

## Research Article

# Analyzing the Impact of Climate Factors on GNSS-Derived Displacements by Combining the Extended Helmert Transformation and XGboost Machine Learning Algorithm

Hanlin Liu <sup>1,2</sup>, Linqiang Yang <sup>3</sup>, and Linchao Li <sup>1,2</sup>

<sup>1</sup>College of Civil and Transportation Engineering, Shenzhen University, Shenzhen 518061, China

<sup>2</sup>Institute of Urban Smart Transportation & Safety Maintenance, Shenzhen University, Shenzhen 518061, China

<sup>3</sup>Department of Civil and Environmental Engineering, University of Hawaii at Manoa, Honolulu, HI 96822, USA

Correspondence should be addressed to Linchao Li; [lilinchao@szu.edu.cn](mailto:lilinchao@szu.edu.cn)

Received 8 March 2021; Accepted 7 May 2021; Published 28 May 2021

Academic Editor: Aijun Yin

Copyright © 2021 Hanlin Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A variety of climate factors influence the precision of the long-term Global Navigation Satellite System (GNSS) monitoring data. To precisely analyze the effect of different climate factors on long-term GNSS monitoring records, this study combines the extended seven-parameter Helmert transformation and a machine learning algorithm named Extreme Gradient boosting (XGboost) to establish a hybrid model. We established a local-scale reference frame called stable Puerto Rico and Virgin Islands reference frame of 2019 (PRVI19) using ten continuously operating long-term GNSS sites located in the rigid portion of the Puerto Rico and Virgin Islands (PRVI) microplate. The stability of PRVI19 is approximately 0.4 mm/year and 0.5 mm/year in the horizontal and vertical directions, respectively. The stable reference frame PRVI19 can avoid the risk of bias due to long-term plate motions when studying localized ground deformation. Furthermore, we applied the XGBoost algorithm to the postprocessed long-term GNSS records and daily climate data to train the model. We quantitatively evaluated the importance of various daily climate factors on the GNSS time series. The results show that wind is the most influential factor with a unit-less index of 0.013. Notably, we used the model with climate and GNSS records to predict the GNSS-derived displacements. The results show that the predicted displacements have a slightly lower root mean square error compared to the fitted results using spline method (prediction: 0.22 versus fitted: 0.31). It indicates that the proposed model considering the climate records has the appropriate predict results for long-term GNSS monitoring.

## 1. Introduction

Within the various remote sensing technologies, Global Navigation Satellite Systems (GNSS) plays an important role in providing fundamental infrastructure and has been successfully implemented in deformation monitoring. The global GNSS, such as the United States' Global Positioning System (GPS), Russia's Global Navigation Satellite System (GLO-NASS), the European Union's Galileo, and China's Beidou Navigation Satellite System (BDS), serve as highly efficient monitoring tools for precise geodetic surveying. Different unions or countries have installed numerous Continuously Operating Reference Stations (CORS) for various monitoring purposes, including the Plate Boundary Observatory (PBO)

maintained by National Science Foundation (NSF) EarthScope, the CORS GPS network maintained by the U.S. National Geodetic Survey, and the GPS Earth Observation Network (GEONET) of Japan. More than 506 worldwide permanent GNSS stations are managed by the International GNSS Service (IGS) group as of December 08, 2019. The original RINEX (Receiver Independent Exchange Format) files for the GNSS stations in different CORS networks are free to download through University NAVSTAR Consortium (UNAVCO) or National Geodetic Survey (NGS) data archiving facilities [1, 2]. The GPS, originally NAVSTAR GPS, is a widely used satellite positioning system, and it especially refers to the GNSS owned and operated by the United States. In the United States, the GPS signal has been widely used in

the monitoring service since 1989 [3]. Although this study only used GPS signals, GNSS is currently used as an umbrella term for all the aforementioned global satellite positioning systems. For this reason, this study uses the term GNSS.

Understanding the climate variations of GNSS time series are important for monitoring applications. In practice, during the period of half, one, or two years, the long-term monitoring time series have cyclical fluctuation and rebound characteristics triggered by climate influence, which makes the users who do not major in geodesy confuse about the precision of GNSS monitoring. The climate factors, mainly including the rainfall, temperature difference, wind speed, visibility, dew, and humidity, have always been a question among the long-term monitoring GNSS operators and data users. Xu et al. [4] proved that climate change led to the periodical variety of thermal expansion of bedrock (TEB). Dong et al. [5] established the hybrid model to remove the seasonal variations from long-term time series. Yan et al. [6] combined the thermal expansion model with the mass loading model to study the observed annual GNSS height changes. Munekane [7] proposed that the large temperature difference induced the maximum vertical displacements of three millimeters within 24 hours and established a model to move it. The climate variations are recognized as impact factors of the target deformation. Moreover, since the climate includes a variety of parts, the effects due to local climate variations on GNSS observations need to be carefully further studied. Thus, it is necessary to quantitatively analyze the weights of impacts caused by different climate factors on the long-term GNSS monitoring time series.

To achieve the quantitative analysis results of the impact on the GNSS time series caused by different daily climate factors, we need high-precision GNSS records. The accuracy and precision of GNSS observations are impacted by the types of GNSS equipment used and the processing method applied. In general, there are two widely used GNSS data postprocessing approaches: relative positioning and absolute positioning. The relative positioning approach uses the carrier-phase double-difference (DD) method to fix differenced phase ambiguities to integer values between stations and between satellites [8]. The DD method uses simultaneous GNSS records from at least two GNSS stations, and at least one of the GNSS stations should be installed in a stable foundation or known position with respect to a specific reference frame. The absolute positioning approach only needs a single GNSS station to determine the position without using any synchronous data from other stations. This method used undifferenced dual-frequency pseudo-range and carrier-phase observations in addition to precise satellite orbits and clock information to determine the position of a stand-alone GNSS station [9].

Precise Point Positioning (PPP) is a typical GNSS postprocessing absolute positioning approach, which uses a single-receiver phase-ambiguity-fixed resolution to calculate daily original raw data. The precision of the PPP solutions has dramatically increased during the last decade, which primarily attributes to highly precise satellite orbits and clock data provided by the International GNSS Service (IGS) and new algorithms used to resolve phase ambiguity within a sin-

gle receiver [10]. Moreover, the PPP algorithm has operational simplicity and can provide consistent accuracy [11, 12]. Considering the risk of lacking high-quality reference GNSS station when using the relative DD method, it is convenient to use the PPP method to process the GNSS data that monitor the long-term structural deformation or surface land movements. In this study, the software package GIPSY/OASIS (V6.4) employing the PPP processing strategy was applied [13]. The estimation of wide lane and phase bias was maintained by the Jet Propulsion Laboratory (JPL). When applying the GIPSY/OASIS software to solve daily positions, they are processed within the  $X$ ,  $Y$ , and  $Z$  geocentric Cartesian coordinate systems. The widely accepted global reference named the International Terrestrial Reference Frame (ITRF) is derived from the method of minimizing the overall horizontal movements of selected global permanent stations [14]. Using the GNSS to study the global tectonic movement, the most common global reference frame based on GNSS sites is required. Currently, the latest earth-centered, earth-fixed global reference frame International GNSS service of 2014 (IGS14) is established and aligned to ITRF2014 datum, which was updated from ITRF2008 in January 2017 [15]. However, the reference frame selection is motivated by the purpose of the study. In a continental-scale study area, such as North American, the NA12 was defined by 299 GPS stations and designed to corotate with the stable interior of the North American tectonic plate [16]. Also, another North American Plate fixed reference frame (NAD83 reference frame) was established resulting from requirements of the coordinates for sites located in the Conterminous United States (CONUS), Alaska, and US territories in the Caribbean. However, the high-precision GNSS monitoring time series inevitably include background tectonic motion within the subcentimeter level in the continental scale reference frame [17]. The millimeter-level internal tectonic movements could be easily obscured and biased within an inappropriate reference frame [18]. For this reason, there is a need to establish a stable regional or local reference frame for studies focusing on regional or local scales.

The scale, the orientation, the origin, and the change of these parameters over time are the main physical and mathematical properties of a reference frame. In geodetic applications, a stable local reference frame is primarily transformed from the latest and well-established global reference frame using the Helmert transformation. A group of GNSS reference stations (common points) are used to tie the target regional or local reference frames to a global reference frame (e.g., IGS14). Pearson and Snay [19] used a 14-parameter transformation method to transform coordinates from the global reference frame IGS08 to the regional reference frame NAD83, which is maintained by NGS. Theoretically, the coordinates of a point referred to the IGS14 global reference frame vary when they are transformed to a regional or local reference frame. Nevertheless, deformation monitoring mainly focuses on the deformation rates/displacements of the target study area rather than the coordinates themselves in both reference frames. In general, relative high precision displacements or deformation velocities are needed in different area monitoring applications.

In practice, after the highly precise GNSS records are processed, it is still a challenge to quantitatively analyze the weights of impact from different climate factors on the GNSS time series. However, with computing science development, the machine learning approach achieves a dramatic developing rate and provides a new tool to explore new analysis methods in geodesy and geosciences. The approach can quantitatively analyze the hypothesis and assist to capture high-dimensional data sets [20, 21]. Furthermore, machine learning methods have been widely applied in geosciences. Rouet et al. [22] evaluate the fault movements in the Cascadia subduction zone. Phrampus et al. [23] predict the probability of encountering global SEAFLEA. Anderson and Lucas [24] proposed the climate ensemble method to predict weather and establish various climate models. Ren et al. [25] used the machine learning algorithm to estimate the fault friction with high precision. The machine learning algorithm has the potential to help researchers make further explanations and explore the theories behind the questions in geosciences. Therefore, we applied the ensemble learning method to analyze the weights of impact from various daily climate factors on the GNSS monitoring time series. In this study, we used an appropriate supervised machine learning (SML) algorithm, Extreme Gradient Boosting (XGBoost), to determine the relationship between various daily climate factors and highly precise GNSS data.

In this paper, we proposed a hybrid method to evaluate the impact of daily climate factors on the GNSS time series, using the extended Helmert transformation method and XGBoost algorithm. The model is trained by high-precision GNSS records and various long-term climate records. The contributions of this paper are shown as follows:

- (1) To remove the background tectonic movements when monitoring local ground deformation, we proposed the extended Helmert transformation to establish the highly stable PRVI19 local reference frame based on ten well-distributed continuously operating GNSS stations with at least five years of data
- (2) By combining the GNSS records with millimeter accuracy and the local climate data with a span of at least five years, we applied the XGboost machine learning algorithm to derive the quantitative results of the weights of impact from different daily climate factors on the GNSS time series
- (3) Based on the model, we predicted the GNSS records and validate them with the real raw GNSS data. The results show that the high accuracy of the prediction and it is expected that this study can provide a new prospect to explore the potential deformation monitoring problem

## 2. Data and Methods

*2.1. Coordinate Transformation Using the Extended Helmert Transformation Method.* The Helmert transformation, also called a 7-parameter transformation, is used to conduct distortion-free reference frame transformations within a

three-dimensional (3D) space in the geodesy area. The 7-parameter approach employs three parameters from translations, three parameters from rotations, and one parameter from the scale at a selected epoch and the rates of these seven parameters over time. For daily GNSS positional coordinate transformation from the PPP solutions, the geocentric coordinates of a site with respect to the local reference frame can be approximated by the following formula:

$$X(t)_{\text{local}} = T_X(t) + [1 + s(t)] \bullet X(t)_{\text{IGS14}} + R_Z(t) \bullet Y(t)_{\text{IGS14}} - R_Y(t) \bullet Z(t)_{\text{IGS14}}, \quad (1)$$

$$Y(t)_{\text{local}} = T_Y(t) + R_Z(t) \bullet X(t)_{\text{IGS14}} + [1 + s(t)] \bullet Y(t)_{\text{IGS14}} + R_X(t) \bullet Z(t)_{\text{IGS14}}, \quad (2)$$

$$Z(t)_{\text{local}} = T_Z(t) + R_Y(t) \bullet X(t)_{\text{IGS14}} - R_X(t) \bullet Y(t)_{\text{IGS14}} + [1 + s(t)] \bullet Z(t)_{\text{IGS14}}, \quad (3)$$

where  $R_X(t)$ ,  $R_Y(t)$ , and  $R_Z(t)$  are counterclockwise rotations along  $X$  axis,  $Y$  axis, and  $Z$  axis of the geocentric coordinate system;  $T_X(t)$ ,  $T_Y(t)$ , and  $T_Z(t)$  are translations about these three axes; and  $s(t)$  is a differential scale change between IGS14 and local reference frame.

Theoretically, the seven parameters for each epoch/day are different. Therefore, to obtain the positional time series in the local reference frame, the seven parameters at each epoch have to be provided. Currently, there are two strategies that can be employed to perform the transformation, daily seven-parameter transformation, and 14-parameter transformation [19]. Continuous observations have made it possible to calculate the seven transformation parameters on a daily basis and transform the daily positional time series from one reference frame to another. However, calculating daily transformation parameters is too complex for most end-users, since users have to include a large number of reference stations for calculating transformation parameters at each epoch. Besides, it is difficult to ensure the data quality and availability for all the reference stations at each day. Thus, the GNSS time series of these seven transformation parameters can be simulated by the linear method with equations:

$$\alpha_t = \begin{bmatrix} T_X(t) \\ T_Y(t) \\ T_Z(t) \end{bmatrix}, \beta_t = \begin{bmatrix} R_X(t) \\ R_Y(t) \\ R_Z(t) \end{bmatrix}, \alpha' = \begin{bmatrix} T'_X \\ T'_Y \\ T'_Z \end{bmatrix}, \beta' = \begin{bmatrix} R'_X \\ R'_Y \\ R'_Z \end{bmatrix}, \quad (4)$$

$$\alpha_t = \alpha_{t_0} + \alpha' \bullet (t - t_0), \quad (5)$$

$$\beta_t = \beta_{t_0} + \beta' \bullet (t - t_0), \quad (6)$$

$$s(t) = s(t_0) + s' \bullet (t - t_0). \quad (7)$$

Here,  $t_0$  denotes a specific epoch, which is set as specific epoch.  $T_X(t_0)$ ,  $T_Y(t_0)$ ,  $T_Z(t_0)$ ,  $R_X(t_0)$ ,  $R_Y(t_0)$ ,  $R_Z(t_0)$ , and  $s$

( $t_0$ ) are the seven Helmert transformation parameters at specific epoch  $t_0$ .  $T'_X, T'_Y, T'_Z, R'_X, R'_Y, R'_Z$  and  $s'$  are the first derivatives (rates) of these seven parameters, which are assumed to be constant; the units of the translational components ( $T'_X(t), T'_Y(t), T'_Z(t)$ ) are meters/year; the units of the three rotational components ( $R'_X(t), R'_Y(t), R'_Z(t)$ ) are radians/year; the units of the rates of translational components ( $T'_X, T'_Y, T'_Z$ ) are meter/year; the units of the rates of the rotational components ( $R'_X, R'_Y, R'_Z$ ) are radian/year, and  $s(t)$  is a unitless scale. The local reference frame is tied to the same origin and scale of the original global reference frame, which is the IGS14 global reference frame in this study. The above formulas also require the use of a scale factor  $s(t)$  to minimize the distortion of point-to-point distances between the two reference frames. In general, the scale factor  $s(t)$  can be set as zero in frame transformations between a global reference frame and a local reference frame [26]. The common points are used to solve the inverse problem by the least square method.

Then, the coordinates of the target GNSS stations referred to the local reference frame can be obtained through:

$$X(t)_{\text{local}} = X(t)_{\text{IGS14}} + T'_X \bullet (\Delta t) + R'_Z \bullet (\Delta t) \bullet Y(t)_{\text{IGS14}} - R'_Y \bullet (\Delta t) \bullet Z(t)_{\text{IGS14}}, \quad (8)$$

$$Y(t)_{\text{local}} = Y(t)_{\text{IGS14}} + T'_Y \bullet (\Delta t) - R'_Z \bullet (\Delta t) \bullet X(t)_{\text{IGS14}} + R'_X \bullet (\Delta t) \bullet Z(t)_{\text{IGS14}}, \quad (9)$$

$$Z(t)_{\text{local}} = Z(t)_{\text{IGS14}} + T'_Z \bullet (\Delta t) + R'_Y \bullet (\Delta t) \bullet X(t)_{\text{IGS14}} - R'_X \bullet (\Delta t) \bullet Y(t)_{\text{IGS14}}, \quad (10)$$

$$\Delta t = (t - t_0). \quad (11)$$

Also, since a linear model is assumed, the changing rates can be easily calculated with two sets of transformation parameters at two different epochs. In this study, the epochs are set as  $t_0$  (i.e., 2015) and  $t_1$  (i.e., 2018). So, the coordinates in both frames at epoch  $t_0$  are the same, and the seven parameters are all zeros. The method aligns the two frames at the epoch  $t_0$ . The coordinates at epoch  $t_0$  are calculated with the following equation:

$$\begin{pmatrix} X(t_0)_{\text{IGS14}} \\ Y(t_0)_{\text{IGS14}} \\ Z(t_0)_{\text{IGS14}} \end{pmatrix} = \begin{pmatrix} X(t_0)_{\text{local}} \\ Y(t_0)_{\text{local}} \\ Z(t_0)_{\text{local}} \end{pmatrix}. \quad (12)$$

In order to calculate the transformation parameters at another epoch  $t_1$ , the coordinates at epoch  $t_1$  in IGS14 and local reference frame can be derived by using the following equations:

$$\begin{pmatrix} X(t_1)_{\text{local}} \\ Y(t_1)_{\text{local}} \\ Z(t_1)_{\text{local}} \end{pmatrix} \approx \begin{pmatrix} X(t_0)_{\text{local}} \\ Y(t_0)_{\text{local}} \\ Z(t_0)_{\text{local}} \end{pmatrix}. \quad (13)$$

Here, the velocity in local reference frame is regarded as

zero since local reference frame is designed to have a velocity of zero relative to the rigid part of the region. With the coordinates at epoch  $t_0$  in both IGS14 and local reference frame, equations (1), (2), and (3) can be solved for the seven parameters at epoch  $t_0$ . By knowing the six parameters for epoch  $t_0$  and  $t_1$ , the parameter rates can be obtained by a simple differential method:

$$T'_X = \frac{(T_X(t_1) - T_X(t_0))}{\Delta t}, \quad (14)$$

$$R'_X = \frac{(R_X(t_1) - R_X(t_0))}{\Delta t}. \quad (15)$$

Then, the coordinates of a GNSS site at any epoch can be transformed from IGS14 to the local reference frame with  $T'_X, T'_Y, T'_Z, R'_X, R'_Y$ , and  $R'_Z$  parameters.

**2.2. Extreme Gradient Boosting (XGBoost) Algorithm.** In this study, we applied the Extreme Gradient Boosting (XGBoost) algorithm to predict ground displacements and to understand which climate factors have more impact on the GNSS monitoring time series. XGBoost was proposed by Chen et al. [27], which can combine many regression trees into one strong ensemble learner. Because of the ensemble, XGBoost can sufficiently capture complex interaction of variables in monitoring time series and then fit the nonlinear dynamic changes of displacements. This is one of the motivations for us to select XGBoost in this paper. Another reason is that the strong explanation power of XGBoost can help us understand the relationship between different daily climate variables and predict the high-precision GNSS displacements.

In XGBoost,  $K$  additive regression trees are together to predict displacements. That is:

$$\hat{y}_i = \sum_{k=1}^K g_k(x_i), g_k \in \mathcal{F}, \quad (16)$$

where  $\mathcal{F} = \{g(x) = w_{q(x)}\}$  ( $q: \mathcal{R}^m \rightarrow T, w \in \mathcal{R}^T$ ) is the space of used regression trees,  $x_i$  is a data set containing historical displacements, each  $g_k$  represents an independent regression tree structure  $q$  and leaf weights  $w$ , and  $T$  represents the number of leaves of the regression tree.

The optimization function of XGBoost can be written as:

$$\mathcal{L} = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(g_k), \quad (17)$$

$$\Omega(g) = \beta T + \frac{1}{2} \lambda \|w\|^2, \quad (18)$$

where  $\beta$  and  $\lambda$  are coefficients,  $l$  is a differentiable convex loss function which represents the differences between measured displacements and predicted displacements, and  $\Omega(g)$  is the penalty term of the objective function which helps avoid the overfitting problem.

However, because of the complex architecture, it is difficult to train the ensemble learner once. An additive strategy

has been widely applied which means trees are trained one by one. The trees that have already been trained will be fixed, and then, a new tree is added at one step. Suppose that the predicted displacement at step  $t$  is  $\hat{y}_i^{(t)}$ , then the optimization function at this step can be written as:

$$\mathcal{L}_t = \sum_i l(y_i, \hat{y}_i^{(t-1)} + g_t(x_i)) + \sum_k \Omega(g_t) + C, \quad (19)$$

where  $C$  is a constant. To approximate the optimization function using a second-order Taylor expansion, it can be rewritten as:

$$\mathcal{L}_t = \sum_{i=1}^n \left[ a_i g_t(x_i) + \frac{1}{2} b_i g_t^2(x_i) \right] + \Omega(g_t), \quad (20)$$

where

$$a_i = \frac{\partial l(y_i, \hat{y}_i^{(t-1)})}{\partial \hat{y}_i^{(t-1)}}, \quad (21)$$

$$b_i = \frac{\partial^2 l(y_i, \hat{y}_i^{(t-1)})}{\partial (y \wedge_i^{(t-1)})^2}. \quad (22)$$

With equations (19) and (20), the optimization function can be revised as:

$$\mathcal{L} = \sum_{j=1}^T \left[ \left( \sum_{i \in I_j} a_i \right) w_j + \frac{1}{2} \left( \sum_{i \in I_j} b_i + \lambda \right) w_j^2 \right] + \beta T, \quad (23)$$

where  $I_j = \{i \mid q(x_i) = j\}$ . Given a tree structure  $q(x)$ , the weight of the  $j$ th node and extreme value of  $\mathcal{L}(q)$  can be obtained as:

$$w_j = - \frac{\sum_{i \in I_j} a_i}{\sum_{i \in I_j} b_i + \lambda}, \quad (24)$$

$$\mathcal{L}(q) = - \frac{1}{2} \sum_{j=1}^T \frac{\left( \sum_{i \in I_j} a_i \right)^2}{\sum_{i \in I_j} b_i + \lambda} + \beta T. \quad (25)$$

As studied by Breiman and Friedman [28], for a single regression tree (RT), the following equation can be used to measure the relative importance of independent variables in the tree:

$$I_d^2(\text{RT}) = \sum_{t=1}^{T-1} \alpha_t^2 I(v(j) = d), \quad (26)$$

where  $t$  is a nonterminal node and  $T$  is the terminal node of tree RT,  $x_d$  is the splitting variable of node  $t$ , and  $\alpha_t^2$  is the improvement of square error of the prediction if  $x_d$  is used as splitting variable. For an ensemble learner, which is a collection of  $M$  regression trees  $\{\text{RT}_m\}_1^M$ , the relative impor-

tance of variable  $x$  can be obtained by its average over all additive trees:

$$I_d^2 = \frac{1}{M} \sum_{m=1}^M I_d^2(\text{RT}_m). \quad (27)$$

The XGBoost is implemented using the “xgboost” package in the R software [29].

**2.3. Data and Selection.** In geodesy, a terrestrial reference frame is realized by selecting a set of reference stations and defining their positions and velocities. The selection of the reference stations is critical for establishing a stable reference frame. Here, it is very hard to set any mathematical or technical standard for selecting appropriate GNSS reference stations. In general, with some previously proposed guidelines, the reference stations are selected based on overall geographic distribution and long-term (e.g., >5 years) continuous records [30, 31]. Also, it requires no considerable subsidence or uplift with vertical velocity magnitude less than 0.5 mm/year referred to the global reference frame IGS14. A reference station that is not locally stable will degrade the overall stability or precision of the reference frame. The selection is mostly based on the availability of long-term CORS in the study area. In addition, if the selected GNSS sites have detailed station logs, which it may help explain the unexpected steps. The step means that the GNSS time series have sudden ascending or descending jump induced by the earthquake, volcano eruption, or/and GNSS equipment change. In this paper, we selected Puerto Rico and the Northern Virgin Islands (PRVI) to be the target area. In the tectonic setting, the motivation results from the results that researchers proposed the exist of rigid block of Puerto Rico and the Northern Virgin Islands (PRVI) between the Caribbean Plate and the North American Plate [32, 33]. The area is recognized as a tectonic stable microplate without notably tectonic movements in the Caribbean area and can minimize the influence from tectonic movements. The PRVI region is recognized as an appropriate area to apply the method.

Also, the PRVI region has established an appropriate GNSS infrastructure and has a long-term land monitoring history. Since 1986, GNSS stations were installed by researchers for studying Caribbean plate tectonic movements [34]. In the PRVI region, as of 2020, 28 GNSS stations are recording the displacement observations with 15 seconds positioning. GNSS station coverage within this region is exceptionally dense. The PRVI region owns the densest GNSS stations in the United States and has been well monitored by numerous permanent GNSS stations [30]. For example, the coverage along the coast of Puerto Rico is approximately 20.0 stations per kilometer. Eight GNSS stations within the PRVI region were installed by the Puerto Rico Seismic Network (PRSN) at the University of Puerto Rico at Mayaguez (UPRM). The installation was primarily funded by a National Science Foundation (NSF) Major Research Instrumentation (MRI) project (EAR-0722540). The daily GNSS records with a sample rate of 15 seconds are archived at UNAVCO public data archiving facility.

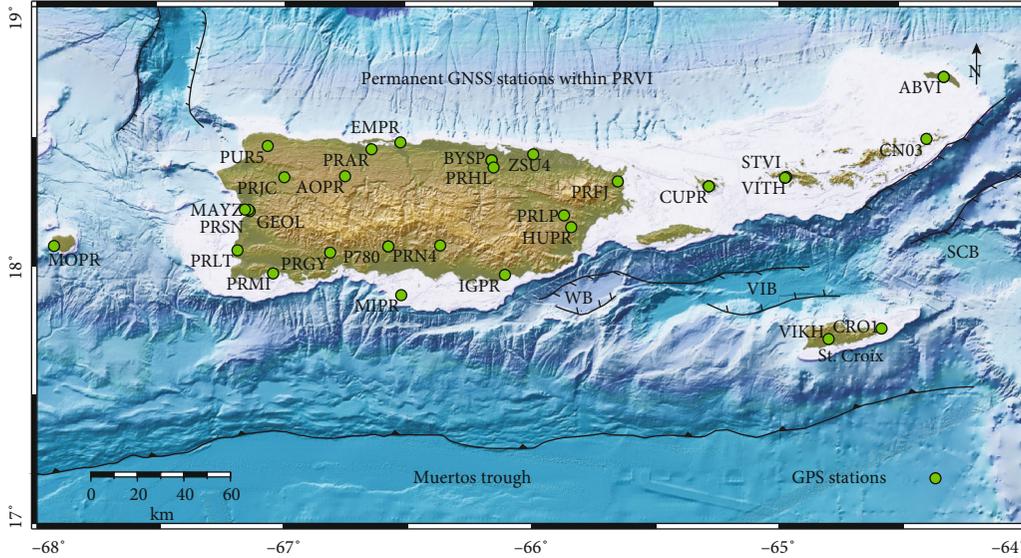


FIGURE 1: Map showing the locations of current GNSS stations in the Puerto Rico and Virgin Islands region. VI represents the Virgin Islands; WB represents the Whiting basin; VIB represents the Virgin Islands basin; SCB represents the St. Croix basin.

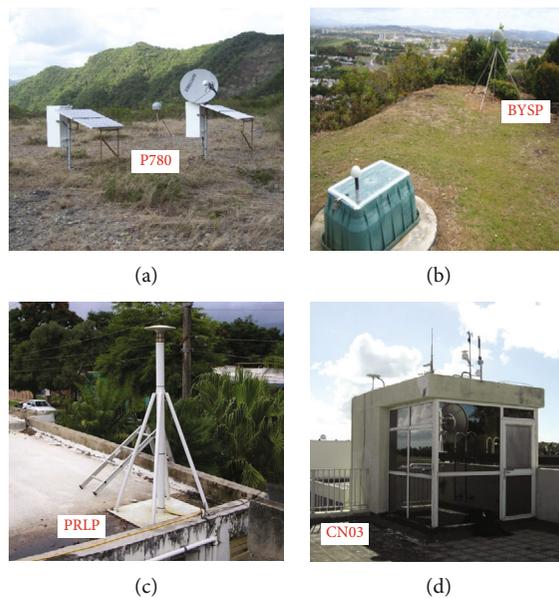


FIGURE 2: Figure showing the site views of four typical continuously operating permanent GNSS stations in the PRVI region. (a) A site view at Cerrillos\_PR2008 GNSS station (ID: P780) installed in Ponce, PR. The antenna is mounted to a short-drilled braced monument (SDBM) designed by UNVCO. (b) A site view at BayamonSciPR2008 GNSS station (ID: BYSP) and a nearby strong motion accelerometer installed in the Bayamón Science Park, Bayamón, PR. (c) A typical building based GNSS station (ID: PRLP) installed by the HLCM Group Inc. in Las Piedras, PR. (d) A site view at DVirgGordaBVI2013 GNSS station (ID: CN03) installed at the top of a mountain in North Sound, Virgin Gorda.

The HLCM Group Inc. also installed eight GNSS stations in the PRVI region for land surveying. The HLCM Group Inc. is a private surveying company, and it archives the GNSS data at the NGS data archiving facility. Other public agencies, such as the U.S. Coast Guard, Federal Aviation Administration (FAA), and Jet Propulsion Laboratory (JPL), also operate GNSS stations in the PRVI region. Therefore, the well-established GNSS network and the complete equipment

maintenance record offer a strong foundation to apply the experiment (Figure 1). Moreover, since the GNSS signals are likely to be blocked by the circumstance, the GNSS stations need to be installed in the open area. The open area, such as the building roof or top of the mountain, could keep minimum interference from GNSS multiple paths and signal block. Figure 2 shows the site views at four typical permanent GNSS stations in the PRVI region, all of which were installed

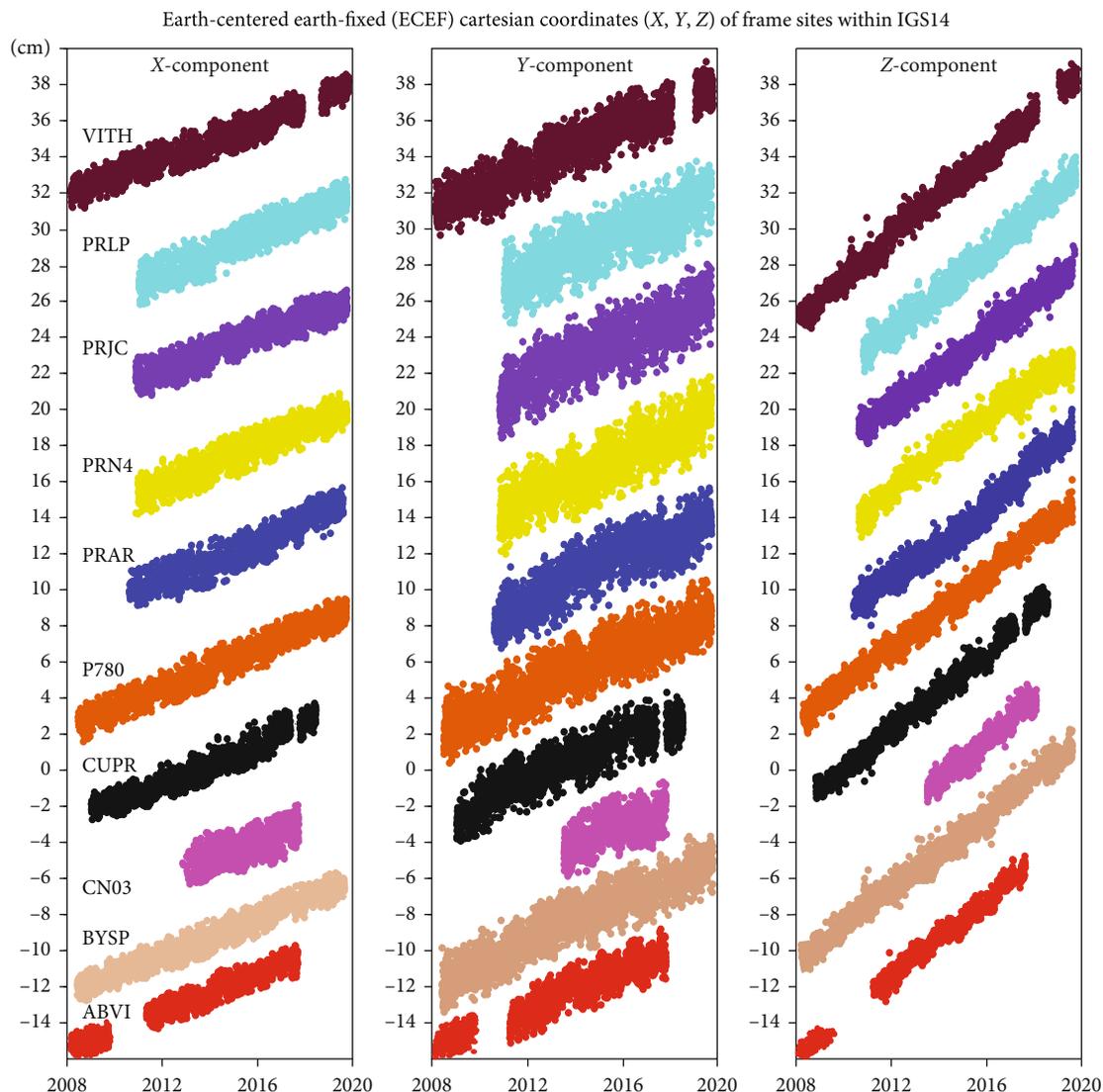


FIGURE 3: Plots showing the geocentric positional time series (X, Y, Z) of the ten frame sites with respect to the global reference frame IGS14.

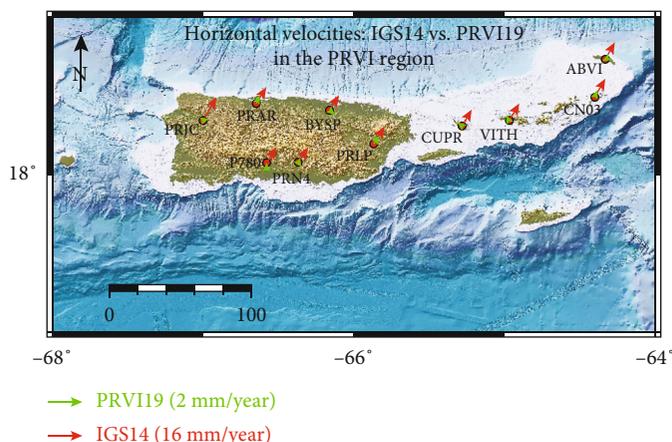


FIGURE 4: Map showing the locations of the ten selected GNSS reference stations within the PRVI region. The red velocity vectors are referred to the IGS14 reference frame. The green velocity vectors are referred to the PRVI19 reference frame.

TABLE 1: Detailed information of the ten reference stations and the site velocities referred to IGS14 and PRVI19 reference frames.

Reference GNSS	Location (degree)			Site velocity						Uncertainty of the velocity (mm/year)**		
				IGS14, mm/year			PRVI19, mm/year					
	Location	Longitude	Latitude	EW	NS	UD	EW	NS	UD	EW	NS	UD
P780	Ponce, PR	-66.5791	18.075	8.9	13.8	-0.3	0.1	0.1	-0.3	0.2	0.2	0.8
BYSP	Bayamón, PR	-66.1612	18.408	8.8	13.8	-0.5	0.2	-0.2	-0.9	0.2	0.2	0.8
PRJC	San Sebastian, PR	-66.9995	18.342	9.0	13.6	-0.8	0.4	0.1	-0.9	0.3	0.3	1.0
PRLP	Las Piedras, PR	-65.8683	18.195	9.3	14.8	0.1	0.5	0.7	-0.1	0.3	0.3	1.0
PRAR	Cercadillo, PR	-66.647	18.450	8.9	14.4	-0.5	0.3	0.7	-0.9	0.3	0.3	0.9
PRN4	Coamo, PR	-66.369	18.079	9.0	13.6	0.6	0.2	-0.3	0.6	0.3	0.3	1.1
CUPR	Culebra Island, PR	-65.2825	18.307	8.5	14.6	-0.5	-0.3	0.2	-0.9	0.3	0.3	1.0
VITH	St. Thomas Island, USVI	-64.9692	18.343	8.6	14.8	-0.1	-0.1	0.2	-0.5	0.2	0.3	0.9
CN03	Gorda Peak, BVI	-64.403	18.490	8.9	15.5	-0.2	0.2	0.6	-0.5	0.4	0.4	1.3
ABVI	Anegada Island, BVI	-64.3325	13.730	8.4	14.9	-0.3	-0.2	-0.1	-0.9	0.3	0.3	1.0

\*PR: Puerto Rico; USVI: U.S. Virgin Islands; BVI: British Virgin Islands. \*\*The uncertainty represents the 95% confidence interval of the velocity estimate.

TABLE 2: Seven parameters for Helmert reference frame transformation from IGS14 to PRVI19.

Parameters	Unit	IGS14 to PRVI19
dTx	m/year	2.652437E-02
dTy	m/year	-2.130450E-02
dTz	m/year	-9.310699E-02
dRx	Radian/year	1.386115E-08
dRy	Radian/year	1.273412E-09
dRz	Radian/year	5.511930E-09
$t_0$	1/year	2015.0

\*Counterclockwise rotations of axes (X, Y, Z) are positive.

in the open area to avoid the GNSS signal block or multipath effects, with solid monuments, i.e., the short-drilled braced monument, and building mount monument.

The climate impacts the seasonal ground deformation, and different climate factors perform different impact weights in the GNSS time series. In this study, PRVI region that presently follows a tropical rainforest climate is selected to test the hybrid method. The area has recorded an annual mean temperature of 28°C, and it has a trend of increasing since the 1950s [35]. The precipitation records show a long-term trend of decreasing precipitation in the northern area and show historically drier with a positive rainfall trend in the southern area [36]. Since the PRVI is one of the islands located in the Caribbean region, it has a wind speed of 10 mph~14 mph which is recorded at 10 meters above the ground. However, hurricane activity has increased since 1995. This increase in hurricane activity may also be a result of natural variability [37]. Also, historical records from tide gauges since 1900 show a Sea Level Rise (SLR) of 1.7 mm/year. Recent satellite-based remote sensing technology shows a SLR rate of 3.2 mm/year since 1992 [38]. Moreover, since the snow/ice has different characteristics in a variety of areas (such as Alaska and Texas in the U.S), it leads to a tremendously different impact weight in GNSS time series [39]. Notably, we expect the hybrid method to be fully

tested and have the characteristics of generality and efficiency. Thus, we select the location without snow/ice to test the hybrid method. Besides, to precisely evaluate the impact caused by different climate factors, the selected study area needs to have consistent climate data. We choose to use the original daily climate monitoring continuous station which is installed near the permanent continuous GNSS station in the PRVI region.

### 3. Results

**3.1. PRVI19 Local Reference Frame.** This study used the GIPSY-OASIS software package (V6.4) to obtain daily solutions using the PPP method. Firuzabadi and King [40] and Wang [41] proposed an outlier detection and removing algorithm to clean the positional or displacement time series. On average, 6~7% of the total samples have been removed as outliers in this study. According to the previous investigations, daily PPP solutions would achieve 3-5 mm horizontal accuracy and 5-8 mm vertical accuracy in the PRVI area [30]. The PPP solutions of the GNSS stations are provided with respect to a global reference frame. The positional time series referred to the IGS14 global reference frame are shown in Figure 3. Also, the raw PPP solutions inevitably include the common background of global tectonic movements. Thus, the GNSS observations referred to a stable local reference frame could minimize the influence induced from the global background movements.

In this study, we used the extended Helmert transformation and selected ten permanent stations to realize the PRVI19 reference frame. The MIDAS method was used to calculate velocity and uncertainty [42]. The velocities of the ten selected reference stations with respect to different reference frames are shown in Figure 4. The locations of these permanent stations and corresponding site velocities referred to IGS14 and PRVI19, respectively, are listed in Table 1. The seven Helmert parameters for transforming the global IGS14 reference frame to the local PRVI19 reference frame are listed in Table 2. Moreover, the plate reconstructions

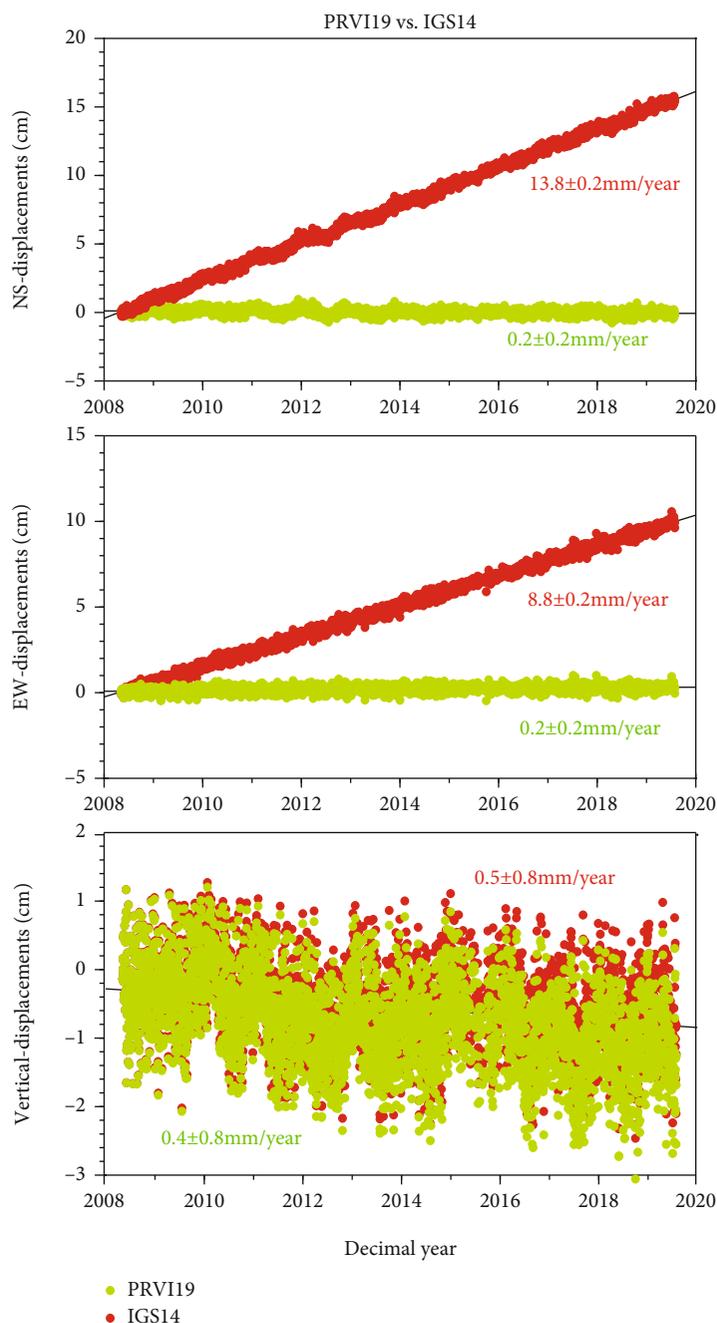


FIGURE 5: Plots showing the three-component displacement time series of the BYSP station with respect to two different scale reference frames: a global reference frame (IGS14) and a local reference frame (PRVI19). The green dots represent the displacements with respect to the PRVI19 reference frame during 2008-2020. The red dots represent the displacements referred to the IGS14 reference frame within the same period of time.

and geodynamic modeling are defined by the rigidity of tectonic plates. The stability of the local reference frame also relies on the assumption of the rigidity of the tectonic plate. The stability or the accuracy of the reference frame can be assessed by the average velocities of all reference stations with respect to the reference frame. Since the reference frame is defined by the reference stations, in an ideal situation, all reference stations should have no relative movement between each other with respect to the defined reference frame. Thus, if the PRVI region is recognized as a rigid microplate in tec-

tonics, the stability of the local reference frame is to be 0 mm/year. However, in practice, a tectonic plate is not completely rigid, and it cannot be strictly stable. For PRVI19, the stability reaches approximately 0.4 mm/year in horizontal and approximately 0.5 mm/year in vertical. This level of accuracy is essential for studying millimeter-per-year ground motions in the PRVI region.

Theoretically, the longer length of GNSS data and better geographical distribution of reference stations can improve the stability of a reference frame. Wang et al. [18] updated

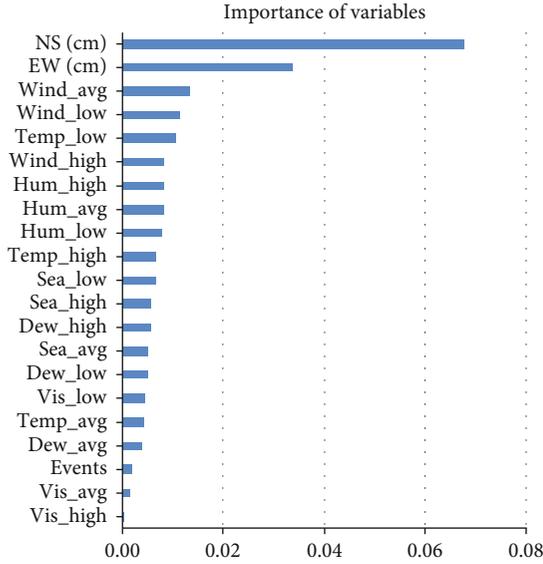


FIGURE 6: Plot showing the impact weights of different climate factors evaluated by the combined method. Sea: sea level; Temp: temperature; Hum: humidity; Vis: visibility. NS represents the North-South displacements. Events represent the operation log of weather station. EW represents the East-West displacements. Avg represents the average of daily values. Low represents the lowest daily value. High represents the highest daily value.

TABLE 3: XGBoost hyperparameters.

nrounds=400	nthread=8	eta=0.01
Gamma=0.01	max_depth=8	min_child_weight=2
Subsample=0.54	colsample_bytree=0.54	Lambda=0.01

the PRVI14 local reference frame to PRVI18 with the latest global reference frame (from IGS08 to IGS14) and reference stations. However, it indicated that there was no considerable improvement in reference frame stability. In this study, we established the local reference frame to PRVI19 with ten reference stations and a longer length of GNSS data. The PRGY, PRLT, and PRMI GNSS stations, which are located at south-western and southern of the PRVI region, are excluded from reference stations. Those stations are located nearby the South Lajas fault which is an active Holocene structure within the past 5,000 years [43]. The Lajas Valley currently experiences a 1.5 mm/year north-south direction extension and minor right-lateral strike-slip [9]. Also, the MIPR station is excluded from the reference stations because it is close to P780 and PRN4 in geographic distribution, and the GNSS station stops recording data since October of 2017. Though the location of MIPR station can form well geographical distribution for the local reference frame, we still excluded it from the reference stations because P780 station has better quality and longer length of data and locates in the same area. However, the stability of the PRVI19 still does not considerably increase compared with the previous two local reference frames, i.e., PRVI14 and PRVI18 [18, 31]. Based on the previous versions of local reference frames and PRVI19, it

proves that the stability of the PRVI19 local reference frame is approaching the rigidity of the PRVI tectonic block.

Figure 5 depicts nearly 13-year North-South, East-West, and vertical displacement time series at BYSP (2008-2020) with respect to the IGS14 and the PRVI19 reference frames. It shows a notable difference in horizontal velocities when referred to the global (IGS14) and local (PRVI19) reference frames. The 13 years of BYSP observations show that it has a velocity of  $13.8 \pm 0.2$  mm/year in NS and a velocity of  $8.8 \pm 0.2$  mm/year in EW with respect to the global IGS14 reference frame; it has a velocity of  $0.2 \pm 0.2$  mm/year in NS and a velocity of  $0.2 \pm 0.2$  mm/year in EW with respect to the local PRVI19 reference frame, respectively. In general, most engineer monitoring applications only focus on the local target deformation displacements and do not consider the movements with respect to a global reference frame. The comparison of the BYSP GNSS time series within IGS14 and PRVI19 reference frames demonstrates that the background global or regional tectonic movements can bias or obscure the local ground deformation when using a global or regional reference frame.

**3.2. Daily Climate Impact Factors.** Climate change is considered as an important external impact factor influencing the GNSS data precision. However, it is a challenge to clarify the relationship between daily climate change and daily GNSS records. The main reasons are because the daily climate change influence could be partially removed by the 24-hour GNSS processing method. Thus, we selected the continuous operating GNSS station BYSP, which is installed nearby a real-time climate-monitoring device. The 5-year continuous climate data are used in the model, which is collected by the weather station (TJSJ) nearby the BYSP (NOAA National Weather Service). Here, we established two models between which the only difference is that one considers the daily climate features and the other not. We used the model to evaluate whether climate change can influence the precision of the 24-hour GNSS time series, which has been transformed to the PRVI19 reference frame using the extended Helmert transformation method. The dimensionless index shows that the model without considering the daily climate change is 0.32 and the other one considering the daily climate change is 0.25. The lower dimensionless index means that the model has better performance. The results prove that daily climate change is one of the impact factors in the GNSS time series.

Furthermore, we determined the quantitative weights of impact from different daily climate factors on the GNSS time series. Figure 6 shows the importance of the 20 selected climate factors on the GNSS time series using the XGBoost model. The XGBoost parameters used to evaluate the displacements are shown in Table 3. Each row displays the impact of the feature, in which the contribution weight is shown on the  $x$ -axis. Here, we used two complementary features, North-South (NS) feature and East-West (EW) feature, which are the physical movements due to the vertical displacements. Theoretically, the NS and EW observations should change when the GNSS station has vertical movements. For this reason, we involve the two features to help

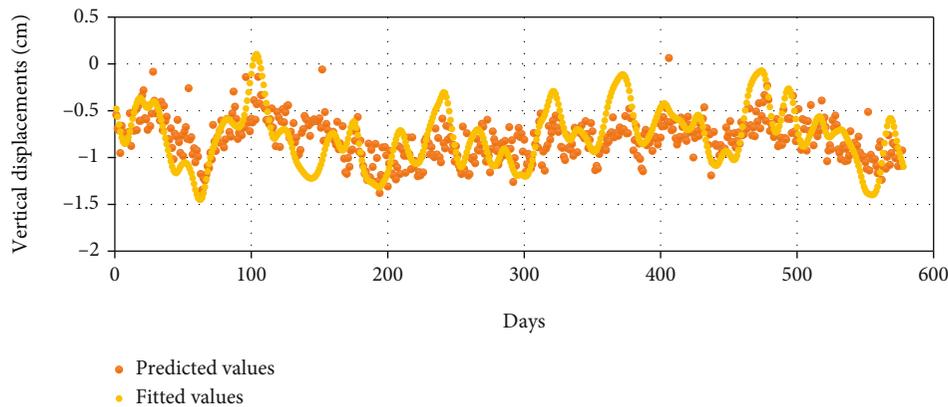


FIGURE 7: Plot showing the detailed comparisons between the fitted values from measured values and predicted values in 600 days. The orange dots represent the results predicted by the combined machine learning method. The yellow dots represent the fitted results based on GNSS records using the cubic spline method.

the hybrid model precisely evaluate different climate features and corresponding impact weights. The relationship between these two physical features is in good agreement with the results reported in Li et al. [44], which indicated that the key feature for the NS movements was the variance of the subsidence or uplift area. Based on the five years of monitoring climate records, we analyze the weight of different climate factors influencing the GNSS displacements. We found that the wind speed (average) and temperature (minimum) were the two climate factors that have the biggest impact on the daily GNSS time series. Interestingly, the humidity, temperature (average), and dew are not sensitive in the model. In the daily GNSS postprocessing records, except for the mega weather conditions, it is hard to determine the relationship between these features that can impact the vertical displacements.

**3.3. GNSS Time Series Prediction.** The hybrid model forecasted the GNSS monitoring displacements by learning the previous data and expected to explore the potential deformation problem in various monitoring applications. The prediction of GNSS observations is derived from the previous displacements that are referred to the stable local PRVI19 reference frame. Here, we used the BYSP GNSS postprocessing records referred to PRVI19 to train the model. The hybrid model was trained by the 1823 days of GNSS displacements. Also, we used the root mean square error (RMSE) as an indicator to evaluate the forecasting precision [45, 46]:

$$\text{RMSE} = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}, \quad (28)$$

where  $x$  is the prediction value,  $\bar{x}$  is the measured value, and  $n$  is the sample size of test dataset. In the experiment, we need to use the prediction results to compare with real deformation results monitored by the GNSS station. However, the GNSS time series still inevitably include the nonlinear change, such as the noise, which is complex and hard to remove. Gazeaux et al. [47] tested different models and methods to minimize the white noise and colored noise such

as the first-order Gauss Markov model and auto-regressive moving average model. However, the best noise model is location dependent, which is impacted by the location of the GNSS site, GNSS raw data processing method, and monument designs [48]. Thus, in the comparison, we used cubic spline function interpolation to fit the GNSS measured solutions, which have been transformed to PRVI19 [49]. Figure 7 shows the predicted model results compared with the post-fitted GNSS displacements using the cubic spline function in the experiment. We used optimized parameters for the cubic spline method and found the predicted values performed well. It shows that the RMSE of the forecasting method is 0.22, which is slightly lower than 0.31 from the fitted results using the cubic spline function. Moreover, we tested all parameters enrolled in the cubic spline method. It also shows that the precision of the forecasted results meets well with the spline fitted results, which are processed using the real raw GNSS records.

## 4. Conclusions

To precisely analyze the impact of various daily climate factors on the GNSS time series, we proposed a hybrid method and applied it in the PRVI area. We used the extended Helmert transformation method to establish the PRVI19 local reference frame, which could help avoid the bias of background global or regional tectonic movements in the GNSS time series when studying local ground deformation. The stability of the PRVI19 reference frame is approximately 0.4 mm/year and 0.5 mm/year in the horizontal and vertical directions, respectively. Also, we adopted the XGBoost algorithm and the highly stable PRVI19 local reference frame to quantitatively assess the effects of daily climate factors on the GNSS daily (24 hours) observations. Based on the 13 years of GNSS records referred to PRVI19 and climate data recorded by a nearby climate-monitoring device, we observed that the wind had the biggest impact on the GNSS time series. The results show that the average, lowest, and highest wind speeds are the first, second, and fourth-largest weights among all the climate factors. Besides, the result also

shows that the lowest temperature also greatly affects the GNSS displacements, which is the third-largest weight among all climate factors. This paper introduces a new method that can quantitatively determine the impact weights of different climate factors on the GNSS time series. Moreover, we used the model to predict the GNSS records and indicate users to explore potential deformation risk. It is hoped that this study can promote the applications of the GNSS techniques and improve the understanding of the impact of different climate factors on the GNSS monitoring time series.

## Data Availability

Data available in a publicly accessible repository. Data is available through <http://geodesy.unr.edu/> and also appreciates the climate data from NOAA National Weather Service <https://www.weather.gov/>.

## Conflicts of Interest

The authors declared that they have no conflicts of interest to this work. We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

## Authors' Contributions

Linchao Li and Hanlin Liu contributed to the conceptualization, methodology, data curation, and writing—original draft preparation. Hanlin Liu contributed to the visualization and investigation. Linchao Li contributed to the supervision. Linchao Li and Linqiang Yang contributed to the writing—reviewing and editing.

## Acknowledgments

The authors acknowledge Guoquan Wang (University of Houston) for providing the PPP solutions data quality check for this study. The first author also appreciates the UNAVCO and the Nevada Geodetic Laboratory (NGL) at the University of Nevada for sharing their GPS products with the public. Data is available through <http://geodesy.unr.edu/> and also appreciates the climate data from NOAA National Weather Service <https://www.weather.gov/>. This research was funded by the National Key R&D Program of China (grant number 2019YFB2102700). This study was supported by the research funds from China Postdoctoral Science Foundation funded project (grant number 2020M682883), the Shenzhen Science and Technology program (grant number KQTD20180412181337494), and the Foundation for Distinguished Young Talents in Higher Education of Guangdong, China (Grant No. 2019KQNCX126).

## References

- [1] R. A. Snay and T. Soler, "Continuously operating reference station (CORS): history, applications, and future enhancements," *Journal of Surveying Engineering*, vol. 134, no. 4, pp. 95–104, 2008.
- [2] UNAVCO (University NAVSTAR Consortium), "Monumentation types," 2014, <http://www.unavco.org/instrumentation/monumentation/types/types.html>.
- [3] G. Blewitt, "Carrier phase ambiguity resolution for the global positioning system applied to geodetic baselines up to 2000 km," *Journal of Geophysical Research: Solid Earth*, vol. 94, no. B8, pp. 10187–10203, 1989.
- [4] X. Xu, D. Dong, M. Fang, Y. Zhou, N. Wei, and F. Zhou, "Contributions of thermoelastic deformation to seasonal variations in GPS station position," *GPS Solutions*, vol. 21, no. 3, pp. 1265–1274, 2017.
- [5] D. Dong, P. Fang, Y. Bock, M. K. Cheng, and S. I. Miyazaki, "Anatomy of apparent seasonal variations from GPS-derived site position time series," *Journal of Geophysical Research: Solid Earth*, vol. 107, no. B4, pp. ETG 9-1–ETG 9-16, 2002.
- [6] H. Yan, W. Chen, Y. Zhu, W. Zhang, and M. Zhong, "Contributions of thermal expansion of monuments and nearby bedrock to observed GPS height changes," *Geophysical Research Letters*, vol. 36, no. 13, 2009.
- [7] H. Muneke, "Sub-daily noise in horizontal GPS kinematic time series due to thermal tilt of GPS monuments," *Journal of Geodesy*, vol. 87, no. 4, pp. 393–401, 2013.
- [8] T. A. Herring, T. I. Melbourne, M. H. Murray et al., "Plate boundary observatory and related networks: GPS data analysis methods and geodetic products," *Reviews of Geophysics*, vol. 54, no. 4, pp. 759–808, 2016.
- [9] W. Bertiger, Y. Bar-Sever, A. Dorsey et al., "GipsyX/RTGx, a new tool set for space geodetic operations and research," *Advances in Space Research*, vol. 66, no. 3, pp. 469–489, 2020.
- [10] L. Yang, G. Wang, V. Huérfano, C. G. von Hillebrandt-Andrade, J. A. Martínez-Cruzado, and H. Liu, "GPS geodetic infrastructure for natural hazards study in the Puerto Rico and Virgin Islands region," *Natural Hazards*, vol. 83, no. 1, pp. 641–665, 2016.
- [11] J. Geng and C. Shi, "Rapid initialization of real-time PPP by resolving undifferenced GPS and GLONASS ambiguities simultaneously," *Journal of Geodesy*, vol. 91, no. 4, pp. 361–374, 2017.
- [12] J. Geng, X. Li, Q. Zhao, and G. Li, "Inter-system PPP ambiguity resolution between GPS and BeiDou for rapid initialization," *Journal of Geodesy*, vol. 93, no. 3, pp. 383–398, 2019.
- [13] W. Bertiger, S. D. Desai, B. Haines et al., "Single receiver phase ambiguity resolution with GPS data," *Journal of Geodesy*, vol. 84, no. 5, pp. 327–337, 2010.
- [14] Z. Altamimi, X. Collilieux, and L. Métivier, "ITRF 2008: an improved solution of the international terrestrial reference frame," *Journal of Geodesy*, vol. 85, no. 8, pp. 457–473, 2011.
- [15] Z. Altamimi, P. Rebischung, L. Métivier, and X. Collilieux, "ITRF 2014: a new release of the International Terrestrial Reference Frame modeling nonlinear station motions," *Journal of Geophysical Research: Solid Earth*, vol. 121, no. 8, pp. 6109–6131, 2016.
- [16] G. Blewitt, C. Kreemer, W. C. Hammond, and J. M. Goldfarb, "Terrestrial reference frame NA12 for crustal deformation studies in North America," *Journal of Geodynamics*, vol. 72, pp. 11–24, 2013.
- [17] G. Blewitt and D. Lavallée, "Effect of annual signals on geodetic velocity," *Journal of Geophysical Research: Solid Earth*, vol. 107, no. B7, p. ETG 9-1, 2002.
- [18] G. Wang, H. Liu, G. S. Mattioli, M. M. Miller, K. Feaux, and J. Braun, "CARIB18: a stable geodetic reference frame for

- geological hazard monitoring in the Caribbean region,” *Remote Sensing*, vol. 11, no. 6, p. 680, 2019.
- [19] C. Pearson and R. Snay, “Introducing HTDP 3.1 to transform coordinates across time and spatial reference frames,” *GPS Solutions*, vol. 17, no. 1, pp. 1–15, 2013.
- [20] E. S. Abdolkarimi, G. Abaei, and M. R. Mosavi, “A wavelet-extreme learning machine for low-cost INS/GPS navigation system in high-speed applications,” *GPS Solutions*, vol. 22, no. 1, p. 15, 2018.
- [21] K. J. Bergen, P. A. Johnson, V. Maarten, and G. C. Beroza, “Machine learning for data-driven discovery in solid Earth geoscience,” *Science*, vol. 363, no. 6433, article eaau0323, 2019.
- [22] B. Rouet-Leduc, C. Hulbert, and P. A. Johnson, “Continuous chatter of the Cascadia subduction zone revealed by machine learning,” *Nature Geoscience*, vol. 12, no. 1, pp. 75–79, 2019.
- [23] B. J. Phrampus, T. R. Lee, and W. T. Wood, “A global probabilistic prediction of cold seeps and associated seafloor fluid expulsion anomalies (SEAFLEAs),” *Geochemistry, Geophysics, Geosystems*, vol. 21, no. 1, 2020.
- [24] G. J. Anderson and D. D. Lucas, “Machine learning predictions of a multiresolution climate model ensemble,” *Geophysical Research Letters*, vol. 45, no. 9, pp. 4273–4280, 2018.
- [25] C. X. Ren, O. Dorostkar, B. Rouet-Leduc et al., “Machine learning reveals the state of intermittent frictional dynamics in a sheared granular fault,” *Geophysical Research Letters*, vol. 46, no. 13, pp. 7395–7403, 2019.
- [26] T. Soler and R. A. Snay, “Transforming positions and velocities between the International Terrestrial Reference Frame of 2000 and North American Datum of 1983,” *Journal of Surveying Engineering*, vol. 130, no. 2, pp. 49–55, 2004.
- [27] T. Chen and C. Guestrin, “Xgboost: a scalable tree boosting system,” in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785–794, New York, NY, USA, 2016.
- [28] L. Breiman and J. H. Friedman, “Predicting multivariate responses in multiple linear regression,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 59, no. 1, pp. 3–54, 1997.
- [29] R. Ihaka and R. Gentleman, “R: a language for data analysis and graphics,” *Journal of Computational and Graphical Statistics*, vol. 5, no. 3, pp. 299–314, 1996.
- [30] G. Wang and T. Soler, “OPUS for horizontal subcentimeter-accuracy landslide monitoring: case study in the Puerto Rico and Virgin Islands region,” *Journal of Surveying Engineering*, vol. 138, no. 3, pp. 143–153, 2012.
- [31] H. Liu and G. Wang, “Relative motion between St. Croix and the Puerto Rico-Northern Virgin Islands block derived from continuous GPS observations (1995–2014),” *International Journal of Geophysics*, vol. 2015, Article ID 915753, 9 pages, 2015.
- [32] U. S. ten Brink and A. M. López-Venegas, “Plate interaction in the NE Caribbean subduction zone from continuous GPS observations,” *Geophysical Research Letters*, vol. 39, no. 10, 2012.
- [33] B. Benford, C. DeMets, B. Tikoff, P. Williams, L. Brown, and M. Wiggins-Grandison, “Seismic hazard along the southern boundary of the Gônavé microplate: block modelling of GPS velocities from Jamaica and nearby islands, northern Caribbean,” *Geophysical Journal International*, vol. 190, no. 1, pp. 59–74, 2012.
- [34] P. Mann, E. Calais, J. C. Ruegg, C. DeMets, P. E. Jansma, and G. S. Mattioli, “Oblique collision in the northeastern Caribbean from GPS measurements and geological observations,” *Tectonics*, vol. 21, no. 6, pp. 7-1–7-26, 2002.
- [35] D. J. Murphy, M. H. Hall, C. A. Hall, G. M. Heisler, S. V. Stehman, and C. Anselmi-Molina, “The relationship between land cover and the urban heat island in northeastern Puerto Rico,” *International Journal of Climatology*, vol. 31, no. 8, pp. 1222–1239, 2011.
- [36] A. E. Lugo, O. Ramos, and C. Rodriguez, *The Río Piedras Watershed and Its Surrounding Environment. FS-980*, vol. 46, US Department of Agriculture, Forest Service, International Institute of Tropical Forestry, 2011.
- [37] J. Nyberg, B. A. Malmgren, A. Winter, M. R. Jury, K. H. Kilbourne, and T. M. Quinn, “Low Atlantic hurricane activity in the 1970s and 1980s compared to the past 270 years,” *Nature*, vol. 447, no. 7145, pp. 698–701, 2007.
- [38] W. W. V. Sweet, R. Kopp, C. P. Weaver et al., “Global and regional sea level rise scenarios for the United States,” *NOAA technical report NOS CO-OPS*, vol. 83, 2017.
- [39] Y. Li, X. Chang, K. Yu, S. Wang, and J. Li, “Estimation of snow depth using pseudorange and carrier phase observations of GNSS single-frequency signal,” *GPS Solutions*, vol. 23, no. 4, p. 118, 2019.
- [40] D. Firuzabadi and R. W. King, “GPS precision as a function of session duration and reference frame using multi-point software,” *GPS Solutions*, vol. 16, no. 2, pp. 191–196, 2012.
- [41] G. Wang, “GPS landslide monitoring: single base vs. network solutions: a case study based on the Puerto Rico and Virgin Islands permanent GPS network,” *Journal of Geodetic Science*, vol. 1, pp. 191–203, 2011.
- [42] G. Blewitt, C. Kreemer, W. C. Hammond, and J. Gazeaux, “MIDAS robust trend estimator for accurate GPS station velocities without step detection,” *Journal of Geophysical Research*, vol. 121, no. 3, pp. 2054–2068, 2016.
- [43] C. S. Prentice and P. Mann, “Paleoseismic study of the South Lajas fault: first documentation of an onshore Holocene fault in Puerto Rico. Active tectonics and seismic hazards of Puerto Rico, the Virgin Islands, and offshore areas,” *Geological Society of America Special Paper*, vol. 385, pp. 215–222, 2005.
- [44] Y. Li, C. Xu, L. Yi, and R. Fang, “A data-driven approach for denoising GNSS position time series,” *Journal of Geodesy*, vol. 92, no. 8, pp. 905–922, 2018.
- [45] L. Li, H. Zhou, H. Liu, C. Zhang, and J. Liu, “A hybrid method coupling empirical mode decomposition and a long short-term memory network to predict missing measured signal data of SHM systems,” *Structural Health Monitoring*, p. 1475921720932813, 2020.
- [46] Y. Lin, L. Li, H. Jing, B. Ran, and D. Sun, “Automated traffic incident detection with a smaller dataset based on generative adversarial networks,” *Accident Analysis & Prevention*, vol. 144, article 105628, 2020.
- [47] J. Gazeaux, S. Williams, M. King et al., “Detecting offsets in GPS time series: first results from the detection of offsets in GPS experiment,” *Journal of Geophysical Research: Solid Earth*, vol. 118, no. 5, pp. 2397–2407, 2013.
- [48] J. Langbein, “Noise in GPS displacement measurements from Southern California and Southern Nevada,” *Journal of Geophysical Research: Solid Earth*, vol. 113, no. B5, 2008.
- [49] C. Li, S. Huang, Q. Chen et al., “Quantitative evaluation of environmental loading induced displacement products for correcting GNSS time series in CMONOC,” *Remote Sensing*, vol. 12, no. 4, p. 594, 2020.