k)}$ in $1-RG\cdot$. RGA values near 0 indicate poor predictions, values near 1 indicate accurate predictions and a value of 0.5 suggests random predictions. An RGE close to 1 indicates that the k_{th} predictor provides maximum explainability, near 0 means it does not contribute. It is important to note that the forecasted values are not directly

Algorithm 1 Reconstruction of the data using copula model and MSTL Components

- 1: **Input:**
 - 2: *strank_rem* : Standardized ranks of the irregulars
 - 3: *copula_model_rem* : Irregulars obtained from the copula model
 - 4: *seasonal_comp* : Seasonal components (weekly, daily)
 - 5: *trend_comp* : Trend component
 - 6: *rem_comp* : Original irregulars component
 - 7: **for** each element i in *strank_rem* **do**
 - 8: *diff_i* \leftarrow *copula_model_rem*[i] $-$ *strank_rem*[i] \triangleright Compute the difference between copula model and standardized ranks
 - 9: *abs_diff_i* \leftarrow |*diff_i*| \triangleright Compute the absolute difference
 - 10: **end for**
 - 11: *min_diff_idx* \leftarrow index of $\min(\text{abs_diff})$ \triangleright Find index with minimum absolute difference
 - 12: **Retrieve the components for the observation with minimum difference:**
 - 13: *seasonal_comp* \leftarrow extract seasonal components for observation
 - 14: at index *min_diff_idx*
 - 15: *trend_comp* \leftarrow extract trend component for observation
 - 16: at index *min_diff_idx*
 - 17: *rem_comp* \leftarrow extract original remainder component for
 - 18: observation at index *min_diff_idx*
 - 19: **Compute the final result:**
 - 20: *final_result* \leftarrow *rem_comp* + *seasonal_comp* + *trend_comp* \triangleright Combine components to reconstruct the data
 - 21: **Output:** *final_result* \triangleright The reconstructed data
-

comparable to the irregular components themselves, but rather to their standardized ranks. Therefore, we adopt a reconstruction procedure to transform the forecasted values back to a scale comparable with the original data. This step is essential to ensure the proper application of performance evaluation metrics. The reconstruction procedure is detailed in Algorithm 1. This algorithm reconstructs a time series observation by identifying where the copula-generated irregulars most closely match the standardized ranks, then it retrieves the corresponding seasonal, trend, and remainder components, and combines them to produce the final reconstructed data point.

Prediction performance of the VT-ARMA model is compared to that of the VARX-DHR model (Metulini and Carpita 2024; Perazzini et al. 2023), that, for $t = 1, \dots, T$, is defined as follows:

$$\mathbf{Y}_t = \boldsymbol{\nu} + \sum_{h_d=1}^{P_d} \mathbf{A}_{h_d} \mathbf{Y}_{t-24 \times h_d} + \sum_{h_w=1}^{P_w} \mathbf{A}_{h_w} \mathbf{Y}_{t-168 \times h_w} + \mathbf{B} \mathbf{x}_t + \boldsymbol{\epsilon}_t \quad (8)$$

where $\mathbf{Y}_t = [\text{Inflow}_t, \text{Outflow}_t, \text{Internal}_t]'$, $\boldsymbol{\nu}$ is a constant vector of length 3, p_d and p_w are, respectively, the daily and the weekly autoregressive parameters, \mathbf{A}_{h_d} and \mathbf{A}_{h_w} are two 3×3 matrices of coefficients to be estimated, and $\boldsymbol{\epsilon}_t$ is the 3×1 vector of the error terms at time t ; \mathbf{x}_t is the vector of the l exogenous variables at time t , and \mathbf{B} is the $3 \times l$ matrix of coefficients of the exogenous variables, so that $\mathbf{B} \mathbf{x}_t$ is a 3×1 vector.

To capture the seasonality, the r th element ($r = 1, 2, 3$) of the vector $\mathbf{B} \mathbf{x}_t$ is defined as the following combination of daily and weekly functions:

$$\beta_0^{(r)} + \sum_{k_d=1}^{K_d} \left[\alpha_{k_d}^{(r)} s_{k_d}(t) + \gamma_{k_d}^{(r)} c_{k_d}(t) \right] + \sum_{k_w=1}^{K_w} \left[\alpha_{k_w}^{(r)} s_{k_w}(t) + \gamma_{k_w}^{(r)} c_{k_w}(t) \right] \quad (9)$$

$$s_{k_a}(t) = \sin\left(\frac{2\pi k_a t}{m_a}\right), \quad c_{k_a}(t) = \cos\left(\frac{2\pi k_a t}{m_a}\right), \quad a = d, w \quad (10)$$

where β_0 is a constant term, K_d and K_w are the optimal numbers of Fourier bases for the daily and the weekly pattern, respectively, α_k and γ_k are regression coefficients to be estimated, $m_w = 24 \times 7 = 168$ is the weekly seasonal period, and $m_d = 24$ is the daily seasonal period.

Specifically, for comparison purposes the chosen model is a VARX($p_d = 3$, $p_w = 4$) with a DHR($K_d = 7$, $K_w = 6$) component, as the one adopted by Perazzini et al. (2023). The list of exogenous variables includes monthly dummies, to control for the possible presence of changes in average levels among months, and weekdays dummies to control for the possible presence of changes in average levels among weekdays.

Note that the model assumes that the three dependent variables (i.e., inflows, outflows, and internal flows) at time t are related to each other. It is also worth noting that the model is trivariate and so it allows to estimate inflows, outflows and internal flows together, differently from the VT-ARMA, where each flow is estimated in a separate model.

5 Results

In this section, we present the main findings of the study, focusing on the forecasting performance of the VT-ARMA models in comparison to the VARX-DHR benchmark. We begin by reporting, in Section 5.1, the outcomes of the blocked k -fold cross-validation, discussing the results related to accuracy metrics from the approach of Babaei et al. (2025). In this section, we also go more in-depth in evaluating the models' performance by identifying and discussing periods where model performance deteriorates. Subsequently, in Section 5.2 we analyze the contribution of individual components in the copula models through the rank-graduation explainability framework. A diagnostic analysis of model residuals is also provided in Section 5.3, including visual inspections and statistical tests to detect potential issues such as normality and residual autocorrelation.

5.1 Forecasting performance

Table 2 shows the RGA values associated with the copula models and with the VARX-DHR model, computed over the entire forecast period (January 1st-August 31st, 2021). We can see that the predictive performance of the copula models is comparable to that of the VARX-DHR, whereas the best model among the three considered is the VT(1)-ARMA(0,1). This can also be seen in Figure 4.

Table 2: Rank-graduation accuracy for all the copula models considered, and the VARX-DHR calculated on Outflow, Inflow and Internal flow.

Model	Outflow	Inflow	Internal Flow
VT(1)-ARMA(2,2)	0.9758	0.9750	0.9775
VT(1)-ARMA(2,1)	0.9758	0.9750	0.9774
VT(1)-ARMA(1,2)	0.9759	0.9750	0.9776
VT(1)-ARMA(2,0)	0.9758	0.9749	0.9779
VT(1)-ARMA(0,2)	0.9863	0.9854	0.9844
VT(1)-ARMA(1,1)	0.9758	0.9750	0.9777
VT(1)-ARMA(1,0)	0.9768	0.9758	0.9786
VT(1)-ARMA(0,1)	0.9921	0.9919	0.9898
VARX-DHR	0.9805	0.9813	0.9820

It is worth highlighting that copula models are generally simpler and more parsimonious than the VARX-DHR framework. Specifically, in the VT(1)-ARMA(0,1) specification, only a limited number of parameters—namely ϕ , θ , and δ —need to be estimated. In contrast, the VARX-DHR model requires a substantially larger set of parameters due to the inclusion of two autoregressive components of order 3 and 4, along with 6 and 7 Fourier basis terms to capture weekly and daily periodicities, respectively. Furthermore, as previously mentioned, the copula model is univariate, adding to its simplicity.

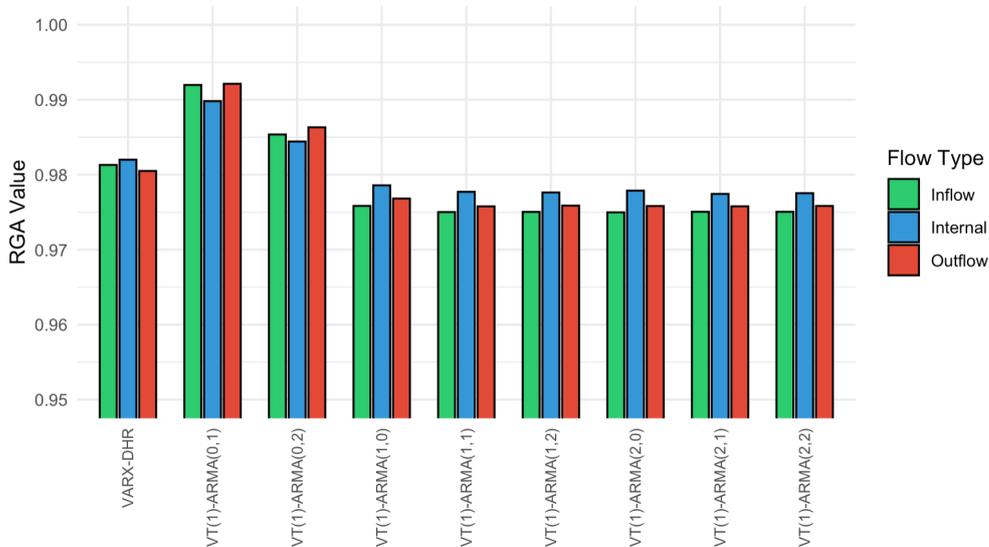


Fig. 4: Results of the rank-graduation accuracy for all the copula models considered, and the VARX-DHR calculated on Outflow, Inflow and Internal flow.

From now on, we will go more in deep with the analysis by considering the VT(1)-ARMA(2,2) and the VT(1)-ARMA(0,1) that are, respectively, the augmented model (i.e., the one with the richer set of parameters) and the one with the highest RGA.

Figure 5a shows the original outflow time series in comparison with the forecasted one for the entire period of validation, which is from January 1st to August 31st, 2021. VT(1)-ARMA(2,2) time series is depicted in green and VT(1)-ARMA(0,1) time series in red. To provide a clearer view of short-term dynamics, Figure 5b just focuses on the month of February. Analogous plots for inflow and internal flow are presented in Figures A2 and A1 in the appendix, with corresponding February zoom-ins.

It is worth noting that the models produce a limited number of negative values.³

A summary of these values is presented in Table 3. It can be observed that the VT(1)-ARMA(0,1) model better performs in terms of limiting the amount of estimated negative values. Nevertheless, the number of negative values is very low in both cases. Note that most of these negative values occur during the night and early morning hours, which are the periods when the amount of traffic is generally the lowest, as can be seen in Figure 1b.

³One reason for the presence of negative values is the incorrect reallocation of components. When the trend and seasonal components are summed with a mismatched irregular component due to a poor alignment in standardized ranks, it can lead to negative values. For instance, if the observation for May 22nd is mistakenly matched with that of January 19th based on their ranks, the combined components may no longer correspond to a coherent time point, potentially resulting in a negative value.

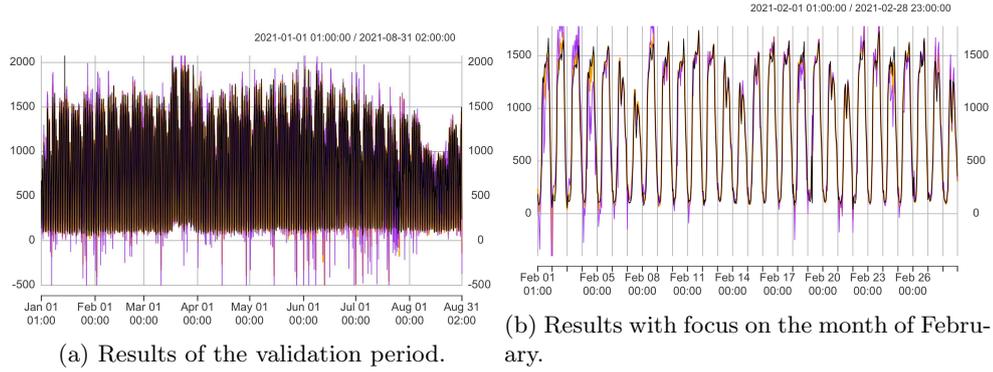


Fig. 5: Results of VT(1)-ARMA(2,2) model (in purple) and VT(1)-ARMA(0,1) (orange) with Student-t, compared to the original outflow data (black), as the whole validation period (left) and a focus on the month of February (right).

Table 3: Estimated negative values on the forecasting period (January 1st-August 31st, 2021) of length 5,808. Presented are both the total negative values and the percentage over the total number of forecasted values.

	VT(1)-ARMA(2,2)			VT(1)-ARMA(0,1)		
	Outflow	Internal Flow	Inflow	Outflow	Internal Flow	Inflow
Negative values	157	125	157	52	53	50
Percentage	2.70%	2.15%	2.70%	0.89%	0.91%	0.86%

While the model has demonstrated strong overall accuracy across the entire forecast period, its performance may vary during specific times of the year. Understanding these variations is important, especially if the model is intended for predictive applications, as it allows users to identify days that may be more challenging to forecast. To investigate this aspect, we computed the rank-graduation accuracy separately for each of the 242 forecast days, using both the VT(1)-ARMA(2,2) and VT(1)-ARMA(0,1) models. The five days with the lowest RGA values, listed from lowest to highest, are January 6th, which is a national holiday (Epiphany), May 14th, March 10th, July 4th (a Sunday), and July 19th. In Figure 6 we display the original time series for January, 6th national holiday (Epiphany) versus the estimated values using VT(1)-ARMA(2,2) and VT(1)-ARMA(0,1). We can observe that for January 6th, especially in the afternoon, the predicted behavior differs from the actual data. The VT(1)-ARMA(2,2) model (green) shows a noticeable drop in the early hours and again during the night, which is not present in the actual data. In contrast, the VT(1)-ARMA(0,1) model (red) tends to overestimate both peaks and troughs, producing forecasts with more pronounced fluctuations than observed in the real series.

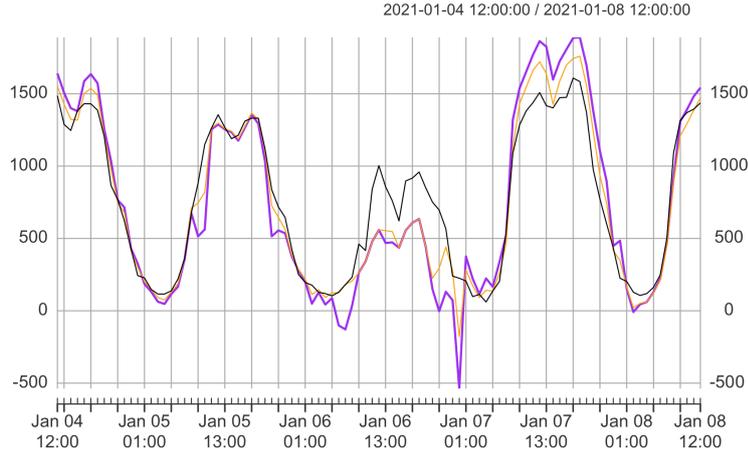


Fig. 6: Representation of the outflows corresponding to the window from midday, January 4th to midday, January 8th, including the national holiday of Epiphany. Original data (black) against forecasted data with VT(1)-ARMA(2,2) model (in purple), and with VT(1)-ARMA(0,1) (orange), with Student-t margins.

5.2 Features explainability

Estimating the explainability of each feature or group of features in statistical models helps identify which components significantly contribute to the model’s predictive power. This generally improves model interpretability, supports feature selection, and guides simplification without lowering performance. So, we assess the contribution of each component of the VT(1)-ARMA(2,2) model using the rank-graduation explainability. We measure the explainability of the copula part, the AR component, and the MA component on the whole validation period. In doing so, we prefer the VT(1)-ARMA(2,2) model, since it incorporates a larger number of terms in each component.

A summary of results is presented in Table 4 and a plot of the results are shown in Figure 7. It is worth noting that the explainability of each component of the model does not necessarily sum up to one. This means that a value such as the one of the copula component, which is averaging around 0.05, should be regarded as high in this context.

We can note that the AR and MA components contribute very little to the model’s explainability, especially the moving average part. This suggests that removing these components individually has minimal impact on the model’s predictive performance. In contrast, the copula component shows significantly higher explainability. This implies that excluding the copula part brings the model’s performance closer to that of a random model, underscoring the critical role of the copula component, surpassing even the importance of the ARMA elements.

Table 4: Rank-graduation explainability for each component of the VT(1)-ARMA(2,2) model: Copula (VT), the autoregressive component (AR) and the moving average component (MA). Outflow, Inflow and Internal flow.

Component	Outflow	Inflow	Internal Flow
VT	0.0530	0.0540	0.0520
AR	0.0040	0.0039	0.0026
MA	0.0001	0.0001	0.0001

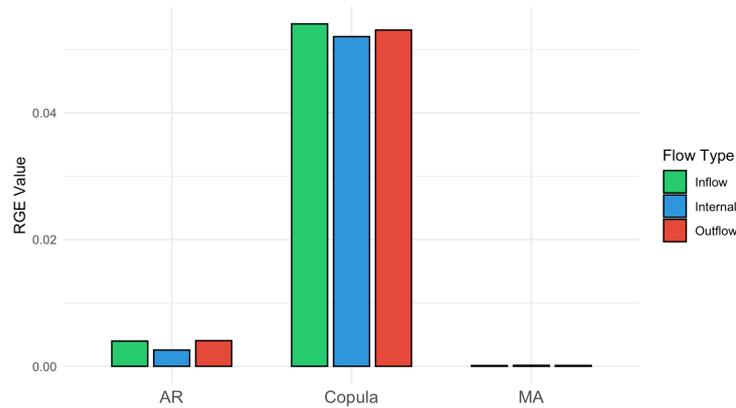


Fig. 7: Results of rank-graduation explainability for each component of the VT(1)-ARMA(2,2) model: Copula (VT), the autoregressive component (AR), and the moving average component (MA).

5.3 Residuals' diagnostic

To assess the adequacy of our VT-copula model, we perform residual diagnostics aimed at verifying two key assumptions: the normality of residuals with zero mean, and the absence of significant serial autocorrelation. These checks are essential to ensure that the residuals behave as random noise, indicating a well-specified model. Normality is evaluated through statistical tests and visual inspections, while autocorrelation is assessed using ACF and partial ACF (PACF). The residual diagnostics for outflows are shown in Figure 8, while similar results for internal flows and inflows are shown and discussed, respectively, in Figure B1 and Figure B2 in the Appendix. Looking at the ACF and PACF shown in Figure 8a and 8b, we can see that there is a lingering autoregressive dynamic, more accentuated for the VT-ARMA(2,2) model. However,

this is significantly reduced with respect to the autocorrelation function of the original data, shown in Figure 1. Looking at the partial autocorrelation function in Figure 8c and 8d, there is still a limited partial autocorrelation, however we can see that it is under the significance level for most of the observations. In the QQ-plots shown in Figure 8e and Figure 8f, the distribution of the residuals is basically normal, with a discrepancy in the tails, which is more accentuated in the VT-ARMA(2,2) model. The same thing is then shown in the histograms of the residuals which show a normal behavior, with heavy tails.

For the model to be well specified, we examined whether the extremely high (low) estimated residuals correspond to extremely high (low) predicted values. In general, this correspondence does not occur. According to the VT-ARMA(0,1) model, the five hourly periods with the highest estimated residuals are: 7 a.m. on February 19th, 7 p.m. on May 14th, 7 p.m. on July 20th, 5 p.m. on July 26th, and 10 a.m. on July 28th. Conversely, the five hourly periods with the lowest estimated residuals correspond to: 3 p.m. on January 24th, 4 p.m. on April 14th, 1 a.m. on June 13th, 8 p.m. on June 30th, and 1 a.m. on July 1st. These periods do not strictly coincide with the hourly intervals corresponding to the highest (lowest) estimated flow values.

6 Conclusions and future developments

The scope of this paper was to develop and evaluate a statistical framework that is both flexible and robust to forecast urban traffic flows, using mobile phone data. Specifically, we applied the Vine-Transform Autoregressive Moving Average (VT-ARMA) copula model to a case study of traffic flows recorded in the municipality of Cellatica, in the province of Brescia. Models that are accurate, interpretable, and parsimonious are needed more and more to capture the complex temporal and non-linear dependencies observed in mobility data. This need is tied to contexts in which mobility planning is crucial for emergency response policies linked to environmental issues.

The methodological strategy proposed consists in incorporating ARMA components into a vine-transformed copula modeling by combining copula theory and traditional time series modeling. To ensure stationarity, the model was applied to the irregulars of the time series obtained decomposing the original data through the MSTL process and isolating the trend and seasonalities of each observation. Out-of-sample forecasting was carried out using a blocked k -fold cross-validation with a moving window approach, and model performance was evaluated in terms of accuracy and explainability with the rank-graduation box framework. The model was compared with the VARX-DHR benchmark model, and it was revealed that the VT-ARMA copula model achieved comparable and in some cases superior predictive accuracy, even by being simpler and requiring fewer parameters. Explainability analysis revealed that the AR and MA components had lower predictive value while the copula component played a central role in capturing traffic flow dynamics, particularly by modeling nonlinearities and tail dependencies often overlooked by traditional models.

The presented findings carry various implications for policies. Reliable traffic forecasting plays a crucial role in smarter mobility planning, like adaptive traffic control and investments in infrastructure. Moreover, the ability to model events of high and

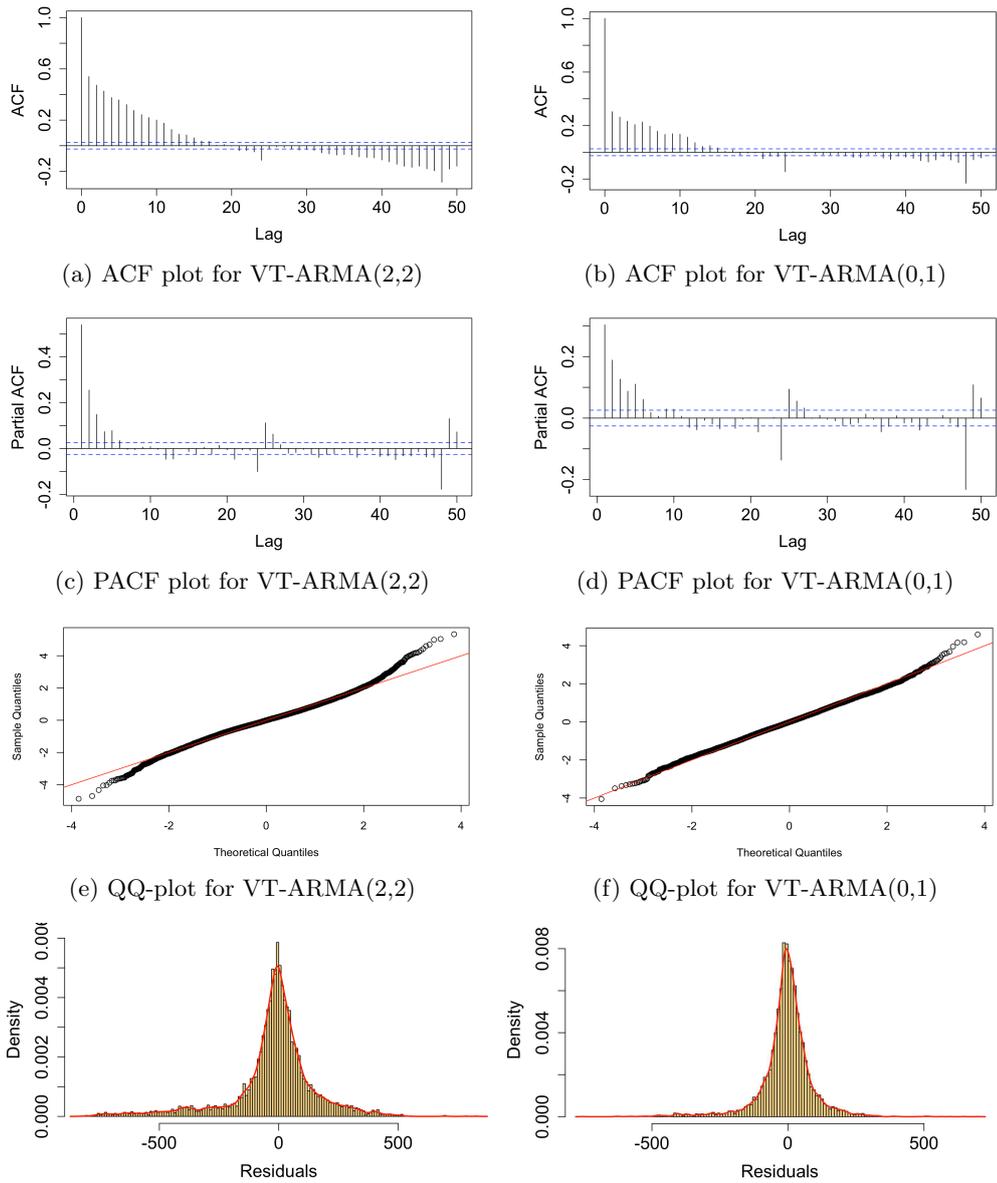


Fig. 8: Residual diagnostics for VT-ARMA(2,2) model (left) and VT-ARMA(0,1) model (right), calculated on outflow data.

low traffic helps in planning the responsiveness of emergency services, crucial in flood-prone or congested urban areas, like Cellatica. Finally, better traffic forecasting can aid in the creation of policies aimed at reducing emissions by redirecting traffic in cases of traffic congestion, or promoting alternative transportation during predicted traffic peaks. In conclusion, the VT-ARMA copula approach transpires as a promising tool for urban traffic forecasting, by being able to balance statistical rigor and simplicity with interpretability.

Future work may extend this methodological framework by evaluating dynamic copulas in the multivariate modeling of interconnected areas. By considering multivariate copula-GARCH models, such as the Dynamic Conditional Correlation GARCH (DCC-GARCH) or the Baba-Engle-Kraft-Kroner GARCH (BEKK-GARCH) with copulas, it becomes possible to capture both time-varying volatility and nonlinear dependencies between locations. This would allow for a more nuanced representation of traffic volatility, which has thus far been overlooked in the current analysis. Another important aspect that warrants further investigation concerns the heavy tails observed in the residuals' distributions. Despite the current model, some extreme residual values persist and may be more appropriately captured by incorporating extreme value distributions, such as the Generalized Extreme Value (GEV) or the Generalized Pareto Distribution (GPD).

Funding

This contribution has been developed for the project "Study of mobile phone siGNals for the evalUation of the interconnections between Mobility and the environment in Lombardia (SIGNUM)" CUP: F53D23010910001- PRIN 2022 PNRR M4C2 - financed by the European Union - Next Generation EU (DD MUR n. 1409 del 14/09/2022).

Ethics declarations

Conflict of interest: all authors declare that they have no conflicts of interest.

References

- Alam, I., Farid, D.M., Rossetti, R.J.: The prediction of traffic flow with regression analysis. In: *Emerging Technologies in Data Mining and Information Security: Proceedings of IEMIS 2018, Volume 2*, pp. 661–671 (2019). https://doi.org/10.1007/978-981-13-1498-8_58 . Springer
- Ahdika, A., Primandari, A.H., Adlin, F.N.: Considering the temporal interdependence of human mobility and covid-19 concerning indonesia's large-scale social distancing policies. *Quality & Quantity* **57**(3), 2791–2810 (2023) <https://doi.org/10.1007/s11135-022-01497-4>
- Babaei, G., Giudici, P., Raffinetti, E.: A rank graduation box for safe ai. *Expert Systems with Applications* **259**, 125239 (2025) <https://doi.org/10.1016/j.eswa.2024.125239>

- Bandara, K., Hyndman, R.J., Bergmeir, C.: Mstl: A seasonal-trend decomposition algorithm for time series with multiple seasonal patterns. *International Journal of Operational Research* (2025) <https://doi.org/10.48550/arXiv.2107.13462>
- Bladt, M., McNeil, A.J.: Time series copula models using d-vines and v-transforms. *Econometrics and Statistics* **24**, 27–48 (2022) <https://doi.org/10.1016/j.ecosta.2021.07.004>
- Bladt, M., McNeil, A.J.: Time series models with infinite-order partial copula dependence. *Dependence Modeling* **10**(1), 87–107 (2022) <https://doi.org/10.1515/demo-2022-0105>
- Balistrocchi, M., Metulini, R., Carpita, M., Ranzi, R.: Dynamic maps of human exposure to floods based on mobile phone data. *Natural Hazards and Earth System Sciences* **20**(12), 3485–3500 (2020) <https://doi.org/10.5194/nhess-20-3485-2020>
- Burzacchi, A., Urbano, V.M., Arena, M., Azzone, G., Secchi, P., Vantini, S.: Spatio-temporal analysis of public transportation undercrowding: Leveraging apc data for a comprehensive evaluation of usage rates. *arXiv preprint arXiv:2410.12618* (2024) <https://doi.org/10.48550/arXiv.2410.12618>
- Clemente, G.P., Della Corte, F., Zappa, D.: Hierarchical spatial network models for road accident risk assessment. *Annals of Operations Research*, 1–36 (2024) <https://doi.org/10.1007/s10479-024-06049-7>
- Carpita, M., De Luca, G., Metulini, R., Zuccolotto, P.: Traffic flows time series in a flood-prone area: modeling and clustering on extreme values with a spatial constraint. *Stochastic Environmental Research and Risk Assessment* **38**(8), 3109–3125 (2024) <https://doi.org/10.1007/s00477-024-02735-x>
- Curci, F., Kërçuku, A., Zanfi, F., Novak, C., *et al.*: Permanent and seasonal human presence in the coastal settlements of lecce. an analysis using mobile phone tracking data. *TEMA* **2**, 57–71 (2022) <https://doi.org/10.6093/1970-9870/8914>
- Cervellera, C., Maccio, D., Reborá, F.: Copula-based scenario generation for urban traffic models. *Expert Systems with Applications* **210**, 118389 (2022) <https://doi.org/10.1016/j.eswa.2022.118389>
- Carpita, M., Simonetto, A.: Big data to monitor big social events: Analysing the mobile phone signals in the brescia smart city. *Electronic Journal of Applied Statistical Analysis: Decision Support Systems and Services Evaluation* **5**(1), 31–41 (2014) <https://doi.org/10.1285/i2037-3627v5n1p31>
- Fernández-Ares, A., Mora, A., Arenas, M.G., García-Sánchez, P., Romero, G., Rivas, V., Castillo, P.A., Merelo, J.: Studying real traffic and mobility scenarios for a smart city using a new monitoring and tracking system. *Future Generation Computer Systems* **76**, 163–179 (2017) <https://doi.org/10.1016/j.future.2016.11.021>

- Fang, Z., Zhu, S., Fu, X., Liu, F., Huang, H., Tang, J.: Multivariate analysis of traffic flow using copula-based model at an isolated road intersection. *Physica A: Statistical Mechanics and its Applications* **599**, 127431 (2022) <https://doi.org/10.1016/j.physa.2022.127431>
- Guardabascio, B., Brogi, F., Benassi, F.: Measuring human mobility in times of trouble: an investigation of the mobility of european populations during covid-19 using big data. *Quality & Quantity* **58**(6), 5181–5199 (2024) <https://doi.org/10.1007/s11135-023-01678-9>
- Genest, C., Ghoudi, K., Rivest, L.P.: A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* **82**(3), 543–552 (1995) <https://doi.org/10.2307/2337532> . Accessed 2024-12-16
- Giorgini, B., Sartori, M.: Human mobility world lines on urban topologies. *Quality & Quantity* **50**(4), 1817–1831 (2016) <https://doi.org/10.1007/s11135-015-0237-6>
- Hyndman, R.J., Athanasopoulos, G.: *Forecasting: Principles and Practice*, 3rd Edition. OTexts (OTexts.com/fpp3), Melbourne, Australia. Accessed on 20/03/2025 (2021)
- Hyndman, R., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O’Hara-Wild, M., Petropoulos, F., Razbash, S., Wang, E., Yasmeeen, F.: *forecast: Forecasting Functions for Time Series and Linear Models*. (2008). R package version 8.23.0. <https://pkg.robjhyndman.com/forecast/>
- Kan, Z., Tang, L., Kwan, M.-P., Ren, C., Liu, D., Li, Q.: Traffic congestion analysis at the turn level using taxis’ gps trajectory data. *Computers, Environment and Urban Systems* **74**, 229–243 (2019) <https://doi.org/10.1016/j.compenvurbsys.2018.11.007>
- Lu, H., Sun, D., Hao, J.: Random traffic flow simulation of heavy vehicles based on r-vine copula model and improved latin hypercube sampling method. *Sensors* **23**(5), 2795 (2023) <https://doi.org/10.3390/s23052795>
- McNeil, A.J., Bladt, M.: *Tscopula: Time Series Copula Models*. (2024). R package version 0.3.9. <https://cran.r-project.org/web/packages/tscopula/index.html>
- Metulini, R., Carpita, M.: A spatio-temporal indicator for city users based on mobile phone signals and administrative data. *Social Indicators Research* **156**(2), 761–781 (2021) <https://doi.org/10.1007/s11205-020-02355-2>
- Metulini, R., Carpita, M.: Modeling and forecasting traffic flows with mobile phone big data in flooding risk areas to support a data-driven decision making. *Annals of Operations Research* **342**, 1629–1654 (2024) <https://doi.org/10.1007/s10479-023-05195-8>
- McNeil, A.J.: Modelling volatile time series with v-transforms and copulas. *Risks* **9** (2021) <https://doi.org/10.3390/risks9010014>

- Mariotti, I., Giavarini, V., Rossi, F., Akhavan, M.: Exploring the “15-minute city” and near working in milan using mobile phone data. *TeMA-Journal of Land Use, Mobility and Environment* **2**, 39–56 (2022) <https://doi.org/10.6093/1970-9870/9309>
- Manfredini, F., Lanza, G., Curci, F., *et al.*: Mobile phone traffic data for territorial research. opportunities and challenges for urban sensing and territorial fragilities analysis. *TEMA* **2**, 9–23 (2022) <https://doi.org/10.6093/1970-9870/8892>
- Nagy, A.M., Simon, V.: Improving traffic prediction using congestion propagation patterns in smart cities. *Advanced Engineering Informatics* **50**, 101343 (2021) <https://doi.org/10.1016/j.aei.2021.101343>
- Perazzini, S., Metulini, R.: Exploring urban mobility patterns in lombardia through advanced analysis of mobile phone data. In: *Scientific Meeting of the Italian Statistical Society*, pp. 422–427 (2024). https://doi.org/10.1007/978-3-031-64350-7_71. Springer
- Perazzini, S., Metulini, R., Carpita, M.: Integration of flows and signals data from mobile phone network for statistical analyses of traffic in a flooding risk area. *Socio-Economic Planning Sciences* **90**, 101747 (2023) <https://doi.org/10.1016/j.seps.2023.101747>
- Po, L., Rollo, F., Bachechi, C., Corni, A.: From sensors data to urban traffic flow analysis. In: *2019 IEEE International Smart Cities Conference (ISC2)*, pp. 478–485 (2019). <https://doi.org/10.1109/ISC246665.2019.9071639>. IEEE
- Snijders, T.A.: On cross-validation for predictor evaluation in time series. In: *On Model Uncertainty and Its Statistical Implications* vol. 307, pp. 56–69. Springer, Berlin Heidelberg (1988). https://doi.org/10.1007/978-3-642-61564-1_4
- Shahriari, S., Sisson, S.A., Rashidi, T.: Copula arma-garch modelling of spatially and temporally correlated time series data for transportation planning use. *Transportation Research Part C: Emerging Technologies* **146**, 103969 (2023) <https://doi.org/10.1016/j.trc.2022.103969>
- Tao, X., Cheng, L., Zhang, R., Chan, W., Chao, H., Qin, J.: Towards green innovation in smart cities: Leveraging traffic flow prediction with machine learning algorithms for sustainable transportation systems. *Sustainability* **16**(1), 251 (2023) <https://doi.org/10.3390/su16010251>
- Triebe, O., Hewamalage, H., Pilyugina, P., Laptev, N., Bergmeir, C., Rajagopal, R.: Neuralprophet: Explainable forecasting at scale. arXiv preprint arXiv:2111.15397 (2021) <https://doi.org/10.48550/arXiv.2111.15397>
- Taylor, S.J., Letham, B.: Forecasting at scale. *The American Statistician* **72**(1), 37–45 (2018) <https://doi.org/10.1080/00031305.2017.1380080>

- Tettamanti, T., Varga, I.: Mobile phone location area based traffic flow estimation in urban road traffic. *Advances in Civil and Environmental Engineering* **1**(1), 1–15 (2014)
- Urban planning, design and management approaches to building resilience – an evidence review: first report on protecting environments and health by building urban resilience. Technical report, Copenhagen: WHO Regional Office for Europe (2022)
- Yuan, J., Zheng, Y., Xie, X.: Discovering regions of different functions in a city using human mobility and pois. In: *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 186–194 (2012). <https://doi.org/10.1145/2339530.2339561>
- Zhang, J., Zheng, Y., Qi, D.: Deep spatio-temporal residual networks for citywide crowd flows prediction. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31 (1) (2017). <https://doi.org/10.48550/arXiv.1610.00081>

Appendix A Forecasting results for inflows and internal flows

In this appendix we report the results of the chosen models VT-ARMA(2,2) and VT-ARMA(0,1) in terms of forecasting performance for internal flows (Figure A1) and inflows (Figure A2). Results show that, similar to the results found by using data regarding outflows, as shown in Figures 5a and 5b, the VT-ARMA(0,1) outperforms the VT-ARMA(2,2) model, in terms of how the predicted data follows the trends of the original data. Moreover, similarly to outflow data, we can see very few negative predicted values. These results are encouraging because they validate the results that were shown regarding outflow data, meaning that copula models perform well also in these cases.

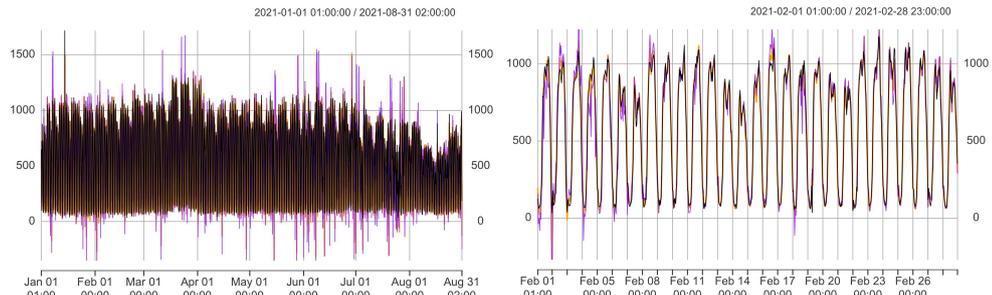


Fig. A1: Results of VT(1)-ARMA(2,2) model (in purple) and VT(1)-ARMA(0,1) (orange) with Student-t, compared to the original internal flow data (black), as the whole validation period (left) and a focus on the month of February (right).

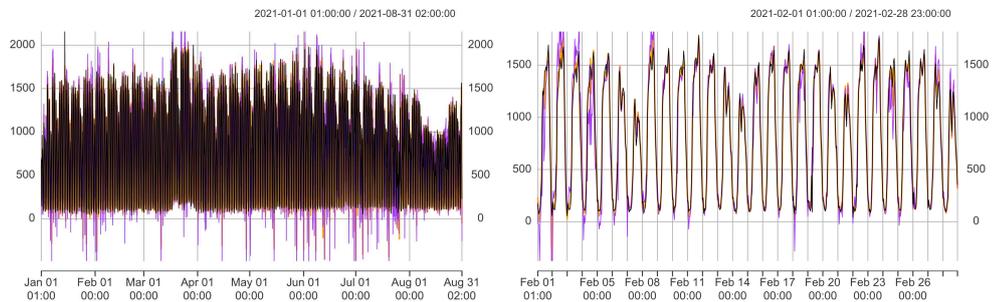
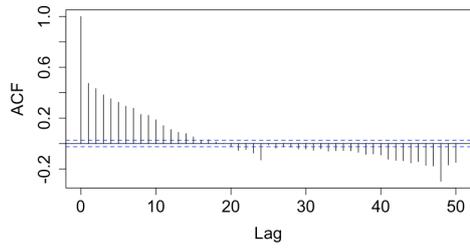


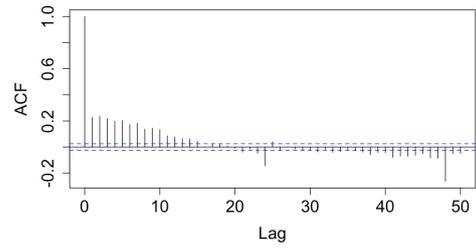
Fig. A2: Results of VT(1)-ARMA(2,2) model (in purple) and VT(1)-ARMA(0,1) (orange) with Student-t, compared to the original inflow data (black), as the whole validation period (left) and a focus on the month of February (right).

Appendix B Residual diagnostics for inflows and internal flows

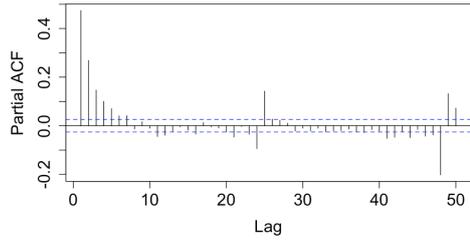
In this appendix we report the diagnostic on model's residuals for inflows (in Figure B2) and internal flows (in Figure B1), similar to what was presented in Figure 8 for the case of outflows. Also in these cases, we can underline how the residuals related to the VT-ARMA(0,1) model follow a trend which behaves more normally than the residuals related to the VT-ARMA(2,2) model. Overall, in both inflow and internal flow cases, a moderate degree of residual serial autocorrelation can be observed, consistent with the pattern identified in outflows. The distribution of residuals appears approximately symmetric and normally distributed; however, deviations in the right and left tails are present and warrant further investigation in future research.



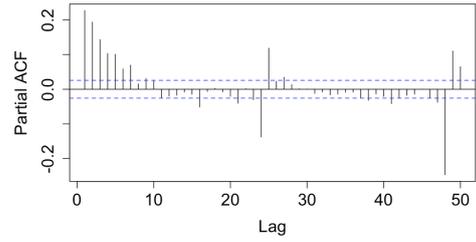
(a) ACF plot for VT-ARMA(2,2)



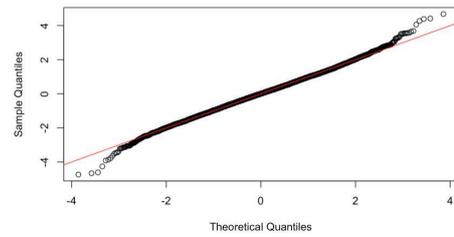
(b) ACF plot for VT-ARMA(0,1)



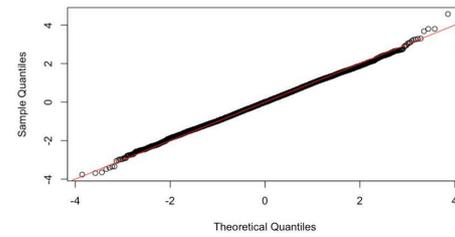
(c) PACF plot for VT-ARMA(2,2)



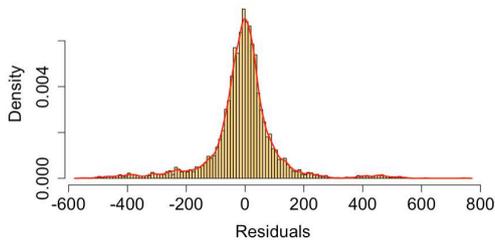
(d) PACF plot for VT-ARMA(0,1)



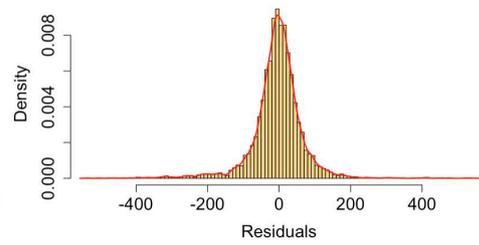
(e) QQ-plot for VT-ARMA(2,2)



(f) QQ-plot for VT-ARMA(0,1)

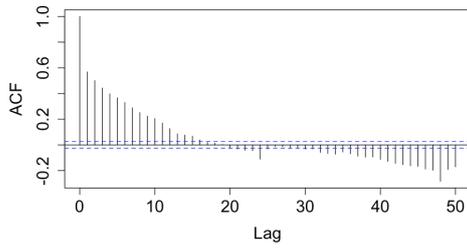


(g) Histogram of residuals for VT-ARMA(2,2)

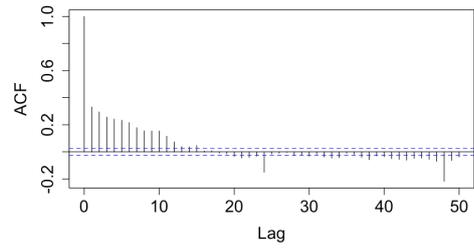


(h) Histogram of residuals for VT-ARMA(0,1)

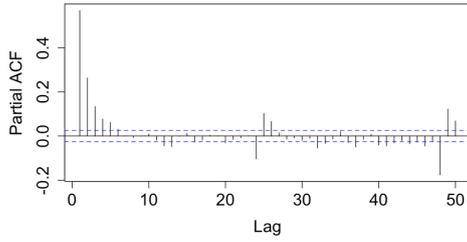
Fig. B1: Residual diagnostics for VT-ARMA(2,2) model (left) and VT-ARMA(0,1) model (right) calculated on internal flow data.



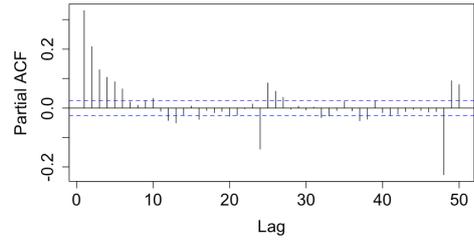
(a) ACF plot for VT-ARMA(2,2)



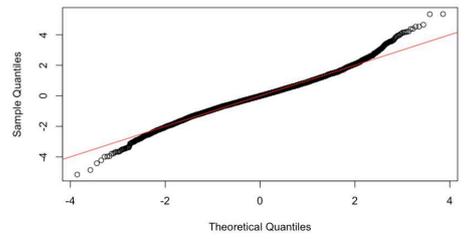
(b) ACF plot for VT-ARMA(0,1)



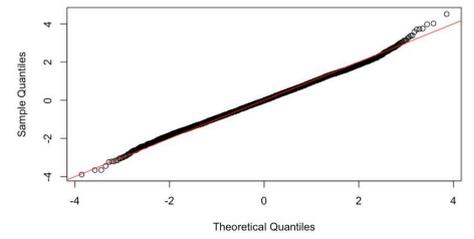
(c) PACF plot for VT-ARMA(2,2)



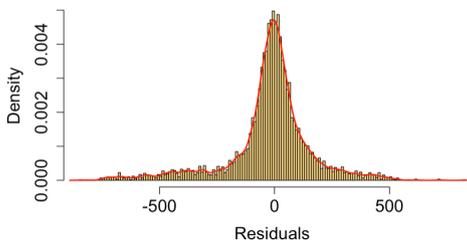
(d) PACF plot for VT-ARMA(0,1)



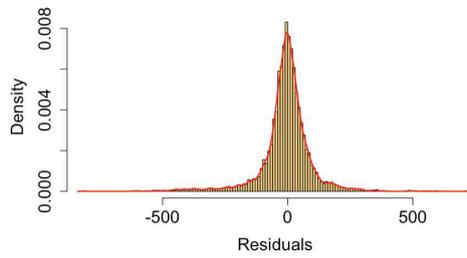
(e) QQ-plot for VT-ARMA(2,2)



(f) QQ-plot for VT-ARMA(0,1)



(g) Histogram of residuals for VT-ARMA(2,2)



(h) Histogram of residuals for VT-ARMA(0,1)

Fig. B2: Residual diagnostics for VT-ARMA(2,2) model (left) and VT-ARMA(0,1) model (right) calculated on inflow data.