ORIGINAL ARTICLE

SVR and ARIMA models as machine learning solutions for solving the latency problem in real-time clock corrections

Mutaz Qafisheh[1,2] · Angel Martín[2] · Raquel M. Capilla[3] · Ana B. Anquela[2]

Abstract

Real-time precise point positioning (PPP) has become a prevalent technique in global navigation satellite systems (GNSS). However, GNSS real-time users must receive space state representation (SSR) products to correct for satellite clock, orbit, and phase biases. The International GNSS Service (IGS) provides GNSS users with real-time services (RTSs) through different real-time correction SSR products. These products arrive at the GNSS users with some latency, which affects the quality of real-time PPP positioning. The autoregressive integrated moving average (ARIMA) and support vector regression (SVR) models are used in this research to predict those corrections to eliminate the latency effect. ARIMA model reduces the standard deviation by 28% and 13% for GPS and GLONASS constellations, respectively, compared to the real-time solution, which includes the latency effect, the research simulated the latency effect and named it a forced-latency solution, and the SVR model reduces the standard deviation by 28% and 23% for GPS and GLONASS constellations, respectively. The results for the permanent GNSS stations used in this study across different years 2013, 2014, 2015, 2019, and 2021 show a mean reduction in the 3D positioning standard deviation by 13% compared with the forced-latency solution for the ARIMA solution and 9% for the SVR solution. The potential of both models to overcome the latency effect is apparent based on the findings.

Keywords Precise point positioning · Real-time positioning · Support vector regression · Autoregressive integrated moving average · Clock corrections

Introduction

The GNSS is widely used to determine positioning through various methods, such as single point positioning, relative positioning, and PPP. The main principles of those methods can be found in Zumberge et al. 1997; Hofmann-Wellenhof et al. 2008; Enge and Misra 2011; Sanz et al. 2013; or Leick et al. 2015. PPP allows GNSS users to locate themselves globally since it does not rely on a local network of GNSS receivers or a base station. This is also free from the local effects resulting from the movement of reference stations.

Nevertheless, users must precisely measure pseudoranges and carrier phases from GNSS satellites to perform post-processing and real-time PPP. GNSS biases must be accurately estimated, such as multipath, ionospheric delay, tropospheric delays, earth tides, relativistic effects, and antenna variation. A detailed study of these effects can be found in Zumberge et al. (1997), Ge et al. (2008), Cai and Gao (2013), Li et al. (2015), and Teunissen and Khodabandeh (2015). Finally, code and phase biases and satellite orbital and clock errors can be mitigated by using adequate biases and orbital and clock correction products (Teunissen and Khodabandeh 2015; Henkel et al. 2018; Ye et al. 2018). The IGS and other Analysis Centers (ACs) provide precise respective products (Dow et al. 2009).

The PPP coordinate accuracy could reach a centimeter or sub-decimeter depending on the operational mode of post-processing or real-time (RT-PPP).

The International GNSS Services (IGS) initiated a pilot project for real-time activities at the beginning of 2001 to improve RT-PPP accuracy (Dow et al. 2009). The goal was to provide the RTS for GNSS users with accurate real-time

✉ Angel Martín
aemartin@upvnet.upv.es

1   College of Engineering, Civil and Architectural Department, Paletine Polytechnic University, PO Box 198, Hebron City, Palestine, Israel

2   Department of Cartographic Engineering, Geodesy, and Photogrammetry, Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain

3   Cartographic Institute of Valencia, 46018 Valencia, Spain

products, mainly orbital and clock corrections, to reduce the satellite ephemeris and the onboard satellite clock errors.

Consequently, some ACs using the GNSS permanent reference network are involved in the real-time tracking, computation, and broadcasting of real-time products (Grinter and Roberts 2013). Various ACs produce and disseminate their real-time products. IGS combines those products to produce official real-time products such as IGS01, IGS02, and IGS03. IGS01 is a single-epoch combination solution, IGS02 is a Kalman filter combination solution, and IGS03 is an experimental Kalman filter combination for GPS and GLONASS solutions.

The orbit and clock products are used by the user with different latencies, i.e., the sum of the time required to generate the products in the ACs, to combine these products for IGS RTS Service (Johnston et al. 2017), broadcast them over the Internet using Networked Transport of RTCM via Internet Protocol (NTRIP) as RTCM state-space representation (SSR) correction streams, and the time required for the local computer where the RT-PPP solution is implemented. As the latency value increases, the corrections become outdated, leading to applying older corrections to real-time observations. Martín et al. (2015b) found a 3D error of approximately 0.15 m and 0.30 m if latencies of 30 and 40 s are employed, respectively. The latency problem is also addressed by Hadas and Bosy (2014). This study showed a negative relationship between latency values and the corrections accuracies. The latency values for the individual products produced by ACs reach 10–12 s. However, this value increases to 30–40 s for the combined IGS products since IGS needs extra time to receive all AC solutions and combine them.

Different prediction models are discussed in the literature. Examples are quadratic polynomial or linear models with sinusoid terms found in Huang et al. (2014), El-Mowafy et al. (2017), El-Mowafy (2019a, b), Yang et al. (2019) to predict orbital and clock corrections during data communication failure or discontinuity periods, and genetic algorithms with autoregressive moving average models to predict 15 min of corrections during data loss of the IGS02 stream in Kim and Kim (2017). A limited number of studies have investigated prediction models for navigation satellite systems other than GPS. In Qafisheh et al. (2020), the GLONASS satellite system is included in the prediction. Still, both Qafisheh et al. (2020) and Kim and Kim (2017) studies are missing the latency effects on coordinate accuracy.

Clock correction values are the most challenging to model compared to orbital corrections because they are highly correlated to the onboard GNSS clock oscillator behavior, which suffers from frequency instability, jumps, offsets, outliers, and frequency drift. Additional research about the characteristics model of the clocks and their behaviors can be found in Daly (1990), Senior et al. (2008), Hauschild et al. (2013), and Maciuk (2019). Different real-time schemes are utilized to adapt clock offsets, jumps, noise, and speed using the Kalman filter (Huang and Zhang 2012).

This research aims to overcome the latency in the IGS03 real-time clock product using the ARIMA and the SVR models over a short period. The prediction model and the methodology used in this research can be extended to cover other combined real-time products made available by IGS or other ACs. Various open-source software can handle SSR correction streams such as RTKLIB (Takasu 2009, http://www.rtklib.com/rtklib_tutorial.htm), GNSSSurfer (SAPOS®-Berlin 2020, http://217.9.43.196/Download/), or BNC from the Kartographie und Geodäsie Agency (BKG) (Weber and Mervart 2007). In this research, BNC software is employed. It can be used to accomplish different GNSS tasks, including satellite coordinate comparison, broadcast, combine and upload corrections, PPP post-processing, RT-PPP, and decode the SSR messages to obtain the required correction values to eliminate orbit and clock errors (Kouba and Héroux 2001; Weber and Mervart 2007; Weber and Mervart 2007). The PPP technique uses ionospheric-free combinations to mitigate the ionospheric effects by combining pseudoranges codes and carrier observations of different frequencies. The tropospheric error, antenna phase center, and other biases are well modeled in the BNC software.

Experimental data

First, the RT-PPP measurements from December 13–16, 2019 were stored. We used the IGS03 correction product, a 5-s sampling interval, and the Brest (France) permanent station observations. Before experimenting, the BNC software stored observation, navigation, and correction files in a local computer synchronized to the Internet time.

Latency values were also stored. The mean value of the latency was 31.68 s. However, the range during the implementation varied between 31.34 and 32.21 s. The stored correction file contains data for 52 observed satellites with approximately 26 thousand correction values. Several satellites are affected by data unavailability during some periods.

The research extended to cover more stations over several years to evaluate the prediction models properly. Consequently, one day of RT-PPP data was used, and nine stations distributed globally were selected. We investigated the following years: 2013, 2014, 2015, 2019, and 2021. Table 1 shows the percentages of the null value for various satellite blocks for different years, and Fig. 1 shows the distribution of the selected IGS stations.

Table 1 Clock correction availability for different satellite blocks

| Year	Satellite block	Percentage of clocks corrections availability
2013	GPS-IIF	98.13
	GPS-IIR	99.97
	GPS-IIRM	99.97
	GLONASS-K	100
	GLONASS-M	96.88
2014	GPS-IIF	96.07
	GPS-IIR	99.90
	GPS-IIRM	99.96
	GLONASS-K	100.00
	GLONASS-M	98.00
2015	GPS-IIF	99.07
	GPS-IIR	99.96
	GPS-IIRM	99.93
	GLONASS-K	98.77
	GLONASS-M	99.62
2019	GPS-IIF	98.34
	GPS-IIR	99.97
	GPS-IIRM	100.00
	GLONASS-K	89.88
	GLONASS-M	99.16
2021	GPS-IIF	99.46
	GPS-IIR	91.48
	GPS-IIRM	99.89
	GLONASS-K	76.11
	GLONASS-M	96.05
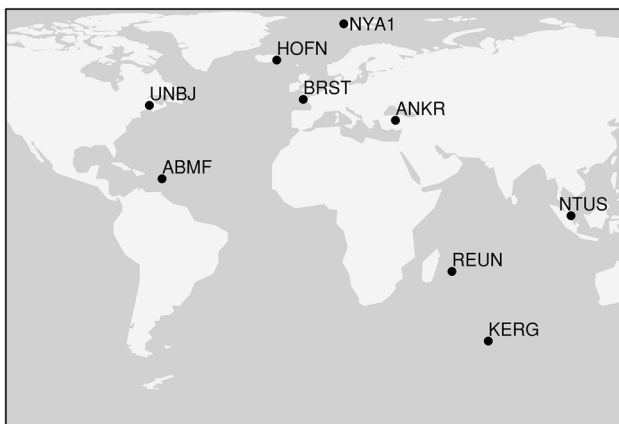
Fig. 1 Distribution of the selected IGS permanent stations (source: Esri, Here, Garmin, FAO, USGS, © OpenStreetMap contributors, and the GIS User Community)

Methodology

One of the most important aspects to take into account in any machine learning or data mining project is the proper choice of the method or mathematical model to use. In this case, it is about making future predictions about the studied signal itself, so the choice of the method will depend on the characteristics of the signal. Based on the results obtained in the signal analysis, there will be a greater foundation in the choice of the prediction method.

Signal analysis

Statistical analysis for stationary inspection of the signal is needed to select the proper machine learning algorithms. Our goal is to predict future values in a temporal series (time series forecasting). If the signal is stationary, that is, having no trends, seasonality, or cyclic patterns, most of the machine learning methods, including random forest, neural networks, or XGB, fail to fit the data. In those cases, the predictions have the same values as the last observation (Bownlee 2017).

Augmented Dickey–Fuller, Phillips–Perron, and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) statistical tests can examine a time-series stationarity property (Kwiatkowski et al. 1992). It is possible to improve the stationarity decision by combining different statistical tests (Schlitzer 1995). However, the KPSS test is applied to the stored GNSS clock correction data to simplify the code and reduce processing time. The stationarity was tested for various time windows of clock data split from the IGS03 clock correction files obtained between December 13-and 16, 2019. Table 2 presents the stationarity results for the 8-min data windows, including the combined results for different satellite blocks and the percentage of stationary and not-stationarity windows. In this research, we conclude that the clock corrections are mostly stationary signals, but these results could

Table 2 Percentage of rolling sliding windows (RSWs) with stationarity and non-stationarity behavior for different satellite blocks

Satellite block	Percentage of RSWs with stationary behavior	Percentage of RSWs with non-stationary behavior
GPS-IIF	76.99	23.01
GPS-IIR	88.28	11.72
GPS-IIRM	89.64	10.36
GLONASS-K	62.11	37.89
GLONASS-M	61.02	38.98

vary with respect to clock oscillators or signal length, considerably restricting the machine learning model for forecasting. Therefore, SVR and ARIMA models are the best candidates (Hyndman and Athanasopoulos 2018, Bownlee 2017). Furthermore, they can also be used to predict non-stationary signals.

Prediction models

Once the prediction models have been chosen based on the signal analysis, we are going to describe them briefly, adjusting the explanation to what is really necessary to understand the implementation and results properly.

Support vector regression model

The support vector machine (SVM) and derived SVR methods have been widely used in machine learning due to their simplicity (Clarkson et al. 2012). SVM development allows applying the traditional support vector classifier in a higher-dimensional space (Drucker et al. 1997). The adaptation of higher-dimensional space can be made by harnessing the mapping function (kernel). Radial-base function, polynomial, linear, and other kernels can transform data into higher-dimensional space.

Given a data set containing $\{(x_i, y_i, i = 1, 2, \ldots, m)\}$ where x_i and $y_i \in R$, x_i are the features and y_i the label, both normalized to $(+1, -1)$, the classifier categorizes data according to its label. For data categorization, the hyperplane must be established to separate the data. Due to the nature of the data, many hyperplanes can classify the same dataset. The chosen hyperplane must maximize the margin, defined using the nearest points near the margin, called support vectors. The margin can be classified as a hard or soft margin. The so-called *soft SVM* is generally used because it allows isolated values and misclassified samples, making it more realistic.

Smola and Schölkopf (2004) defined the SVM formula as:

$$\text{minimize } w, b, \xi = \frac{1}{2} w^T w + C \sum_{i=1}^{m} \xi_i \qquad (1)$$

Subjected to

$$y_i \left(w^T x_i + b \right) \geq 1 - \xi_i, \quad \xi_i \geq 0, \quad \text{for } 1 \leq i \leq m \qquad (2)$$

where w represents the margin width, b the bias, ξ denotes the slack variable, which is associated with the defined soft SVM, allowing some value to fall in the margin, and C is the trade-off margin width.

ARIMA model

The ARIMA model has been extensively used to forecast time series (Sneeuw et al. 2012; Ye et al. 2012; Moreira et al. 2013; Xin et al. 2018; Van Le and Nishio 2019). The ARIMA is a combination of autoregressive (AR), moving average (MA), and the differencing term (I). The ARIMA model depends on the well-known Box–Jenkins methodology (Box et al. 2011; Hyndman and Athanasopoulos 2018).

The I-term differentiates the time series to ensure the stationarity, and the AR indicates that it is a regression of the variable against itself. For example, an autoregressive model of order p can be defined as:

$$y_t = a + b_1 y_{t-1} + b_2 y_{t-2} + \cdots + b_p y_{t-p} + \varepsilon_t \qquad (3)$$

where a is a constant, ε_t is the noise, and b_p are the parameters to be identified. This expression is similar to a multiple regression but considers the lagged values of y_t as a predictor.

The MA suggests a moving average model that uses past forecast errors in a regression model. For example, a model average model or order q can be expressed by:

$$y_t = c + \varepsilon_t + r_1 \varepsilon_{t-1} + r_2 \varepsilon_{t-2} + \cdots + r_q \varepsilon_{t-q} \qquad (4)$$

where c is a constant, ε_t is the noise, and r_p are the parameters to be determined. In this case, the prediction y_t can be computed based on a weighted moving average of the past q forecast errors.

Finally, the integration of AR, I, and MA generates the ARIMA model described as:

$$\begin{aligned} y'_t = {} & m + b_1 y'_{t-1} + b_2 y'_{t-2} + \cdots + b_p y'_{t-p} \\ & + r_1 \varepsilon_{t-1} + r_2 \varepsilon_{t-2} + \cdots + r_q \varepsilon_{t-q} + \varepsilon_t \end{aligned} \qquad (5)$$

where m is a constant, ε_t is the noise, and y'_t signifies that the series is differenced, and it may have been differenced by more than one.

The ARIMA model is usually denoted as ARIMA (p, d, q) where the p, d, and q represent the number of lags applied to the MA model, the differencing degree, and the order of the MA model, respectively (Piccolo 1990; Box et al. 2011; Hyndman and Athanasopoulos 2018).

Implementation

The stored RT-PPP correction data contain orbital and clock corrections. The prediction is centered on the clock corrections since orbital corrections can be well projected based on the radial, along-track, and cross-track components and their

velocities broadcast in the navigation messages (Hadas and Bosy 2015). Clock corrections also need to be estimated at a high rate (Ge et al. 2012) due to the frequency instability of onboard GNSS clocks, mainly caused by temperature and gravitational variations.

A balance between the training set data length and computational time is required to control the prediction accuracy and process time if real-time predictions are required. Rolling sliding window (RSW), expanding window, or fixed splitting are various testing and training methods to accomplish this. The clock corrections experience a period of high oscillations, with periodic jumps due to the clocks' frequency drift, instability, and changes in gravitational forces. Consequently, the forward validation and the RSW technique is more suitable testing and training method. It guarantees computational time reduction and mitigates the effect of outdated data, which clearly emerges by using an expanding window. Using a rolling window instead of an expanding window allows GNSS users to keep the computational load at the minimum level and improve the prediction accuracy.

This research aims to predict clock corrections in approximately 30 s to overcome the latency effect, so three essential aspects must be addressed. First, we must fix the impact of choosing different RSWs with various lengths on both SVR and ARIMA model predictions in the prediction accuracy. Second, the model parameterization in the prediction model definition must be updated by searching and adjusting the optimal hyperparameters in the model definition based on the stored data. The question here is to search for the best rate of updating hyperparameters and then search for the best values for these hyperparameters. This led to evaluating the prediction accuracy with a different combination of C and *gamma* hyperparameters in the SVR model and (p,q,d) in the ARIMA model, considering a large hyperparameter search space in both models to ensure more stable predictions.

Regarding SVR, a large C value means that the model did not allow errors to violate the margin, resulting in margin shrinking; the smaller it is, the more isolated values are allowed in the margin area (soft SVM). The *gamma* parameter establishes the variance of the Gaussian function used for class separation; it must be tuned to control interpolation, extrapolation, and nonlinearly separable classes (Guyon et al. 1993). For the ARIMA model, p controls the number of lags required to implement the linear regression, d controls the degree of differencing mandatory to ensure signal stationarity, and q is devoted to controlling signal error propagation. Finally, the third essential aspect is to compare, in a post-process experiment, the clock values stored in real-time with the SVR and ARIMA predictions, computing the standard deviation and the range for residuals and the effect on the final PPP computed coordinates.

This experiment was performed by reprocessing the stored RT-PPP files using BNC software in the static mode by eliminating the latency effect since the clock and orbital corrections production and transmission are eliminated. So, the latency effect was eliminated to obtain the so-called latency-free solution.

The stored clock correction file was modified to hold new clock correction values from SVR and ARIMA predictions. BNC is rerun again with the files containing predictions to compare the coordinates between the stored files in post-processing (latency-free solution) and SVR and ARIMA predictions, solutions with the latency eliminated by prediction. A simulation of the latency effect is computed to complete the comparisons by shifting the stored correction file ahead by 30 s, which more or less represents the stored mean latency, the so-called forced-latency solution, and the normal RT-PPP situation. Figures 2 and 3 show an overview of the two central parts of the methodology applied in this research. Figure 2 includes the methodology part related to real-time implementation. The BNC software used the broadcasted corrections and navigation information with the station observation stream to produce the real-time coordinates solution. The clock corrections predictions were obtained from both models according to the rolling sliding values. The clock corrections residuals were calculated from different solutions with respect to real-time corrections.

The post-processing was described in Fig. 3. The stored navigation and station observation information with the stored clock corrections predictions and clock corrections stored in real-time and the simulated clock corrections were used to produce the station coordinates with respect to the different solutions or clock corrections. The sampling value for real-time and forced-latency clock corrections files was 10 s. SVR and ARIMA models predict the clock corrections with the same sampling interval.

Results and discussion

The initial portion of the results, following the implementation section, is related to the RSW length effect. Different sliding windows are probed with sizes of 1, 2, 4, 8, and 16 min to fit the SVR and ARIMA models. One hour of clock correction data was implemented for this experiment on December 13, 2019. The standard deviation for clock correction residuals obtained by subtracting the SVR and ARIMA predictions was compared with the latency-free solution. Based on the results, an 8-min-long RSW for the ARIMA model and a 1-min RSW for the SVR model long can be selected as a good compromise between prediction accuracy and processing time. Tables 3 and 4 include

Fig. 2 Overview of the RT-PPP prediction flow and implementation
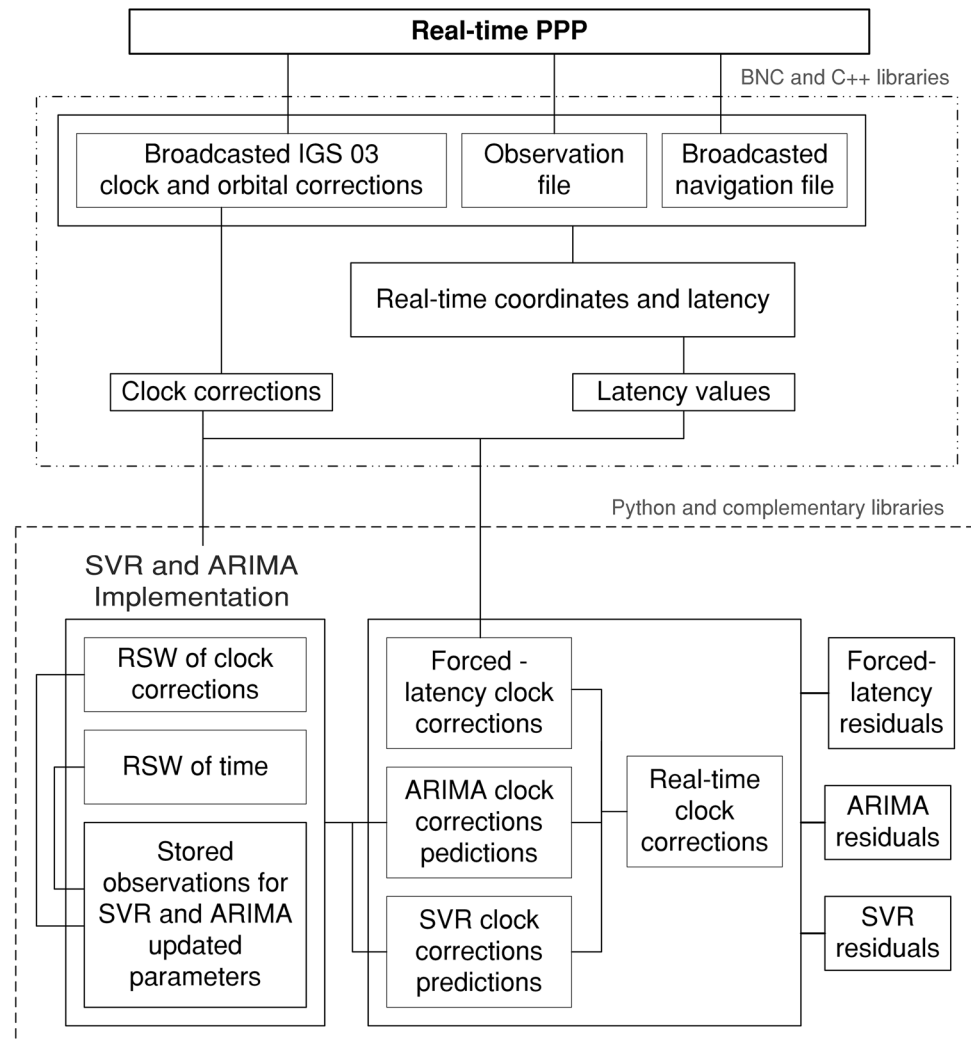
the standard deviation comparison for both models. One satellite represents each satellite block, and the required time for processing is also included. Those tables also contain, for comparison purposes, the standard deviation of the residual between the forced-latency solution, as a reproduction of the real-time process, and the latency-free solution. It should be mentioned that the results are the same if we choose different one-hour intervals of the stored data. It can be seen from the tables that there is a negative relation between the RSW length and the error values obtained by the ARIMA model. This shows that the ARIMA model is highly dependent on the length of previous observations to adjust the hyperparameters properly and construct the prediction model. STDP in Tables 3 and 4 denote the standard deviation of predictions in comparison with the free-latency solution, and STDL denotes the forced-latency standard deviation in comparison with the free-latency solutions. For both tables, the standard deviation unit is in meters, and the processing time corresponds to a one-hour of predictions.

The SVR model can be implemented through different kernels such as linear, polynomial, and radial base functions. The reason behind picking the radial base function as a kernel is that it is more suitable for short-term predictions like latency. In this research, the radial base function was used, and as the RSW length increased, the influence of outdated observations became more notable. It is worth mentioning that the different types of clocks used in different satellite blocks that led to negative and positive trends are not valid for all RSWs.

According to the implementation section, the second vital position to consider is related to selecting the best updating rate for hyperparameters. Updating rates of 0.25, 0.50, 1, 2, 3, 4, 5, and 6 h were examined from December 13 to December 16, 2019. Based on the experiments, a one-hour updating rate for SVR is selected. The list of hyperparameters to resolve in the SVR method should be quite extensive in order to ensure that correct values are chosen. Consequently, a one-hour rate is a good balance between

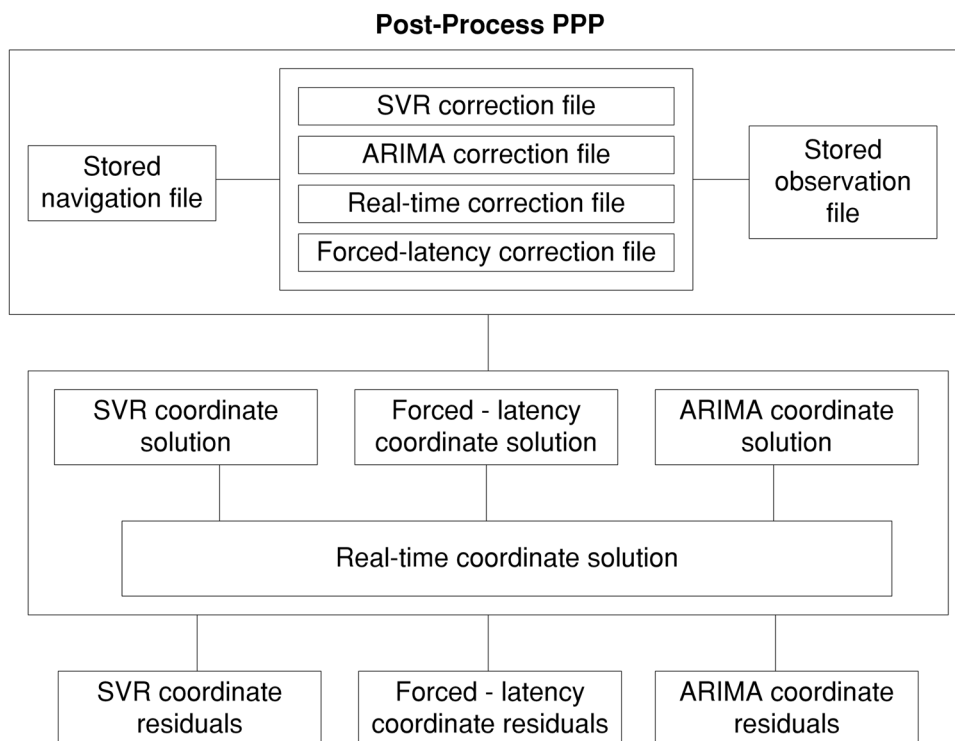**Fig. 3** Post-processing PPP flow and implementation used to apply the clock correction predictions

Post-Process PPP

Stored navigation file	SVR correction file	Stored observation file
	ARIMA correction file	
	Real-time correction file	
	Forced-latency correction file	

SVR coordinate solution	Forced - latency coordinate solution	ARIMA coordinate solution

Real-time coordinate solution

SVR coordinate residuals	Forced - latency coordinate residuals	ARIMA coordinate residuals

Table 3 Comparison of the standard deviation of the ARIMA prediction using various RSW lengths with the no-latency solution

Satellite PRN/block number	Sliding window (minutes)				
	1	2	4	8	16
G32/IIF					
STDP	0.015	0.016	0.016	0.009	0.012
STDL	0.016	0.016	0.016	0.016	0.015
Processing time	2m56s	4m35s	4m40s	5m26s	6m16s
G05/IIRM					
STDP	0.040	0.036	0.033	0.024	0.024
STDL	0.040	0.040	0.040	0.041	0.041
Processing time	3m20s	6m45s	7m02s	7m46s	10m59s
G23/IIR					
STDP	0.016	0.015	0.013	0.010	0.007
STDL	0.017	0.017	0.017	0.017	0.017
Processing time	3m34s	6m13s	6m52s	7m04s	8m39s
R09/K					
STDP	0.026	0.026	0.019	0.016	0.015
STDL	0.028	0.028	0.028	0.030	0.029
Processing time	2m45s	4m18s	3m54s	4m13s	5m13s
R15/M					
STDP	0.014	0.016	0.012	0.022	0.019
STDL	0.016	0.016	0.016	0.028	0.027
Processing time	2m41s	4m59s	5m02s	4m53s	4m58s

Table 4 Comparison of the standard deviation of the SVR prediction using various RSW lengths with the no-latency solution

Satellite PRN/block number	Sliding window (minutes)				
	1	2	4	8	16
G32/IIF					
STDP	0.011	0.016	0.022	0.029	0.034
STDL	0.016	0.016	0.016	0.016	0.015
Processing time	20 s	20 s	21 s	24 s	26 s
G05/IIRM					
STDP	0.027	0.033	0.039	0.047	0.051
STDL	0.040	0.040	0.040	0.041	0.041
Processing time	20 s	20 s	21 s	24 s	25 s
G23/IIR					
STDP	0.012	0.018	0.025	0.028	0.031
STDL	0.017	0.017	0.017	0.017	0.017
Processing time	20 s	20 s	15 s	23 s	26 s
R09/K					
STDP	0.018	0.025	0.031	0.041	0.058
STDL	0.028	0.028	0.028	0.030	0.029
Processing time	20 s	20 s	21 s	23 s	25 s
R15/M					
STDP	0.010	0.015	0.023	0.041	0.051
STDL	0.016	0.016	0.016	0.028	0.027
Processing time	20 s	23 s	20 s	23 s	26 s

Fig. 4 IGS03 range differences in meters for GPS satellites

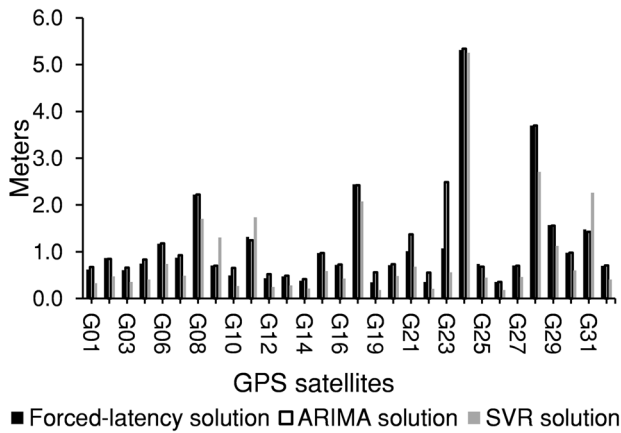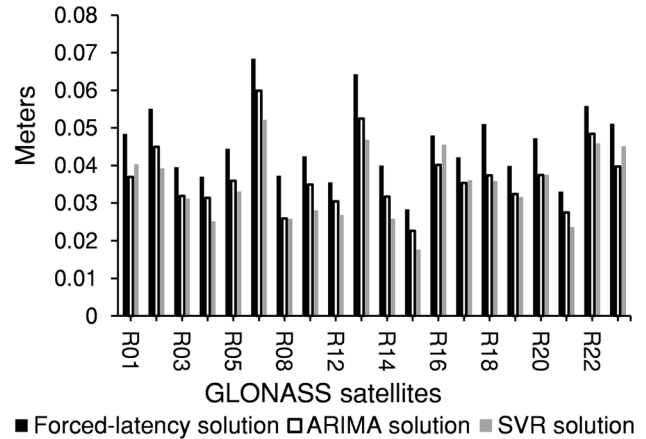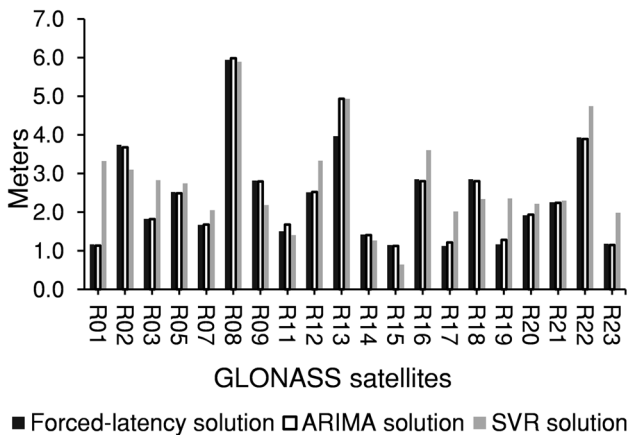**Fig. 5** IGS03 range differences in meters for GLONASS satellites

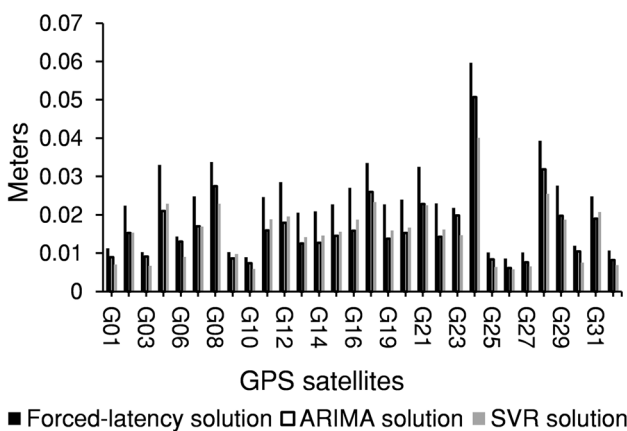**Fig. 6** IGS03 standard deviation in meters for GPS satellites

Fig. 7 IGS03 standard deviation in meters for GLONASS satellites

computation time and accuracy. However, the ARIMA model hyperparameters list can be considered fairly small: 0 to 3 for p and d parameters and 0 to 1 in the q parameter. Thus, the hyperparameter search is a high-speed operation, and their fixation can be executed for every RSW of 8 min without loss of computational time. It has also been realized that the rolling sliding window length has a greater impact on the prediction accuracy than the parameter updating rate. The same data and products for the ARIMA model examination were employed to analyze the SVR model. The third and final aspect to consider is related to the comparisons between prediction models and the forced-latency solution regarding the latency-free clock correction values. Figures 4 and 5 demonstrate the range calculated for the clock correction residuals. The residuals were calculated by subtracting the clock correction values of forced-latency, ARIMA prediction, and SVR prediction clock files with respect to the latency-free file. Figures 6 and 7 illustrate the standard deviation for the same residuals. It is apparent from those figures that the range reduction is negligible due to clock correction jumps. However, a remarkable reduction in standard deviation, especially for GPS satellites, has been discovered. For Figs. 4, 5, 6, and 7, the differences are computed with respect to the latency-free clock correction values.

From the previous experiments, it can be concluded that the ARIMA model has lower standard deviations than the SVR model, except for the GLONASS M satellite block, where SVR is superior. However, the SVR model shows a faster execution time than the ARIMA model, and Tables 3 and 4 refer to the processing time needed for one-hour predictions.

Table 5 Average of the mean, the standard deviation, and range of the 3D coordinate residual for all stations

Solution	Average of all years	2013	2014	2015	2019	2021
Forced-latency solution (mean)	0.23	0.04	0.07	0.51	0.40	0.14
ARIMA model (mean)	0.19	0.04	0.04	0.22	0.42	0.21
SVR model (mean)	0.20	0.06	0.04	0.40	0.41	0.10
Forced-latency solution (standard deviation)	0.84	0.60	0.47	1.26	0.76	1.11
ARIMA model (standard deviation)	0.73	0.66	0.44	0.75	0.70	1.12
SVR model (standard deviation)	0.58	0.58	0.38	1.43	0.69	0.75
Forced-latency solution (range)	11.52	13.74	7.74	7.79	11.01	17.33
ARIMA model (range)	11.90	11.68	8.92	4.71	10.94	23.27
SVR model (range)	10.25	11.95	8.20	6.75	11.34	13.20
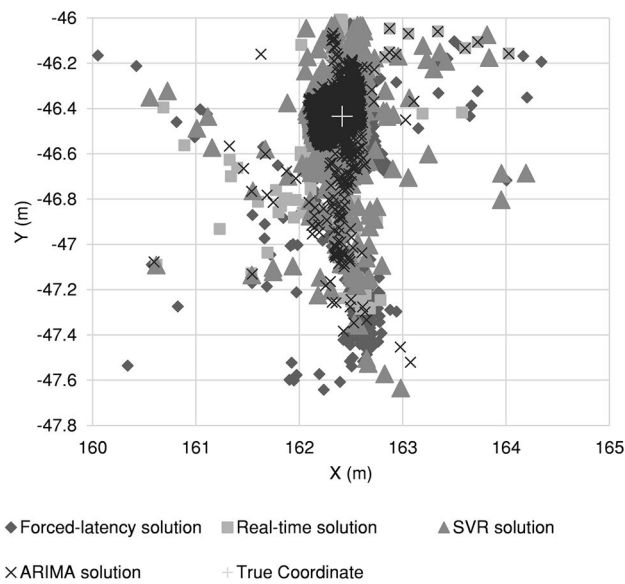
Fig. 8 BREST station coordinates for all solutions in the ITRF 2014 reference frame. December 13, 2019

Coordinates free from latency (latency-free solution), simulated real-time (forced-latency solution), SVR predictions, and ARIMA predictions were acquired using the postprocess BNC software module for the permanent stations of Fig. 1, and one day of real-time data for the years 2013, 2014, 2015, 2019, and 2021. Table 5 shows the average of the 3D residuals statistical summary in terms of the mean, standard deviation, and range of the forced-latency, SVR, and ARIMA solutions compared with the latency-free solution. Figure 8 is a scatter plot that expresses the X and Y planar coordinates of all four solutions and includes the fixed coordinates of station IGS BREST during one day period. Both of the prediction models show remarkable improvement in the accuracy of coordinates in comparison with the forced-latency solution, which simulates the real-time situation. Figure 8 confirms that the SVR and ARIMA coordinate solutions are more precise and much denser around the true coordinate than the forced-latency solution, especially the ARIMA solution. For visualization purposes, X and Y coordinates were shifted with (423,100 and 332,700) meters.

Table 6 The average standard deviation of the clock corrections was obtained from three solutions with respect to five satellite blocks

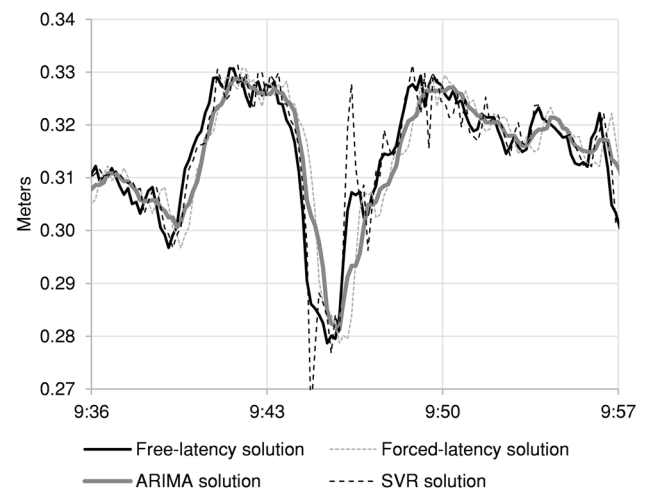| Solution	Satellite block	2013	2014	2015	2019	2021
Forced-latency solution	GPS-IIF	0.39	0.04	0.02	0.03	0.01
	GPS-IIR	0.11	0.03	0.03	0.02	0.02
	GPS-IIRM	0.12	0.03	0.03	0.03	0.02
	GLONASS-K	0.04	0.07	–	0.14	0.04
	GLONASS-M	0.06	0.06	0.06	0.05	0.05
ARIMA model	GPS-IIF	0.34	0.03	0.01	0.02	0.01
	GPS-IIR	0.09	0.02	0.02	0.02	0.01
	GPS-IIRM	0.10	0.02	0.02	0.02	0.01
	GLONASS-K	0.04	0.07	–	0.13	0.03
	GLONASS-M	0.05	0.05	0.05	0.04	0.04
SVR model	GPS-IIF	0.32	0.03	0.01	0.02	0.01
	GPS-IIR	0.09	0.02	0.02	0.02	0.01
	GPS-IIRM	0.11	0.02	0.02	0.02	0.01
	GLONASS-K	0.03	0.06	–	0.09	0.03
	GLONASS-M	0.05	0.05	0.04	0.04	0.04

Table 7 The average of clock corrections ranges obtained from three solutions with respect to five satellite blocks

Solution	Satellite Block	2013	2014	2015	2019	2021
Forced-latency solution	GPS-IIF	18.79	1.54	0.51	1.06	0.38
	GPS-IIR	4.41	0.84	0.77	0.68	0.40
	GPS-IIRM	4.55	0.87	0.82	1.25	0.37
	GLONASS-K	1.11	2.57	–	9.72	0.78
	GLONASS-M	2.49	2.54	2.14	2.52	1.99
ARIMA model	GPS-IIF	13.80	1.59	0.53	0.90	0.40
	GPS-IIR	4.03	0.84	0.74	0.71	0.41
	GPS-IIRM	4.13	0.88	0.88	1.21	0.47
	GLONASS-K	1.11	2.57	–	9.03	0.74
	GLONASS-M	2.70	2.25	2.05	2.49	2.06
SVR model	GPS-IIF	18.59	1.81	0.43	0.92	0.23
	GPS-IIR	4.07	0.76	0.69	0.38	0.26
	GPS-IIRM	4.36	0.68	0.98	0.91	0.20
	GLONASS-K	0.82	3.95	–	10.17	0.46
	GLONASS-M	3.03	3.04	2.74	3.07	2.46
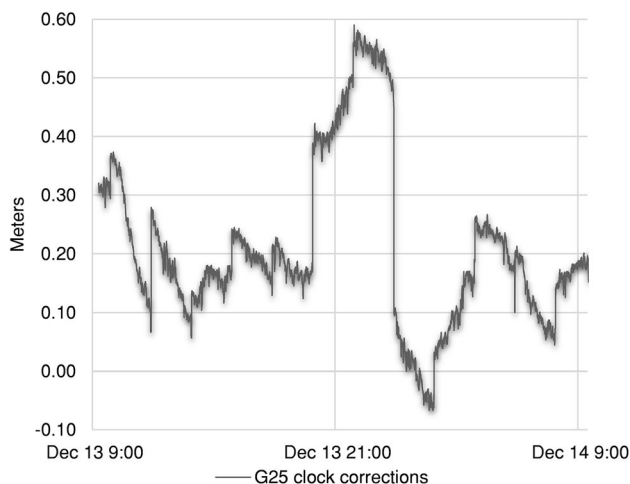
Fig. 9 GPS satellites (PRN: G25) clock corrections values. December 13 and 14, 2019

Fig. 10 GPS satellites (PRN: G25) clock corrections and prediction models values

Tables 6 and 7 show the statistical summary throughout the research days of the years into consideration of standard deviation and range for the residuals obtained by subtracting the clock corrections values from different solutions (forced-latency, ARIMA, and SVR predictions) with respect to latency-free values. Tables 6 and 7 does not include the mean value analysis as the mean values for the residuals are near zero for all solutions.

Finally, the data collected with the highest clock correction jumps were examined. Each clock correction value was subtracted from the previous values so that the maximum clock correction jumps could be determined; consequently, the results were grouped for each satellite block by calculating the mean of the maximum values. Figure 9 shows the

clock corrections values for 24 h between December 13 and 14, 2019. The figure shows the value for corrections concerning the GPS satellite G05. It can be seen from the figure the clock correction values dropped around 0.40 m within seconds. Figure 10 represents 30 min of clock corrections, the black line represents the free-latency solution, and the gray dash line denotes the forced-latency solution, while the thick gray and black dash lines show the prediction models. Table 8 presents the mean value of the maximum consecutive difference between successive clock correction values. The results proved that the jumps could reach 11.68 m for satellites in GPS block GPS-IIF meters in 2013. Table 8 shows that the clock corrections consecutive jumps have fewer values in recent years, especially for GPS satellites.

Table 8 Mean of the maximum differences for the clock corrections for different satellite blocks (units in meters)

Year	Satellite block	Mean of maximum consecutive differences
2013	GPS-IIF	11.68
	GPS-IIR	2.16
	GPS-IIRM	2.15
	GLONASS-K	0.50
	GLONASS-M	1.27
2014	GPS-IIF	0.75
	GPS-IIR	0.26
	GPS-IIRM	0.48
	GLONASS-K	1.25
	GLONASS-M	1.04
2015	GPS-IIF	0.27
	GPS-IIR	0.37
	GPS-IIRM	0.29
	GLONASS-K	0.49
	GLONASS-M	1.41
2019	GPS-IIF	0.11
	GPS-IIR	0.31
	GPS-IIRM	0.19
	GLONASS-K	0.43
	GLONASS-M	0.63
2021	GPS-IIF	0.16
	GPS-IIR	0.16
	GPS-IIRM	0.22
	GLONASS-K	0.41
	GLONASS-M	1.95

From the GNSS real-time user point of view, the implementation should be as follows. If the SVR method is selected, the user should store observations for the first 2 min to generate the first valid SVR model. An initial search for the hyperparameters and fixed model prediction is carried out. Next, the fitted model can predict the next clock correction according to the stored latency value, keeping the size of the rolling sliding windows constant at 1 min. Thus, the new observations are stored, and the old observations are automatically deleted in the SRW window. In addition, the GNSS users need to update the *C* and *gamma* parameters for the SVR model each hour to ensure accuracy. Therefore, the user is required to store one-hour latency values continuously. For the first hour of observations, the hyperparameters can be updated every 15–20 min to ensure accuracy in the prediction.

In case the ARIMA method is selected, the user should store the observation for the first 2 min. The first search for hyperparameters should be executed for the first 2 min of observations. The respective hyperparameters are utilized for the first 8 min of observation. Afterward, an SRW window of 8 min can be employed, including the search-and-fix of hyperparameters. In both models, the GNSS user should use the clock corrections as they are received in the first 2 min with a latency effect; then, the GNSS user can use a prediction model and results.

Conclusions

The SVR and ARIMA prediction models are applied to IGS03 clock corrections. Both models are used to estimate the clock corrections to overcome the latency effect. In this research, the latency effects on orbital corrections were not studied, as they vary at a low rate compared to the clock corrections and are not severely affected by latency. Three days of real-time data were utilized initially in this study in order to obtain the correct time dimension for the rolling sliding window and the correct rate to update the ARIMA and SVR hyperparameters. Similarly, one day of real-time data for the years 2013, 2014, 2015, 2019, and 2021 is used to confirm the validity of the proposed methods. The results prove that both models can be used to overcome the latency.

According to Tables 6 and 7, the ARIMA model reduces the standard deviation by 28% and 13% for GPS and GLONASS constellations, respectively, compared to the forced-latency solution. The SVR model reduces the standard deviation by 28% and 23% for GPS and GLONASS constellations, respectively. Both models showed robust behavior during clock correction jumps because the models rely on 1 min of clock correction observations for the SVR model and 8 min for the ARIMA model. The uses for the RSWs mitigate the effect of jumps and improve the availability of the RT-PPP solution derived from the BNC software and maintain the coordinate convergence during jump periods.

Finally, based on the results of Table 5, the average of the 3D standard deviations, both models are successful in eliminating the latency effect, especially the ARIMA model. However, the SVR model displays superiority in processing time compared to ARIMA. It is approximately eight to nine times faster. Thus, it could be an excellent solution to overcome the latency due to simplicity and computational speed, which is important since GNSS receivers need to predict clock corrections for approximately 10–20 satellites that are above the horizon during the observation when using GPS and GLONASS signals. The proposed prediction models could also forecast clock corrections during periods of data loss or discontinuity. In this research, the range analysis is included in order to investigate outlier predictions. Both prediction models experienced this phenomenon to prevent such behavior in a limited number of predictions. They affect both clock corrections and 3D calculated residuals. To eliminate possible outliers, a threshold detector must be implemented

for both predicting models. Additionally, the research methodology could improve the accuracy of dynamic GNSS receivers such as self-driving vehicles and mobile users. Those users rely on mobile data connections where the internet connections are more affected by the latency.

Data availability The 2014 data were chosen as a sample among the different years researched. The data have been published through the Mendeley data repository through the following link https://data.mendeley.com/datasets/dd5prwnhpv/1 it contains the 2014 observation, navigation, and clock corrections files obtained from the four different solutions. The shared data included a step-by-step readme file explaining results reproduction. Additionally, the shared data have a figures folder containing scatter plots for the researched IGS stations.

References

Asari K, Saito M, Mikami I (2019) Capturing method of millimeter displacement in centimeter class PPP-RTK measured data. In: Proceedings of the ION GNSS 2019, Pacific PNT Meeting, Honolulu, Hawaii, April 2019, pp 367–375

Box GEP, Jenkins GM, Reinsel GC (2011) Time series analysis: forecasting and control, vol 734. Wiley, New York

Brownlee J (2017) Introduction to time series forecasting with python: how to prepare data and develop models to predict the future. In: Brownlee J (ed) Machine learning mastery

Cai C, Gao Y (2013) Modeling and assessment of combined GPS/GLONASS precise point positioning. GPS Solut 17(2):223–236

Clarkson KL, Hazan E, Woodruff DP (2012) Sublinear optimization for machine learning. J ACM. https://doi.org/10.1145/2371656.2371658

Daly IDK (1990) Characterization of NAVSTAR GPS and GLONASS on-board clocks. In: IEEE symposium on position location and navigation. A decade of excellence in the navigation sciences, Las Vegas, NV, USA, March 1990

Dow JM, Neilan RE, Rizos C (2009) The international GNSS service in a changing landscape of global navigation satellite systems. J Geod 83(3–4):191–198

Drucker H, Surges CJC, Kaufman L, Smola A, Vapnik V (1997) Support vector regression machines. Adv Neural Inf Process Syst 1:155–161

El-Mowafy A (2019a) Impact of predicting real-time clock corrections during their outages on precise point positioning. Surv Rev 51(365):183–192

El-Mowafy A (2019b) Predicting real-time orbit and clock corrections for positioning using GPS, GLONASS, and QZSS in natural hazard warning systems. J Appl Geod 13(2):69–79

El-Mowafy A, Deo M, Kubo N (2017) Maintaining real-time precise point positioning during outages of orbit and clock corrections. GPS Solut 21(3):937–947

Enge P, Misra P (2011) Global positioning system: signals, measurements, and performance—revised, 2nd edn. Ganga-Jamuna Press, Nanded

Ge M, Gendt G, Rothacher M, Shi C, Liu J (2008) Resolution of GPS carrier-phase ambiguities in Precise Point Positioning (PPP) with daily observations. J Geod 82(7):389–399. https://doi.org/10.1007/s00190-007-0187-4

Ge M, Chen J, Dousa J, Gendt G, Wickert J (2012) A computationally efficient approach for estimating high-rate satellite clock corrections in realtime. GPS Solut 16(1):9–17. https://doi.org/10.1007/s10291-011-0206-z

Georg Weber G, Mervart L, Stürze A, Stöcker D (2016) BNC Ntrip Client (BNC). A toolkit for retrieving, decoding, converting and processing real-time GNSS data streams. https://software.rtcm-ntrip.org/export/HEAD/ntrip/trunk/BNC/src/bnchelp.html. Accessed Jan 2020

Grayson B, Penna NT, Mills JP, Grant DS (2018) GPS precise point positioning for UAV photogrammetry. Photogramm Rec 33(164):427–447

Grinter T, Roberts C (2013) Real time precise point positioning: are we there yet? In: Proceedings of the international global navigation satellite systems society. IGNSS symposium 2013, Outrigger Gold Coast, Qld, Australia

Guyon I, Boser B, Vapnik V (1993) Automatic capacity tuning of very large VC-dimension classifiers. Adv Neural Inf Process Syst 5:147–155

Hadas T, Bosy J (2014) IGS RTS precise orbits and clocks verification and quality degradation over time. GPS Solut 19(1):93–105. https://doi.org/10.1007/s10291-014-0369-5

Hadas T, Bosy J (2015) IGS RTS precise orbits and clocks verification and quality. GPS Solut 19(1):93–105

Hauschild A, Montenbruck O, Steigenberger P (2013) Short-term analysis of GNSS clocks. GPS Solut 17(3):295–307

Henkel P, Psychas D, Günther C, Hugentobler U (2018) Estimation of satellite position, clock and phase bias corrections. J Geod 92(10):1199–1217

Hofmann-Wellenhof B, Lichtenegger H, Wasle E (2008) GNSS, global navigation satellite systems: GPS, GLONASS, GALILEO & more. Springer, New York

Huang G, Zhang Q (2012) Real-time estimation of satellite clock offset using adaptively robust Kalman filter with classified adaptive factors. GPS Solut 16(4):531–539. https://doi.org/10.1007/s10291-012-0254-z

Huang GW, Zhang Q, Xu GC (2014) Real-time clock offset prediction with an improved model. GPS Solut 18(1):95–104. https://doi.org/10.1007/s10291-013-0313-0

Hyndman RJ, Athanasopoulos G (2018). Forecasting: principles and practice. Online-open-access-textbooks

Johnston G, Riddell A, Hausler G (2017) The international GNSS service. In: Teunissen P, Montenbruck O (eds) Handbook of global navigation satellite systems. Springer, New York, pp 967–982. https://doi.org/10.1007/978-3-319-42928-1_33

Kim M, Kim J (2017) GA-ARMA model for predicting IGS RTS corrections. Int J Aerosp Eng. https://doi.org/10.1155/2017/6316590

Kouba J, Héroux P (2001) Precise point positioning using IGS orbit and clock products. GPS Solut 5(2):12–28. https://doi.org/10.1007/PL00012883

Kwiatkowski D, Phillips PCB, Schmidt P, Shin Y (1992) Testing the null hypothesis of stationarity against the alternative of a unit root. J Econom 54(1–3):159–178

Leick A, Rapoport L, Tatarnikov D (2015) GPS satellite surveying, 4th edn. Wiley, Hoboken

Li X, Ge M, Dai X, Ren X, Fritsche M, Wickert J, Schuh H (2015) Accuracy and reliability of multi-GNSS real-time precise positioning: GPS, GLONASS, BeiDou, and Galileo. J Geod 89(6):607–635

Maciuk K (2019) Satellite clock stability analysis depending on the reference clock type. Arab J Geosci. https://doi.org/10.1007/s12517-018-4069-2

Martín A, Anquela AB, Dimas-Pagés A, Cos-Gayón F (2015a) Validation of performance of real-time kinematic PPP. A possible tool for deformation monitoring. Measurement 69:95–108. https://doi.org/10.1016/j.measurement.2015.03.026

Martín A, Hadas T, Dimas A, Anquela AB, Berné JL (2015b) Influence of real-time products latency on kinematic PPP results. In: 5th international colloquium scientific and fundamental aspects of the galileo programme. Braunschweig, Germany

Nie Z, Liu F, Gao Y (2020) Real-time precise point positioning with a low-cost dual-frequency GNSS device. GPS Solut. https://doi.org/10.1007/s10291-019-0922-3

Piccolo D (1990) A distance measure for classifying ARIMA models. J Time Ser Anal 11(2):153–164. https://doi.org/10.1111/j.1467-9892.1990.tb00048.x

Qafisheh MWA, Martín A, Torres-Sospedra J (2020) Support vector regression machine learning tool to predict GNSS clock corrections in real-time PPP technique. ICL-GNSS 2020 WiP proceedings, June 2–4, 2020, Tampere, Finland. http://ceur-ws.org/Vol-2626/paper11.pdf

Sanz J, Juan JM, Zornoza J Hernández-Pajares M (2013) GNSS data processing. Volume I: fudamentals and algorithms. Ed. ESA communications. https://gssc.esa.int/navipedia/GNSS_Book/ESA_GNSS-Book_TM-23_Vol_I.pdf

Schlitzer G (1995) Testing the stationarity of economic time series: further Monte Carlo evidence. Ric Econ 49(2):125–144

Senior KL, Ray JR, Beard RL (2008) Characterization of periodic variations in the GPS satellite clocks. GPS Solut 12(3):211–225. https://doi.org/10.1007/s10291-008-0089-9

Smola AJ, Schölkopf B (2004) A tutorial on support vector regression. Stat Comput 14(3):199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88

Sneeuw N, Novák P, Crespi M, Sansò F (2012) Analysing time series of GNSS residuals by means of AR (I) MA processes. In: VII Hotine-Marussi symposium on mathematical geodesy, international association of geodesy symposia 137. Springer, Berlin

Teunissen PJG, Khodabandeh A (2015) Review and principles of PPP-RTK methods. J Geod 89(3):217–240

Van Le H, Nishio M (2019) Structural change monitoring of a cable-stayed bridge by time-series modeling of the global thermal deformation acquired by GPS. J Civ Struct Health Monit 9(5):689–701. https://doi.org/10.1007/s13349-019-00360-9

Weber G, Mervart L (2007) The BKG ntrip client (BNC). EUREF symposium 2007, London

Xin J, Zhou J, Yang SX, Li X, Wang Y (2018) Bridge structure deformation prediction based on GNSS data using Kalman-ARIMA-GARCH model. Sens. https://doi.org/10.3390/s18010298

Yang H, Xu C, Gao Y (2019) Analysis of GPS satellite clock prediction performance with different update intervals and application to real-time PPP. Surv Rev 51(364):43–52. https://doi.org/10.1080/00396265.2017.1359473

Ye Q, Szeto WY, Wong SC (2012) Short-term traffic speed forecasting based on data recorded at irregular intervals. IEEE Transact Intell Transp Syst 13(4):1727–1737. https://doi.org/10.1109/TITS.2012.2203122

Ye S, Zhao L, Song J, Chen D, Jiang W (2018) Analysis of estimated satellite clock biases and their effects on precise point positioning. GPS Solut. https://doi.org/10.1007/s10291-017-0680-z

Zumberge JF, Heflin M, Jefferson DC, Watkins M, Webb F (1997) Precise point positioning for the efficient and robust analysis of GPS data from large networks. J Geophys Res. https://doi.org/10.1029/96JB03860

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Mutaz Qafisheh is a Ph.D. student at the Department of Cartographic Engineering, Geodesy, and Photogrammetry at Universitat Politècnica de València (UPV). His main research areas are devoted to GNSS, Physical Geodesy, Big Data, and Machine Learning applied to Geosciences.

Angel Martin received his Ph.D. degree in Geodesy and Cartography at Universitat Politècnica de València (UPV) in 2001. He is a senior scientist and a lecturer professor at the Department of Cartographic Engineering, Geodesy, and Photogrammetry at UPV. His main research areas are devoted at GNSS, Physical Geodesy, Big Data, and Machine Learning applied to Geosciences.

Raquel M. Capilla Raquel M. Capilla received a Ph.D. degree in Geodesy, Cartography, and Geographic Information Systems at the Universitat Politècnica de València (Spain) in 2015. She works at the Cartographic Institute of Valencia as a GNSS projects technician. Her current research interests include Satellite Geodesy, Global Navigation Satellite Systems (GNSS), realtime GNSS services, and PPP algorithms and applications in Cartography, Geomatics, Engineering, and Natural Hazards.

Ana B. Anquela received her Ph.D. degree in Geodesy and Cartography at Universitat Politècnica de València (UPV) in 2001, where she is a senior scientist and a lecturer professor at the Department of Cartographic Engineering, Geodesy, and Photogrammetry. Her main research areas are GNSS and geodetic monitoring. She is also the director of the geomatic school at UPV.